Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the RuBisCO chaperone RbcX from the thermophilic cyanobacterium *Thermosynechococcus elongatus*

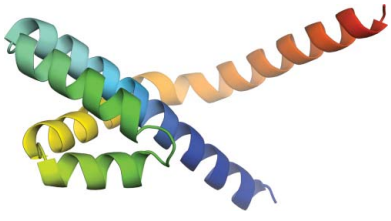
Miroslaw Tarnawski,[a] Szymon Krzywda,[b] Wojciech Bialek,[a] Mariusz Jaskolski[b,c]* and Andrzej Szczepaniak[a]*

[a]Department of Biophysics, Faculty of Biotechnology, University of Wroclaw, Przybyszewskiego 63/77, 51-148 Wroclaw, Poland, [b]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Grunwaldzka 6, 60-780 Poznan, Poland, and [c]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 60-704 Poznan, Poland

Correspondence e-mail: mariuszj@amu.edu.pl, aszczep@ibmb.uni.wroc.pl

Received 25 February 2011
Accepted 17 May 2011

PDB Reference: RbcX, 3q20.



The crystal structure of TeRbcX, a RuBisCO assembly chaperone from the cyanobacterium *Thermosynechococcus elongatus*, a thermophilic organism, has been determined at 1.7 Å resolution. TeRbcX has an unusual cysteine residue at position 103 that is not found in RbcX proteins from mesophilic organisms. Unlike wild-type TeRbcX, a mutant protein with Cys103 replaced by Ala (TeRbcX-C103A) could be readily crystallized. The structure revealed that the overall fold of the TeRbcX homodimer is similar to those of previously crystallized RbcX proteins. Normal-mode analysis suggested that TeRbcX might adopt an open or closed conformation through a hinge movement pivoted on a kink in two long α4 helices. This type of conformational transition is presumably connected to RbcL (the large RuBisCO subunit) binding during the chaperone function of the RuBisCO assembly.

1. Introduction

Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), an enzyme that accounts for 30–50% of the total soluble protein in the chloroplast, is the most abundant protein on Earth (Dhingra *et al.*, 2004). It fixes inorganic carbon dioxide, which enters the biosphere, and is therefore regarded as one of the most important proteins (Ellis, 1979).

RuBisCO form I, which is found in plants, green algae and cyanobacteria, is composed of eight large subunits (RbcL) and eight small subunits (RbcS), which are encoded by two genes, *rbcL* and *rbcS*, respectively. Chaperone factors enabling polypeptide-chain folding, as well as specific assembly chaperones (Ellis, 2006), are required for the RuBisCO assembly process (Andersson & Taylor, 2003; Tabita, 1999). Interestingly, in some cyanobacteria an extra gene, *rbcX*, is located between the *rbcL* and *rbcS* genes. Owing to its localization inside the RuBisCO operon, it has been suggested that the RbcX protein may act as a chaperone during RuBisCO assembly. Indeed, co-expression of *rbcX* with the *rbcL* and *rbcS* genes enhanced the formation of active RuBisCO from *Anabaena* sp. CA (Li & Tabita, 1997), *Synechococcus* sp. PCC 7002 (Onizuka *et al.*, 2004) and *Thermosynechococcus elongatus* (Tarnawski, Gubernator *et al.*, 2008). Thus, it seems that the presence of the *rbcX* gene within or outside the RuBisCO operon is closely related to its significance in RuBisCO assembly. In the case of *Synechococcus* sp. PCC 7942, in which the *rbcX* gene is located about 100 kb away from the *rbcLS* operon, insertional inactivation of the *rbcX* gene in the cyanobacterium showed no perturbation of the growth rate or RuBisCO content and activity (Emlyn-Jones *et al.*, 2006), in apparent contrast to the situation in *Synechococcus* sp. PCC 7002 (Onizuka *et al.*, 2004).

The RbcX proteins show a high level of sequential and structural conservation, invariably forming dimeric assemblies, which suggests that their mechanism of action must also be similar. Recently, it has been shown that folding of the RbcL subunit by the GroEL/GroES chaperonin is tightly coupled to an assembly process mediated by the RbcX chaperone. Further structural analysis of an $RbcL_8$–$RbcX_8$ assembly intermediate by cryo-electron microscopy provided a detailed insight into the mechanism of RuBisCO holoenzyme

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Crystal dimensions (mm)	1.3 × 0.5 × 0.5
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 45.78, b = 67.73, c = 94.54
Mosaicity (°)	0.6
Protein molecules per asymmetric unit	2
Matthews coefficient V_M (Å^3 Da^{-1})	2.51
Solvent content (%)	51.1
Radiation source	BESSY BL14.1
Wavelength (Å)	0.9184
Temperature (K)	100
Detector	MX-225 CCD
Resolution range (Å)	50–1.71 (1.76–1.71)
No. of observed reflections	161694
No. of unique reflections	31245
Multiplicity	5.2 (2.9)
Completeness (%)	96.2 (71.8)
R_{merge}† (%)	4.7 (43.0)
$\langle I/\sigma(I)\rangle$	29.9 (2.2)
Refinement	
No. of reflections in working set	29886
No. of reflections in test set	1280
Resolution range (Å)	41.20–1.71
No. of non-H atoms	2075
Protein	1843
Solvent	232
R‡ (%)	18.2
R_{free}§ (%)	21.1
R.m.s. deviations from ideal	
Bonds (Å)	0.018
Angles (°)	1.59
Mean B factor (Å^2)	45.8
Protein	44.9
Solvent	52.8
Ramachandran statistics (%)	
Most favoured regions	92.5
Additional allowed regions	7.5

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement of the intensity of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$ for all reflections, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. § R_{free} is calculated analogously to R for test reflections which were randomly selected and excluded from the refinement.

formation (Liu *et al.*, 2010). In particular, Saschenbrecker *et al.* (2007) showed that RbcX acts as a specific RuBisCO assembly chaperone to promote the formation of the $(RbcL_2)_4$ core complex. The authors presented evidence that a specific C-terminal fragment of RbcL (E_{459}IKFEFD$_{465}$) binds in a central groove of the RbcX molecule. This fragment adopts an extended conformation, with the polar side chains facing outwards and two large hydrophobic side chains of phenylalanine extending into cavities formed by the side chains of Thr13, Tyr17, Tyr20, Ile50 and Gln51.

The crystal structures of RbcX from three mesophilic cyanobacteria, namely *Synechococcus* sp. PCC 7002 (Saschenbrecker *et al.*, 2007), *Anabaena* sp. CA (Saschenbrecker *et al.*, 2007) and *Synechocystis* sp. PCC 6803 (Tanaka *et al.*, 2007), have been solved. Regardless of the species, each RbcX protein is composed almost exclusively of α-helices, which form an unusual four-helix bundle. In addition, all known RbcX proteins exist as homodimers.

The main goal of the present project is to elucidate the structure of the RbcX protein from the thermophilic cyanobacterium *T. elongatus* (TeRbcX). Amino-acid sequence analysis of the RbcX proteins highlighted the presence in the TeRbcX sequence of an unusual single Cys residue at position 103 that is only found in two other thermophilic RbcX proteins: those from *Phormidium laminosum* and *T. vulcanus* (Tarnawski, Krzywda *et al.*, 2008). This cysteine residue maps onto a surface Arg of the known RbcX crystal structures (PDB entries 2peq, 2peo and 2py8; Saschenbrecker *et al.*, 2007; Tanaka *et al.*, 2007) but at a position that precludes its biological role. However, recombinant TeRbcX could not be crystallized owing to uncontrolled aggregation *via* intermolecular disulfide-bond formation (Tarnawski, Krzywda *et al.*, 2008). Recently, we have constructed and crystallized two TeRbcX variants, which form the basis for the present structural analysis (Tarnawski, Krzywda *et al.*, 2008). Whereas Cys103 prevented the crystallization of wild-type TeRbcX even in the presence of reducing agents, the C103A (TeRbcX-C103A) and C103R mutants did not form higher-order aggregates and could be readily crystallized. Importantly, it was shown that TeRbcX-C103A could still form homodimers analogously to the previously crystallized native RbcX proteins from other species.

Here, we report the crystal structure of TeRbcX-C103A, which is the first example of an RbcX structure from a thermophilic organism. Since the presence of the native Cys103 hampered crystallization, we describe the structure of TeRbcX containing the site-specific mutation C103A. TeRbcX-C103A could be obtained in two polymorphic modifications and crystals of the orthorhombic form diffracted X-rays to high resolution (1.7 Å). This is the highest resolution structure of any RbcX protein available to date.

2. Experimental procedures

2.1. Mutagenesis, expression and purification

The coding sequence of the *rbcX* gene from *T. elongatus* was cloned into pET-16b vector (Novagen). The Cys103 residue was mutated to Ala (C103A) by means of the QuikChange (Stratagene) method. Recombinant TeRbcX protein was expressed at 310 K in *Escherichia coli* BL21 (DE3) cells and purified by ammonium sulfate fractionation, thermal fractionation, ion-exchange and size-exclusion chromatography as reported previously (Tarnawski, Krzywda *et al.*, 2008).

2.2. Protein crystallization

Crystallization of TeRbcX mutant proteins using the hanging-drop vapour-diffusion method was performed as described previously (Tarnawski, Krzywda *et al.*, 2008). High-quality single crystals of TeRbcX-C103A were obtained using a reservoir solution consisting of 0.1 *M* sodium acetate pH 4.5, 1.0 *M* 1,6-hexanediol and 0.01 *M* $CoCl_2$.

2.3. X-ray diffraction data collection

A single crystal was harvested in a cryoloop; it was briefly soaked in a cryoprotectant solution consisting of the reservoir solution with 1,6-hexanediol added to a final concentration of 1.8 *M* prior to flash-vitrification in a stream of cold nitrogen gas. X-ray diffraction data were collected using synchrotron radiation on the BL14.1 beamline at the BESSY synchrotron (Berlin). The diffraction data were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The final statistics are summarized in Table 1.

2.4. Structure solution and refinement

The structure was solved by molecular replacement by application of the program *REMO* (Burla *et al.*, 2007; Caliandro *et al.*, 2006) using previously reported diffraction data (data set A-1) in the 50–2 Å resolution range (Tarnawski, Krzywda *et al.*, 2008) and the RbcX model from *Synechococcus* sp. PCC 7002 (PDB entry 2peq; Saschenbrecker *et al.*, 2007). Subsequently, the solution was confirmed in *MOLREP* (Vagin & Teplyakov, 2010) using a new data set with a resolution range of 50–1.7 Å. Model rebuilding and refinement

were carried out using *Coot* (Emsley & Cowtan, 2004) and *REFMAC*5 (Murshudov *et al.*, 2011). TLS parameters calculated by the *TLS Motion Determination* server (Painter & Merritt, 2006*a*,*b*) and bulk-solvent correction were applied towards the end of the refinement. The final structure of TeRbcX-C103A is characterized by an *R* factor of 18.2% and an R_{free} of 21.1% (Table 1).

Atomic coordinates and structure factors have been deposited in the PDB with accession code 3q20. Structural figures were generated with *PyMOL* (DeLano, 2002).

2.5. Normal-mode analysis

Putative conformational changes of TeRbcX were detected by *elNémo*, the web interface to *The Elastic Network Model*, a fast and simple tool for the computation of low-frequency normal modes of a protein (Suhre & Sanejouand, 2004). The following program settings were applied: NMODES = 5, DQMIN = −100, DQMAX = 0, DQSTEP = 20.

3. Results and discussion

3.1. Overall structure

The overall architecture of the RbcX protein from *T. elongatus* is similar to other known cyanobacterial RbcX structures. TeRbcX is nearly entirely α-helical (88.2%). The fold consists of four α-helices (α1–α4) per monomer which form a characteristic helix bundle (Fig. 1*a*). The N-terminal helix α1 consists of residues Val3–Thr33 and is followed by helix α2 (Pro35–Glu48) in an antiparallel orientation. The α1–α2 turn is formed by only one residue (Asp34). The junction between helices α3 (Gly53–Glu63) and α4 (Pro65–Gln111) is arranged in a similar way. The relatively short helices α2 and α3 are connected by a loop-like linker, the $S_{49}IQD_{52}$ loop, with the Ile50-Gln51-Asp52 segment in a reverse γ-turn conformation. The longest C-terminal helix α4 is bent in the middle and the residues Ala83-Asp84-Tyr85 form a 3_{10}-helical turn. The C-terminal helix α4 extends away from the compact core of the bundle, so that its C-terminal half (beyond the kink) has no packing interactions with the core.

The final model consists of two polypeptide chains in the asymmetric unit that form a boomerang-like homodimer (Figs. 1*b* and 1*c*). The crystal structure is consistent with the expectation of a dimer as the biological unit. Electron-density maps show clear tracing of residues 1–114 of polypeptide chain *A* and residues 2–107 of chain *B*. Additionally, two partially occupied molecules of 1,6-hexanediol and two partially occupied acetate ions are present in the asymmetric unit.

The quality of the electron density is poorer in the Thr33–Asp52 fragment of molecule *A*, where it is continuous but lacks the main- and side-chain details that are clearly visible for the rest of the structure. In particular, electron density is completely missing for the side chains of residues Leu39, Gln43, Ser46, Gln47 and Ser49. This fragment covers residues of helix α2 (residues 33–48) and the loop between helices α2 and α3 (residues 49–52), which are implicated in

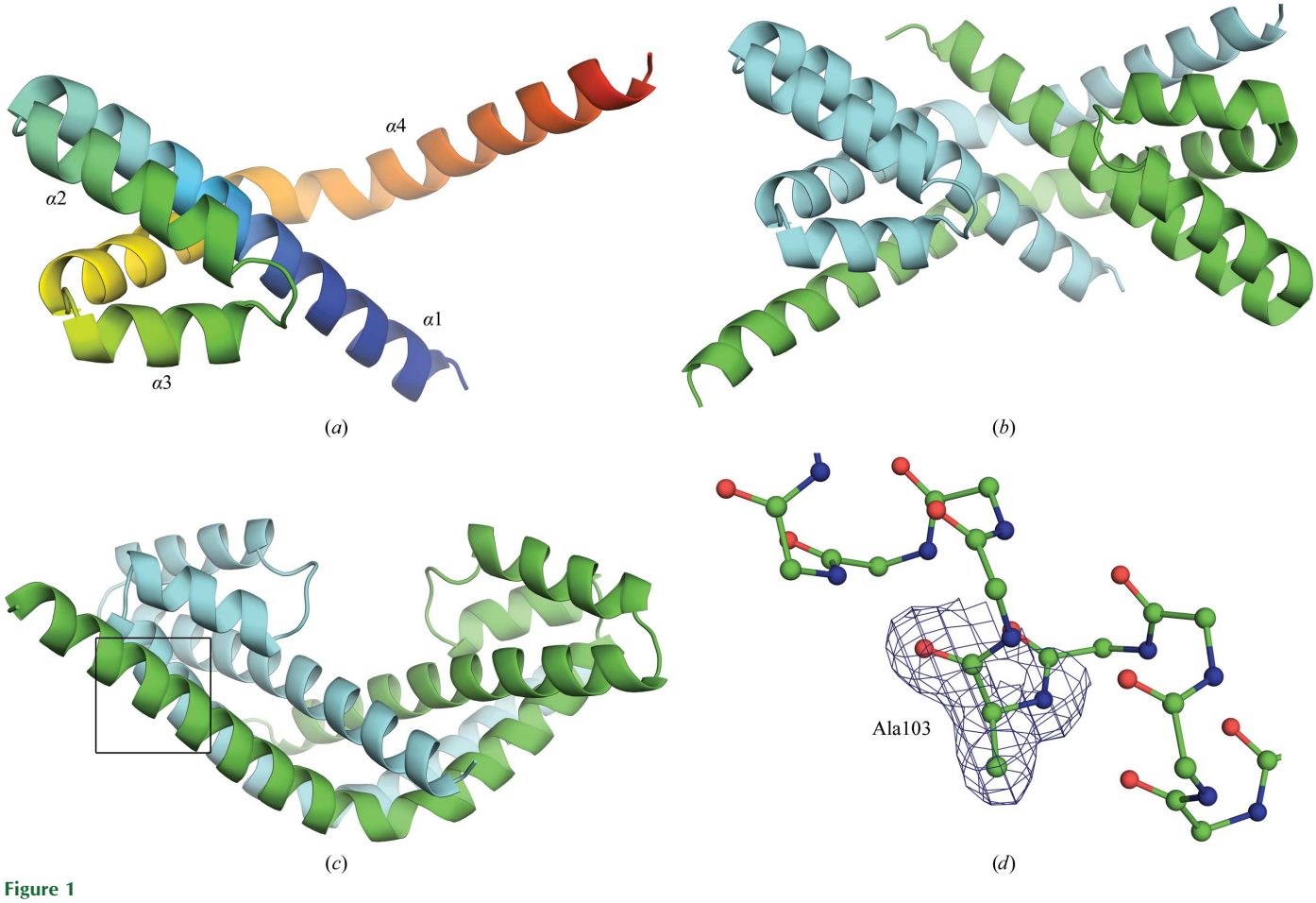
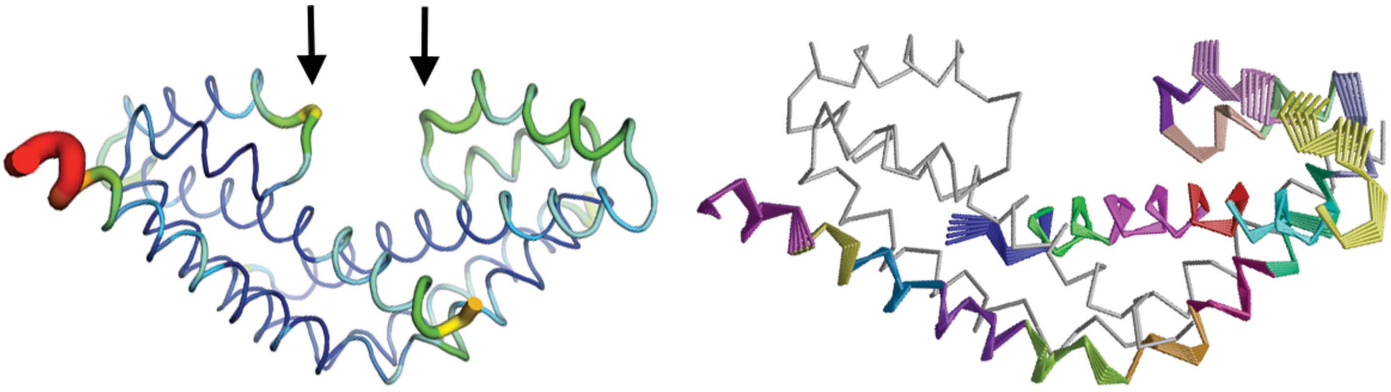


Figure 1
Three-dimensional structure of the RbcX protein from *T. elongatus*. (*a*) The four-helix-bundle fold of a protomer (chain *A*) in rainbow colours from blue (N-terminus) to red (C-terminus). (*b*, *c*) The TeRbcX dimer, with chain *A* shown in green and chain *B* shown in cyan: (*b*) top view, (*c*) side view. (*d*) Ball-and-stick representation of the region boxed in (*c*). The mutated residue C103A of chain *A* is shown as $2F_o - F_c$ electron density contoured at 1.0σ.



(*a*) (*b*)

Figure 2
Flexibility of the TeRbcX structure. (*a*) *B* factors of backbone C^{α} atoms of the TeRbcX dimer shown using a colour gradient (blue, low *B* values; red, high *B* values) and a chain diameter proportional to *B*. The arrows indicate the $S_{49}IQD_{52}$ loop between helices α2 and α3. (*b*) TLS analysis of TeRbcX chain *A* using 19 partition groups shown in different colours. Chain *B* (single segment) is shown in grey. (*b*) was generated with the *TLSMD* server (Painter & Merritt, 2006*a*,*b*).

important contacts with the C-terminal peptide of RbcL (Saschenbrecker *et al.*, 2007). Interestingly, a very similar picture is observed for the corresponding fragments of the 2peo structure, which are the most mobile regions of the protein (Saschenbrecker *et al.*, 2007). The corresponding fragment of the present molecule *B* is well ordered and the electron density is very clear. Stabilization is achieved through several hydrogen bonds to a crystallographic copy of molecule *B*. Four of the five side chains that are disordered in molecule *A* contribute to the stabilization of molecule *B* by forming six protein–protein hydrogen bonds and three hydrogen bonds to water molecules.

As in the other RbcX structures, residues 115–126 are not visible in the electron-density maps. It has been shown that the removal of these poorly conserved C-terminal residues does not affect RbcX activity in RuBisCO assembly (Saschenbrecker *et al.*, 2007).

The two four-helix bundles of the dimeric structure occupy opposite ends of the boomerang-shaped molecule. The α4 helices of the two protomers in a dimer are arranged in an antiparallel fashion. The α4 helix is so long that its C-terminal half reaches the opposite bundle, creating in reality a five-helix bundle with one noncovalent intermolecular component.

Helices α1 and α2 of both protomers outline a diagonal groove on the surface of the molecule which is implicated in binding of the C-terminal peptide of RbcL during RuBisCO assembly. The loops between helices α2 and α3 form a well defined clamp-like structure constraining access to the central groove, which is filled with several water molecules. The cleft leading to the groove is less than 10 Å wide, but the internal diameter of the groove (about 15 Å) is sufficient for binding of a polypeptide chain in an extended conformation.

Within the dimer, the protomers are related by a pseudo-twofold axis (179.3°) that is approximately parallel to the crystallographic **b** axis. The two polypeptide chains forming the dimer are structurally very similar. The r.m.s.d. value for their backbone C^{α} atoms is 0.45 Å, with the largest deviations being located in the region of helix α2. The atoms of the $S_{49}IQD_{52}$ loops are characterized by relatively high displacement parameters and must therefore be considered as flexible elements of the structure (Fig. 2*a*). It is of note that the side chains of the Ile50 residues in the two polypeptide chains adopt different conformations, thus breaking the twofold symmetry of the dimer. Also, rigid-body TLS analysis (Painter & Merritt, 2006*a*,*b*) identified helices α2 and α3 as mobile elements of the structure (Fig. 2*b*).

A cluster of six glutamine residues (95, 96, 100, 101, 104 and 105), the side chains of which adopt alternative conformations, is located within the long C-terminal helix α4. The role of this glutamine tract is not clear. The side chains of four of these glutamine residues (95, 96, 100 and 104) are clearly solvent-exposed, with average relative residue accessible surface areas of about 50%.

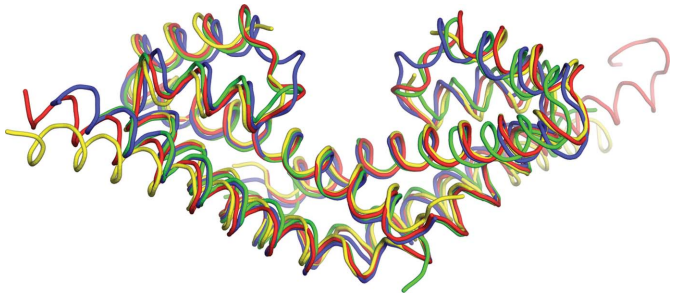


Figure 3
Structural alignment of RbcX dimers from *T. elongatus* (this work; PDB entry 3q20, chains *AB*, blue), *Synechococcus* sp. PCC 7002 (2peq, chains *AB*, green), *Anabaena* sp. CA (2peo, chains *AB*, yellow) and *Synechocystis* sp. PCC 6803 (2py8, chains *BC*, red). The superpositions were calculated in *SSM* (Krissinel & Henrick, 2004) using C^{α} coordinates and the structure of TeRbcX as the target.

3.2. Structural similarities

All cyanobacterial RbcX proteins crystallized to date are structurally similar and share a common fold (Table 2). Most of the differences in the conformation of the polypeptide chain are observed in the region of helices α2 and α3, including the loop-like linker (Fig. 3). In addition, electron-density maps showed that, apart from the highly disordered C-terminus, the $S_{49}IQD_{52}$ loop region is the least well ordered fragment of the polypeptide chain (see above). It is of note that in the model of RbcX from *Anabaena* sp. CA (PDB entry 2peo; Saschenbrecker *et al.*, 2007), the residues Lys48-Val49 (corresponding to the $K_{48}VQD_{51}$ fragment of the present protein) are missing altogether owing to an absence of detectable electron density.

Based on multiple alignment of 21 cyanobacterial RbcX sequences, the *ConSeq* (Berezin *et al.*, 2004) and *ConSurf* (Glaser *et al.*, 2003; Landau *et al.*, 2005) servers identified a number of conserved residues in the TeRbcX structure, mostly in the central region (Met1, Thr10, Tyr17, Thr19, Gln21, Val74, Arg75, Glu76, Pro87 and Asn98) of the molecule, but also in the peripheral 'corner' regions (Glu32, Thr33, Pro35, Arg70, Arg102, Glu107 and Arg108) located at opposite ends

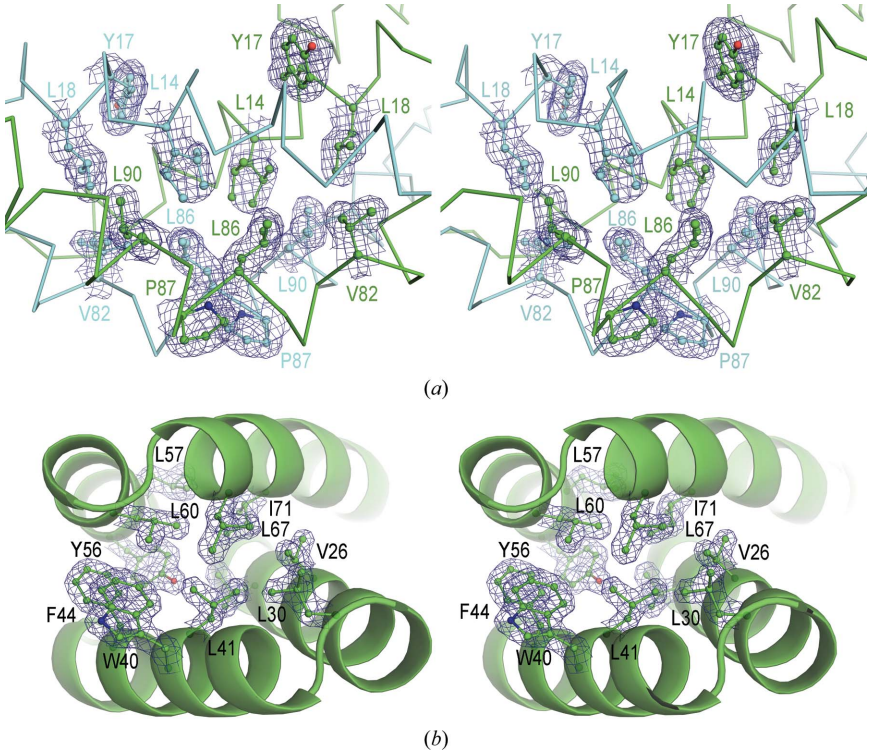


Figure 4
Stereoviews of the hydrophobic interactions in the TeRbcX structure. (*a*) The central core region at the dimer interface formed by helices α1 and α4 of both protomers (chains coloured as in Fig. 1*b*): $2F_o - F_c$ electron density contoured at 1.0σ. (*b*) The core of the four-helix bundle of chain *A* in a $2F_o - F_c$ electron-density map contoured at 2.0σ.

of the dimer. The conserved residues forming the central core region are responsible for the stabilization of the dimeric RbcX fold and those lining the bottom of the central groove are involved in binding of the C-terminal RbcL peptide, as shown by Saschenbrecker *et al.* (2007). In addition, some residues in the peripheral regions have been shown to be functionally critical (Saschenbrecker *et al.*, 2007) and their importance in surface interactions with the N-terminal domain of the RbcL subunit has been established (Liu *et al.*, 2010).

3.3. Water structure

The electron-density maps allowed the identification of 208 water molecules, including 109 with full occupancy. The present TeRbcX structure is distinguished from the other RbcX structures by a high number of water molecules per dimer. Comparison of the solvent contents and Matthews coefficients of the present structure (51.1% and 2.51 Å^3 Da^{-1}, respectively) and 2peq (30.8% and 1.78 Å^3 Da^{-1}, respectively), solved at resolutions of 1.71 and 1.90 Å, respectively, showed a higher water content in the TeRbcX crystal structure. It is of note that in the TeRbcX structure water molecules form well ordered clusters or water chains around the side chains of certain residues, *e.g.* Arg102(*A*) and Gln29(*B*), at the protein surface (Supplementary Fig. 1[1]).

3.4. Cys103 of RbcX from *T. elongatus*

In the sequence of wild-type TeRbcX a single cysteine residue is present at position 103. Since its presence hampered our initial crystallization trials (Tarnawski, Krzywda *et al.*, 2008), we substituted this residue with alanine, creating the TeRbcX-C103A mutant, which readily crystallized. The side chain of the mutated C103A residue, which is located on the surface of the C-terminal part of helix α4, is clearly visible in electron density (Fig. 1*d*).

The presence of Cys103 in the TeRbcX protein is rather puzzling. It might appear that it would be preferable to avoid a thermolabile and helix-disfavouring Cys residue in a protein such as TeRbcX, but it is remarkable that cysteine is also present in this position in two other thermophilic RbcX proteins, namely those from *P. laminosum* and *T. vulcanus*. The occurrence of cysteine at position 103 is directly connected with the observed TeRbcX aggregation *via* intermolecular S—S bond formation. Its presence must be justified by its function, but since the role of Cys103 in thermophilic RbcX proteins is currently unknown this interesting aspect needs to be investigated further.

3.5. Structure stabilization

The structure of RbcX is mainly stabilized by hydrophobic interactions. The dimer interface is predominantly uncharged and hydrophobic. The central region of the molecule is formed by symmetrically disposed residues from helices α1 and α4 of both protomers (Fig. 4*a*). The highly hydrophobic core of the four-helix bundle is formed by residues from all four helices (Fig. 4*b*). A tight packing of conserved hydrophobic residues (Leu14, Tyr17, Leu18, Val82, Leu86, Pro87, Met89 and Leu90) at the central junction between the two protomers is necessary for integrity of the dimer (Saschenbrecker *et al.*, 2007).

The dimer stability appears to be enhanced by polar interactions around Arg75, which forms hydrogen bonds to Thr19, Glu54 and Asn98 from the complementary protomer (Fig. 5*a*). A similar network has been described as contributing to dimer stability of the 2peq structure (Saschenbrecker *et al.*, 2007). The system of inter-

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: WD5153).

actions formed by Arg75 of TeRbcX is much stronger, as it involves a salt bridge with Glu54, which replaces the contact with Ser97 reported for 2peq (Table 3). In the 2peq structure the Asp54 side chain is too short for effective interaction with Arg75. In TeRbcX, the Ser97 position of 2peq is occupied by Ala, which makes no important contacts with the neighbouring residues. Other polar contacts are formed by residues from the opposing helices α4 with the participation of an extended network of water molecules (Fig. 5*b*). It should be noted that in the TeRbcX structure an additional salt bridge is formed involving the Asp66–Arg37 ion pair. It staples the helices α2 and α4 within the helix bundle, thus increasing its integrity.

The organization of the structural elements in the RbcX dimer resembles the situation arising upon three-dimensional domain swapping (Bennett *et al.*, 1995; Gronenborn, 2009; Jaskólski, 2001). However, in its *bona fide* meaning the term 'three-dimensional domain swapping' refers to cases in which the monomeric fold is recreated from elements contributed by two or more protein chains. Since a monomeric RbcX protein, in which helix α4 would fold upon itself (*i.e.* with the C-terminal half packed against the N-terminal half), is not known to exist, the term 'three-dimensional domain swapping' can only be used in a figurative sense in this case.

3.6. RbcX hinge regions

Upon binding to another molecule, protein motions that have functional importance may occur, including in particular motions of flexible hinge elements. Using the *HingeProt* server employing elastic network models (Emekli *et al.*, 2008), we have identified a likely hinge region in all known RbcX structures. The hinge residue is Tyr85 of TeRbcX; Phe is present at the corresponding position in the other proteins. Similar results were obtained using the *StoneHinge* approach (Keating *et al.*, 2009), which identified hinge regions around residues Leu86-Pro87 (chain *A*) and Asp84-Tyr85 (chain *B*) together with two hinge domains.

The dynamics and flexibility of TeRbcX was further examined by normal-mode analysis carried out using the *ElNémo* server (Suhre & Sanejouand, 2004). We found that one dominant mode described most of the conformational changes of the hinge motion. This hinge movement follows the first nontrivial lowest frequency normal mode (mode 7) with the perturbation set to 100 arbitrary units from the crystallographic state. Mode 7 has a collective character, as expected for functional conformational changes. The movement at the hinge element results in opening and closing of the binding groove of

Table 2
Sequence identity (%, upper triangle) and backbone r.m.s.d. values (Å, lower triangle) for cyanobacterial RbcX proteins from *T. elongatus* (PDB entry 3q20; this work), *Synechococcus* sp. PCC 7002 (2peq), *Anabaena* sp. CA (2peo) and *Synechocystis* sp. PCC 8603 (2py8).

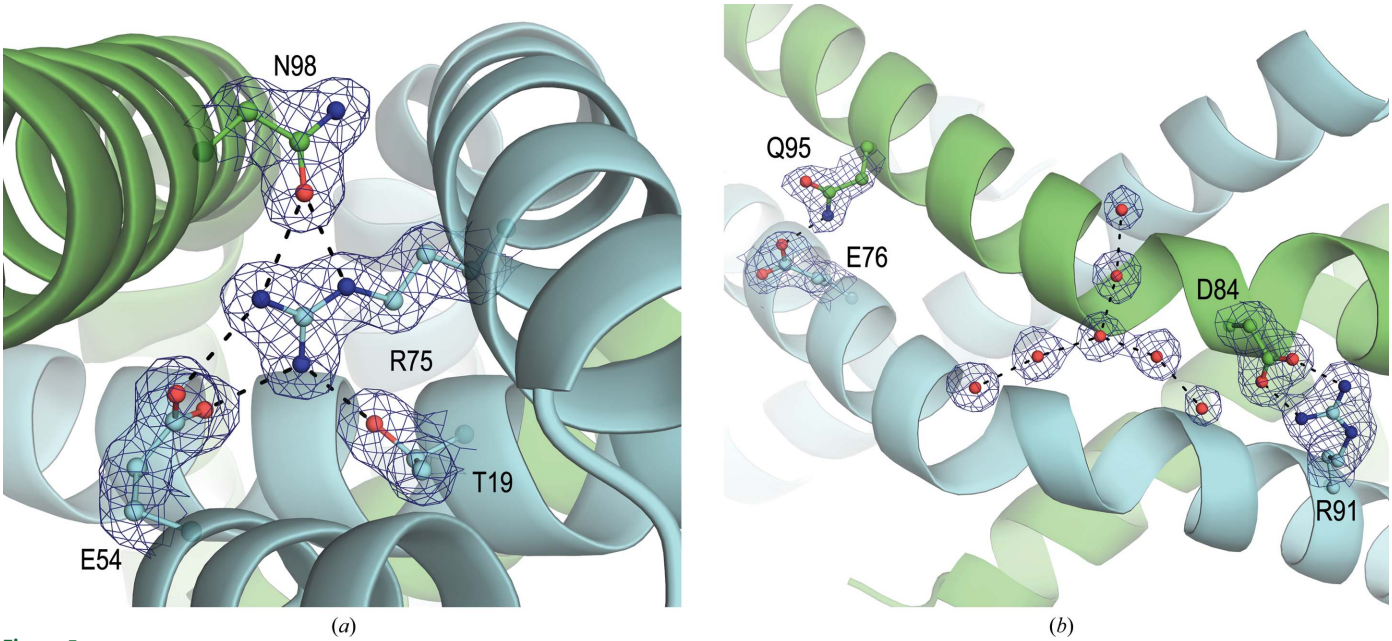
Values were obtained with the *SSM* program (Krissinel & Henrick, 2004). The resolution of each crystal structure determination is specified in parentheses.

Structure	3q20, chain *AB*	2peq, chain *AB*	2peo, chain *AB*	2py8, chain *BC*
3q20, chain *AB* (1.71 Å)	—	49.3	63.5	44.7
2peq, chain *AB* (1.90 Å)	1.39	—	51.6	51.3
2peo, chain *AB* (2.50 Å)	1.69	1.64	—	46.8
2py8, chain *BC* (2.45 Å)	1.40	1.42	1.55	—

Table 3
Hydrogen bonding at the Arg75 side chain in the 3q20 and 2peq structures as calculated by *WHAT IF* (Hooft *et al.*, 1996).

Distances are given in Å. Unrealistically short interatomic distances are indicated in italics.

	Chain *A*						Chain *B*					
3q20	Arg75	N^{ε}	Asn98	$O^{\delta1}$	*B*	2.86	Arg75	N^{ε}	Asn98	$O^{\delta1}$	*A*	2.83
		$N^{\eta1}$	Glu54	$O^{\varepsilon2}$	*A*	2.56		$N^{\eta1}$	Glu54	$O^{\varepsilon2}$	*B*	2.72
		$N^{\eta1}$	Thr19	$O^{\gamma1}$	*A*	2.86		$N^{\eta1}$	Thr19	$O^{\gamma1}$	*B*	2.83
		$N^{\eta2}$	Glu54	$O^{\varepsilon1}$	*A*	3.01		$N^{\eta2}$	Glu54	$O^{\varepsilon1}$	*B*	3.14
		$N^{\eta2}$	Asn98	$O^{\delta1}$	*B*	2.86		$N^{\eta2}$	Asn98	$O^{\delta1}$	*A*	2.86
2peq	Arg75	N^{ε}	Thr19	$O^{\gamma1}$	*A*	3.20	Arg75	N^{ε}	Asn98	$O^{\delta1}$	*A*	2.88
		$N^{\eta1}$	Ser97	O^{γ}	*B*	3.40		$N^{\eta1}$	Wat20	O	—	2.68
		$N^{\eta1}$	Asn98	$O^{\delta1}$	*B*	3.24		$N^{\eta1}$	Thr19	$O^{\gamma1}$	*B*	2.92
		$N^{\eta2}$	Wat11	O	—	*2.33*		$N^{\eta2}$	Ser97	O^{γ}	*A*	2.90
		$N^{\eta2}$	Thr19	$O^{\gamma1}$	*A*	2.92		$N^{\eta2}$	Asn98	$O^{\delta1}$	*A*	2.83



(*a*) (*b*)

Figure 5
Hydrogen bonds stabilizing the dimeric structure of TeRbcX (*a*) around Arg75 and (*b*) between the antiparallel helices α4. Water molecules are shown as red spheres. $2F_o - F_c$ electron-density maps are shown with the contour level set to 1.5σ in (*a*) and 1.0σ in (*b*). In (*b*), the kink in the antiparallel helix α4 is clearly visible. The protomers of the TeRbcX dimer are coloured green (molecule *A*) and cyan (molecule *B*). Hydrogen bonds are shown as dashed lines.

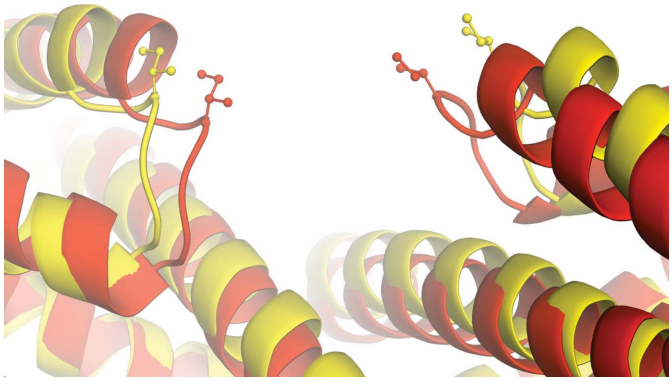


Figure 6
RbcX from *T. elongatus* (PDB entry 3q20; red) and the open (yellow) conformation as modelled by *ElNémo* (Suhre & Sanejouand, 2004). This enlarged view shows the $S_{49}IQD_{52}$ loops and the Ile50 side chain (ball-and-stick representation).

TeRbcX. Upon this 'butterfly-type' motion, which is pivoted on the kinks in helices α4, major rearrangements occur at the sites involved in RbcL binding. According to the obtained results, it appears that the RbcX protein can adopt two conformations. The present TeRbcX crystal structure is in fact very similar to the experimental ligand-bound form (PDB entry 2pem). In such a closed conformation the Ile50 side chains of the $S_{49}IQD_{52}$ loops of molecules *A* and *B* are about 10 Å apart, whereas in the open conformation resulting from the normal-mode analysis this distance is dramatically increased (Fig. 6). There is a good correlation between computed (derived form normal-mode analysis) and observed (crystallographic) *B* factors. The correlation factor of 0.749 for 221 C^{α} atoms (computed for 100 lowest frequency normal modes) indicates that the normal modes are a realistic description of the protein flexibility. An animation showing the structural changes leading from the closed to the open conformation is available as Supplementary Fig. 2. It can be speculated that the closed–open conformational changes occur upon binding and release by TeRbcX of the C-terminal fragment of RbcL.

4. Conclusions

In general, the present structure of the first RbcX protein from a thermophilic organism resembles those of previously described RbcX proteins from mesophilic species, implying a dimer as the minimal structural unit in the functional role of the protein as a chaperone in the assembly of the $RbcL_8$ module of the RuBisCO complex. The C_2 dimeric structure of TeRbcX is supported by extended hydrophobic interactions between the two four-helix-bundle cores and by the exchange of the C-terminal parts of the long α4 helices between the two subunits.

The TeRbcX protein contains an unusual and unique Cys103 residue in its sequence. However, the location of this cysteine residue on the protein surface, in the distal part of helix α4, excludes its involvement in physiological dimerization or in interactions with the cognate C-terminal peptide of RbcL. The TeRbcX-C103A mutant remains functionally competent in the assembly of $RbcL_8$ in *E. coli* co-expression (data not shown).

Comparison with other RbcX models suggests that the TeRbcX dimer can adopt an open conformation with a very wide entrance to the substrate-binding groove that is formed at the dimer interface. It can be predicted that in order to facilitate substrate docking the RbcX molecule assumes a fully open conformation and then undergoes a conformational change through a hinge movement pivoted in the middle of the long α4 helices. It might thus be concluded that the RbcX molecule is capable of closing–opening motions upon ligand binding.

This work was supported in part by a grant from the Ministry of Science and Higher Education (No. N N303 330334) and by Research Grant 1013/S/WB from the University of Wroclaw awarded to AS.

References

Andersson, I. & Taylor, T. C. (2003). *Arch. Biochem. Biophys.* **414**, 130–140.
Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. (1995). *Protein Sci.* **4**, 2455–2468.
Berezin, C., Glaser, F., Rosenberg, J., Paz, I., Pupko, T., Fariselli, P., Casadio, R. & Ben-Tal, N. (2004). *Bioinformatics*, **20**, 1322–1324.
Burla, M. C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G., Siliqi, D. & Spagna, R. (2007). *J. Appl. Cryst.* **40**, 609–613.
Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Mazzone, A. M. & Siliqi, D. (2006). *J. Appl. Cryst.* **39**, 185–193.
DeLano, W. (2002). *PyMOL*. http://www.pymol.org.
Dhingra, A., Portis, A. R. & Daniell, H. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 6315–6320.
Ellis, R. J. (1979). *Trends Biochem. Sci.* **4**, 241–244.
Ellis, R. J. (2006). *Trends Biochem. Sci.* **31**, 395–401.
Emekli, U., Schneidman-Duhovny, D., Wolfson, H. J., Nussinov, R. & Haliloglu, T. (2008). *Proteins*, **70**, 1219–1227.
Emlyn-Jones, D., Woodger, F. J., Price, G. D. & Whitney, S. M. (2006). *Plant Cell Physiol.* **47**, 1630–1640.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E. & Ben-Tal, N. (2003). *Bioinformatics*, **19**, 163–164.
Gronenborn, A. M. (2009). *Curr. Opin. Struct. Biol.* **19**, 39–49.
Hooft, R. W., Sander, C. & Vriend, G. (1996). *Proteins*, **26**, 363–376.
Jaskólski, M. (2001). *Acta Biochim. Pol.* **48**, 807–827.
Keating, K. S., Flores, S. C., Gerstein, M. B. & Kuhn, L. A. (2009). *Protein Sci.* **18**, 359–371.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. & Ben-Tal, N. (2005). *Nucleic Acids Res.* **33**, W299–W302.
Li, L.-A. & Tabita, F. R. (1997). *J. Bacteriol.* **179**, 3793–3796.
Liu, C., Young, A. L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B. V., Rao, K. V., Berninghausen, O., Mielke, T., Hartl, F. U., Beckmann, R. & Hayer-Hartl, M. (2010). *Nature (London)*, **463**, 197–202.
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.
Onizuka, T., Endo, S., Akiyama, H., Kanai, S., Hirano, M., Yokota, A., Tanaka, S. & Miyasaka, H. (2004). *Plant Cell Physiol.* **45**, 1390–1395.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Painter, J. & Merritt, E. A. (2006*a*). *Acta Cryst.* D**62**, 439–450.
Painter, J. & Merritt, E. A. (2006*b*). *J. Appl. Cryst.* **39**, 109–111.
Saschenbrecker, S., Bracher, A., Rao, K. V., Rao, B. V., Hartl, F. U. & Hayer-Hartl, M. (2007). *Cell*, **129**, 1189–1200.
Suhre, K. & Sanejouand, Y. H. (2004). *Nucleic Acids Res.* **32**, W610–W614.
Tabita, F. R. (1999). *Photosynth. Res.* **60**, 1–28.
Tanaka, S., Sawaya, M. R., Kerfeld, C. A. & Yeates, T. O. (2007). *Acta Cryst.* D**63**, 1109–1112.
Tarnawski, M., Gubernator, B., Kolesinski, P. & Szczepaniak, A. (2008). *Acta Biochim. Pol.* **55**, 777–785.
Tarnawski, M., Krzywda, S., Szczepaniak, A. & Jaskolski, M. (2008). *Acta Cryst.* F**64**, 870–874.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.