Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Chaperone Assisted RuBisCO Folding and Assembly-Role of RbcX

Karnam Vasudeva Rao

aus

Chintrapalli, Karnataka, India

2009

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Professor Dr. F. Ulrich Hartl betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfen erarbeitet.

München, am ....

....

Karnam Vasudeva Rao

Dissertation eingereicht am   ……………….2009

1. Gutacher   Prof. Dr. F. Ulrich Hartl

2. Gutachter   PD Dr. Konstanze F. Winklhofer

Mündliche Prüfung am   27.02.2009………………

Acknowledgements

First of all, I wish to express my gratitude to Prof. Dr. F. Ulrich Hartl and Dr. Manajit Hayer-Hartl for giving me the opportunity to conduct my PhD research here. My most earnest acknowledgment must go to them for their encouragement, continual support and invaluable suggestions during my study. I appreciate all their contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating.

I would like to thank Prof. Dr. Konstanze Winklhofer for being in my thesis committee as second Gutachter.

I would like to thank Dr. Andreas Bracher, Sandra and Markus Stemp for their friendly cooperation, help and insightful discussions. Special thanks to Andrea and Silke for their help and support in the administrative matters. A lot of thanks to Bernd, Emannuel, Nadine and Elisabeth for their careful organization which helped my research move smoothly.

Thanks to Prof. Dr. Jürgen Soll for providing us *A. thaliana* cDNA. Thanks to Dr. Frank Siedler for the mass spectrometry analysis mentioned in this thesis. I very much appreciate his time and suggestions.

I wish to extend my warmest thanks to Prof. Dr. G. R. Naik, Department of Biotechnology, Gulbarga University, India and Dr. V. R. Patil, Rallis India Ltd. India, for all the support and motivation.

One of the most important persons who has been with me in every moment of my PhD tenure is my wife Bharathi. I would like to thank her for the many sacrifices she has made to support me in undertaking and completing my doctoral studies. By providing her steadfast support in hard times, she has once again shown the

true love and dedication she has always had towards me. Similarly, uncountable thanks to my parents, Bharathi's parents and our family members for their love and support throughout my PhD research, as always.

Contents

Summary

Protein stability is intimately connected with protein folding; proteins have to be folded into their final active state (and maintain it) to be stable. The resulting three-dimensional structure is determined by the sequence of the amino acids. Many proteins, especially large multidomain proteins often refold inefficiently, resulting in misfolded states that tend to aggregate. In the cell, the folding of many proteins is therefore guided by molecular chaperones.

RuBisCO is one of the key factors for photosynthesis as it is responsible for the fixation of atmospheric carbon dioxide. RuBisCO is generally inefficient as a catalyst. The efficiency of RuBisCO varies among photosynthetic organisms. Red algae RuBisCOs have higher specificity for CO_2 than those of higher plants. One of the attempts to improve RuBisCO is through *in vitro* mutagenesis and protein engineering. In cyanobacteria and higher plants, RuBisCO (Form I) is a ~520 kDa complex composed of eight large subunits (RbcL) and eight small subunits (RbcS). The classical GroEL and GroES chaperones have been implicated in the folding of newly translated RbcL subunits. However, the mechanism of folding and assembly of RuBisCO is poorly understood, and thus attempts to reconstitute Form I RuBisCO have been unsuccessful so far.

The formation of $RbcL_8$ core complexes is thought to be one of the major steps in RuBisCO assembly. In the present study, RbcX emerged as an assembly chaperone, assisting in the productive assembly of cyanobacterial RbcL to $RbcL_8$, which otherwise would misassemble in the absence of RbcX. RbcX acts downstream of GroEL-mediated RbcL folding. Structural and functional analyses revealed that RbcX forms a homodimer with two cooperating RbcL binding regions, a central cleft and a peripheral surface. The central cleft specifically binds the exposed C-terminal peptide of RbcL subunits, enabling the peripheral surface of RbcX to mediate $RbcL_8$ assembly. Due to the dynamic nature of these interactions, RbcX could be displaced from $RbcL_8$ complexes by RbcS, producing the active

enzyme. Species-specific co-evolution of RbcX with RbcL and RbcS accounts for a limited interspecies exchangeability of RbcX and for RbcX-supported or -dependent assembly modes. However, the RbcX homolog from *Arabidopsis thaliana* could efficiently complement for RbcX from cyanobacteria in promoting the assembly of cyanobacterial RbcL to $RbcL_8$ core complexes, which suggests a universal role of RbcX in RuBisCO assembly among organisms possessing Form I RuBisCO.

The present study contributes to a better understanding of the role of chaperones in the assembly of complex proteins and provides insights that may be helpful in future attempts to improve RuBisCO performance by *in vitro* mutagenesis.

2 Introduction

2.1 Proteins

Proteins are large biological molecules with a molecular weight up to few million Daltons. Proteins are made up of amino acids linked into linear chains, called polypeptide chains. Amino acids are linked with each other by peptide bonds – which are formed between the carboxyl and amino groups of neighboring amino acids. Proteins are formed by one or several polypeptide chains. The sequence of the polypeptide chain is defined by the nucleotide sequence of its gene. There are only 20 standard amino acids existing in living organisms. Sometimes these amino acids are chemically modified in the proteins after protein synthesis.

2.2 Protein structure and stability

The structural organization of proteins can be divided into four different levels.

Primary structure or protein sequence: The protein sequence or amino acid sequence of a polypeptide chain defines its primary structure. DNA (or RNA in viruses) codes for the primary structure of protein and thus provides the comprehensive information for the protein structure and function.

Secondary structure: One of the main conformational parameters of the amino acid structure is the value of the Φ and Ψ angles. These angles completely define the conformation of the polypeptide chain. With some special values for these angles, the main chain can adopt specially classified conformations such as the alpha-helix and beta-strand. The other main feature of the protein secondary structure is the local stabilization by hydrogen bonds.

Ternary structure, protein 3D structure or protein fold: Ternary structure or protein fold completely defines the structural organization of the protein molecule in 3-dimensions.

Quaternary structure: The interaction between several protein molecules to form protein complexes is defined as quaternary structure.

Protein stability is intimately connected with protein folding - proteins have to be folded into their final active state (and maintain it) to be stable. Protein folding is the physical process by which a polypeptide folds into its characteristic three-dimensional structure. A sequence of mRNA is translated to a linear chain of amino acids, a polypeptide which lacks any developed three-dimensional structure. However each amino acid in the chain can be thought of having certain 'gross' chemical features. These may be hydrophobic, hydrophilic, or charged and so on. To reach the well defined folded state of protein called the 'native state', these residues interact with each other and their environment. The resulting three-dimensional structure is determined by the sequence of the amino acids (Anfinsen, 1972).

2.3   Protein aggregation and diseases

Protein folding studies were initially carried out in *in vitro* conditions using purified and chemically denatured peptides as substrates. Under appropriate *in vitro* conditions many proteins can fold spontaneously into their native states. However, a major difference between the conditions used for *in vitro* folding studies and those relevant *in vivo* is the highly crowded cytosolic environment. For example: *E.coli* cytosol has 300-400 g/liter of proteins and other macromolecules (Ellis, 2001a). The resulting 'macromolecular crowding' or 'excluded volume effect' can influence many reactions in the cell such as increased intermolecular interactions which can lead to protein aggregation. Aggregation is mediated by exposed hydrophobic residues of nonnative protein chains in a concentration-dependent manner (van den Berg et al., 1999). It takes 1-2 minutes to synthesize an average protein in eukaryotes at a speed of ~4 amino acids per second (Thulasiraman et al., 1999). Folding into a stable structure is only possible once a complete domain (~100-300 residues) has emerged from the ribosome. During synthesis the nascent

chains are therefore prone to aggregation as they expose hydrophobic surfaces for a considerable length of time.

Large multidomain proteins often refold inefficiently resulting in misfolded states that tend to aggregate. Aggregation of soluble proteins or peptides into insoluble amyloid fibrils (multimerization of misfolded proteins into structured, fibrillar aggregates) (Figure 1) with cross-ß structure has been shown to be associated with a variety of diseases. Over 20 diseases have been linked to deposits of amyloid fibrils or plaques derived from different precursor proteins. These amyloidogenic diseases include Alzheimer's disease, Parkinson's disease, type II diabetes, and spongiform encephalopathies. The human health impact of these diseases has motivated intensive study of the structure and growth of amyloid fibrils (Dobson, 1999; Murphy, 2002; Schenk, 2002 ; Sipe, 2000; Temussi, 2003; Walsh, 2003).

In certain solutions and under some conditions proteins will not fold into their biochemically functional forms. Temperatures above, and sometimes those below the optimum for cells can cause proteins to unfold or "denature". High concentrations of solutes, extremes of pH, mechanical forces, and chemical denaturants can cause similar effects. A fully denatured protein lacks both tertiary and secondary structure, and exists as a 'random coil'. Under certain conditions some proteins can refold; however, in many cases denaturation is irreversible (Shortle, 1996). Aggregation may occur between productive folding intermediates or between kinetically trapped misfolded intermediates exposing hydrophobic surfaces. To avoid these problems, living cells possess extensive machinery of protein quality control. Molecular chaperones are a major component of protein quality control and many proteins depend on them for efficient folding.

2.4   Molecular chaperones

Living cells have developed systems for preventing misfolding and aggregation of newly synthesized proteins and for controlling the folding status of polypeptides under normal and stress conditions. Molecular chaperones are known to prevent inter- and intra-molecular interactions by binding to the hydrophobic surfaces of non native proteins, thereby influencing the partitioning between productive and non productive folding pathways. Chaperones do not form part of the final structures of folded proteins. They cycle substrates between bound and free states and provide an opportunity for folding whenever the substrate is released.
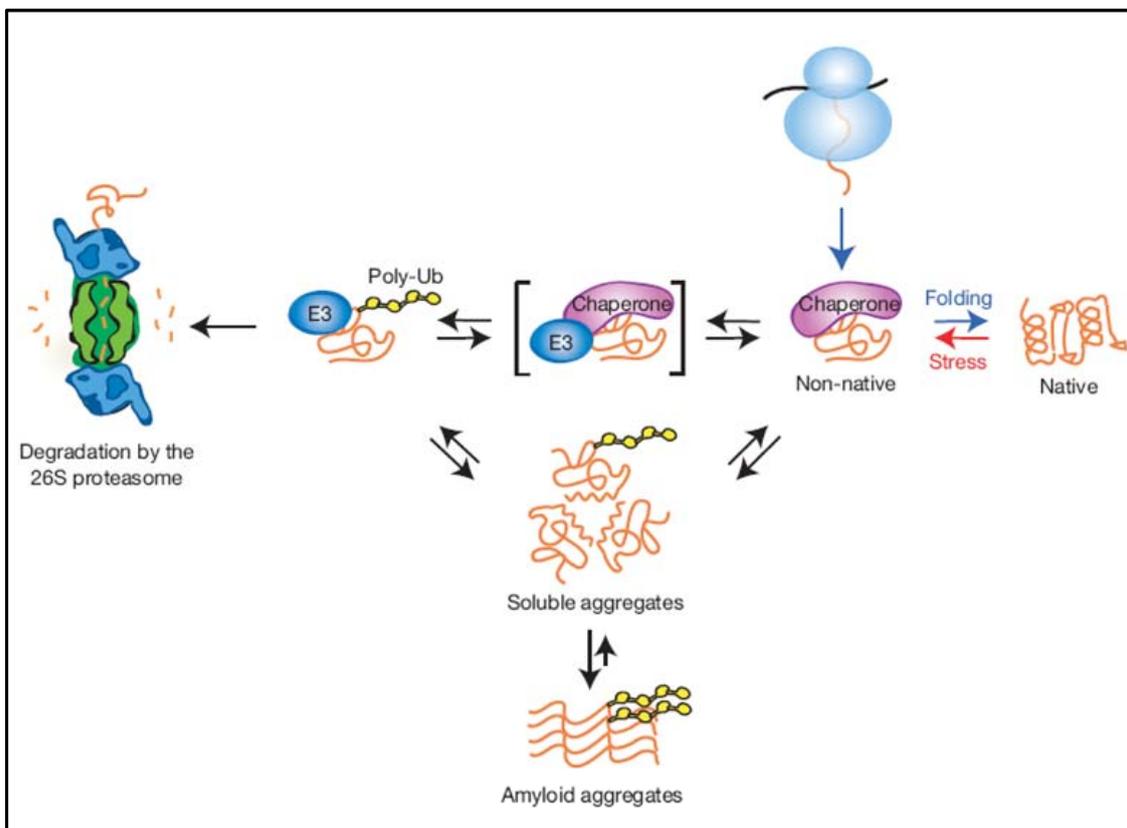
Figure 1. Life of newly synthesized proteins and role of molecular chaperones in protein quality control.
Molecular chaperones facilitate folding of newly synthesized polypeptides to the native state (blue arrow). Chaperones also bind to non-native intermediates that are generated when native proteins are denatured, for example by stress (red arrow). Cellular surveillance results in either refolding

(blue 'folding' arrow) or elimination by the ubiquitin–proteasome system (arrows pointing to the left, which ultimately end in 'degradation'). Degradation generally requires tagging by ubiquitination via a quality control E3 ligase. In some cases, ubiquitination of the chaperone-bound substrate is directed by specific interaction of certain chaperones, for example Hsp70 or Hsp90, with E3s containing chaperone-interaction domains, such as CHIP (complex in brackets). Under some circumstances, such as stress or aging, quality control efforts may fail, either before or after adding a polyubiquitin chain. In these cases, misfolded proteins may form small soluble aggregates, which, if unresolved by refolding or degradation, go on to form heat- and detergent-resistant amyloidogenic aggregates (McClellan et al., 2005b).

Chaperones that participate broadly in *de novo* protein folding, such as Hsp70 and chaperonins (GroEL, thermosome, TRiC) promote the folding process through substrate binding and release regulated by their ATPase activity and by cofactor proteins. Chaperones such as Hsp90 and Hsp100 also interact with substrates in an ATP dependent manner while others such as Trigger factor (TF), sHsps interact with substrates in an ATP independent manner (Frydman et al., 1994; Langer et al., 1992; Siegers et al., 1999). The basic features of the major chaperone systems are described below.

2.4.1 Chaperones of three domains of life

Cytosolic chaperones can be devided into 'holding' chaperones (e.g., Hsp70s) and 'folding' chaperones (e.g., the cylindrical chaperonins) depending on their abilities to either mainly prevent protein aggregation or to promote proper protein refolding, respectively. These two types of chaperone exist in all the three domains of life, Eubacteria, Archaea and Eukarya (Hartl and Hayer-Hartl, 2002) (Figure 2).

In eubacteria, TF is a ribosome associated chaperone providing protection for nascent polypeptide chains against misfolding and aggregation (Figure 2A). The majority of small proteins may fold subsequently spontaneously without any assistance from chaperones. Longer polypeptide chains need assistance from the second class of nascent-chain binding chaperones (classical Hsp70s and prefoldin), which are not directly associated with ribosomes. These chaperones may either associate co-translationally to stabilize elongating chains and /or act

post-translationally in assisting folding or transfer of newly synthesized chains to downstream chaperones (Figure 2A). For 10-15% of newly made polypeptides (which include mostly slow folding and aggregation-sensitive proteins), the chaperonins (GroEL/ES, Thermosome or TRiC) assist folding post-translationally in both prokaryotes and eukaryotes (Genevaux et al., 2004). Additionally, ATP-dependent chaperones such as Hsp90 is necessary for the folding of a subset of eukaryotic proteins.
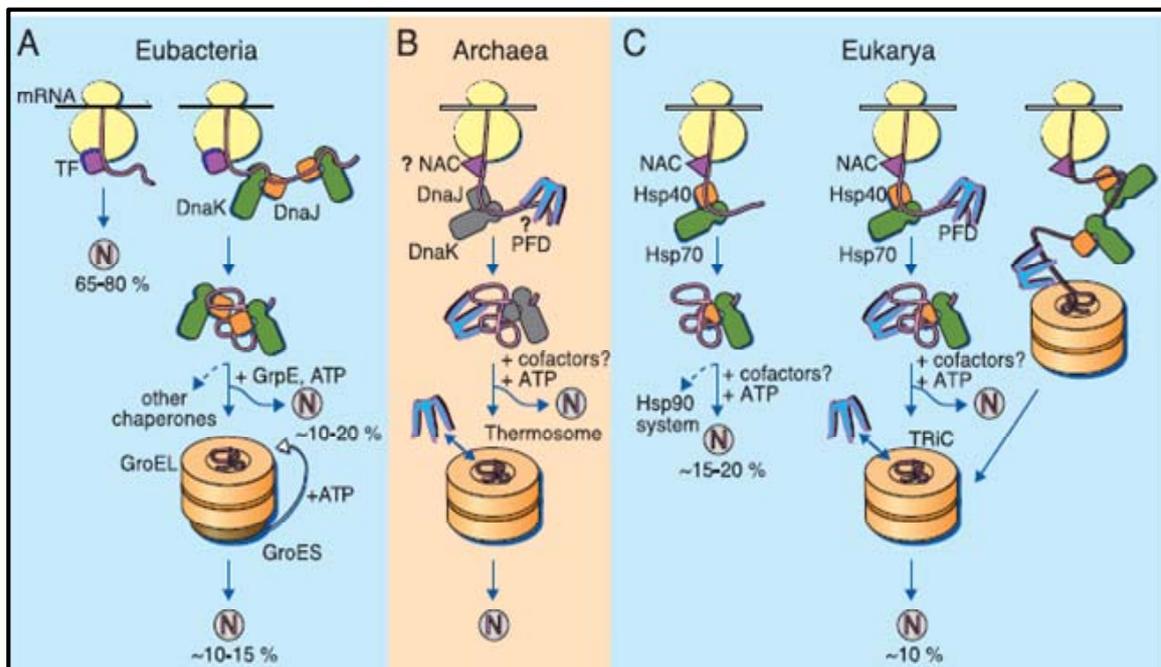
Figure 2. Models for the chaperone-assisted folding of newly synthesized polypeptides in the cytosol.
(A) Eubacteria. TF, trigger factor; N, native protein. Nascent chains probably interact generally with TF, and most small proteins (approximately 65 to 80% of total) fold rapidly upon synthesis without further assistance. Longer chains (10 to 20% of total) interact subsequently with DnaK and DnaJ and fold upon one or several cycles of ATP-dependent binding and release. About 10 to 15% of chains transit the chaperonin system-GroEL and GroES-for folding. GroEL does not bind to nascent chains and is thus likely to receive an appreciable fraction of its substrates after their interaction with DnaK.
(B) Archaea. PFD, prefoldin; NAC, nascent chain-associated complex. Only some archaeal species contain DnaK/DnaJ. The existence of a ribosome-bound NAC homolog, as well as the interaction of PFD with nascent chains, has not yet been confirmed experimentally.

(C) Eukarya-the example of the mammalian cytosol. Like TF, NAC probably interacts generally with nascent chains. The majority of small chains may fold upon ribosome release without further assistance. About 15 to 20% of chains reach their native states in a reaction assisted by Hsp70 and Hsp40, and a fraction of these must be transferred to Hsp90 for folding. About 10% of chains are co- or posttranslationally passed on to the chaperonin TRiC in a reaction mediated by PFD (Genevaux et al., 2004; Hartl and Hayer-Hartl, 2002).

2.4.2 Ribosome-bound, nascent polypeptide-interacting chaperones

TF is the first chaperone to interact with nascent polypeptides as they emerge from the bacterial ribosome (Figure 2A). It binds to the ribosome as a monomer but forms dimers in free solution (Agashe et al., 2004) . Several studies have shown a general role of TF in cytosolic protein folding that overlaps partially with that of the bacterial Hsp70 chaperone system DnaK, DnaJ and GrpE (Deuerling et al., 1999; Pfund et al., 1998; Teter et al., 1999) (Figure 2A). *E. coli* cells lacking TF (Δ*tig*) or DnaK (Δ*dnak*) exhibit no apparent folding defects at 37°C; however, deletion of *dnaK* in a Δ*tig* strain is lethal mainly because of large scale protein aggregation (Genevaux et al., 2004). The combined deletion of both systems is possible at lower temperature, and the resulting Δ*tig* Δ*dnaK* cells can be adapted to growth up to 30°C. Other chaperones, such as SecB and GroEL are thought to compensate partially for the loss of TF and DnaK under these conditions (Genevaux et al., 2004; Vorderwulbecke et al., 2004). TF has been shown to increase the folding efficiency of certain multidomain proteins concomitant with delaying their folding relative to translation (Agashe et al., 2004).

Nascent chain-associated complex (NAC) is the first known eukaryotic cytosolic factor that interacts with the nascent polypeptide chains emerging from the ribosome. (Figure 2C) (Moller et al., 1998; Wiedmann et al., 1994). NAC associates with short nascent chains and dissociates upon chain release from the ribosome (Beatrix et al., 2000; Hartl, 1996b). NAC is known to prevent mistargetting of nascent polypeptide chains to the endoplasmic reticulum (Lauring et al., 1995a; Lauring et al., 1995b). NAC is highly conserved from yeast to

humans. A role of stimulating protein import into yeast mitochondria has also been proposed for NAC (Funfschilling and Rospert, 1999). Mutations in beta-NAC in Drosophila could lead to the development of embryonic lethality (Markesich et al., 2000). Eukaryotic NAC is a heterodimer with two subunits sharing substantial homology (Makarova et al., 1999). The archaebacterial NAC homolog (Figure 2B) forms a homodimer and the crystal structure revealed the presence of a ubiquitin-associated domain that suggests a yet to be identified role in ubiquitination pathway (Spreter et al., 2005).

Ssb1 and Ssb2 are specialized nascent chain-chaperones in yeast and other fungi. These Hsp70 chaperones have been shown to interact with the ribosome and short nascent chains. They functionally cooperate with another Hsp70, Ssz1, which forms a stable complex called RAC (Ribosome Associated Complex) with zuotin, an Hsp40 homologue (Gautschi et al., 2001; Gautschi et al., 2002).

2.4.3 The Hsp70 chaperone family

Hsp70 homologues are found in both prokaryotes and eukaryotes. They occur in the cytosol, mitochondria, chloroplasts and the endoplasmic reticulum. DnaK is the major Hsp70 protein found in *E. coli* (Figure 2A). ATP binding and hydrolysis allow DnaK to shift between low and high affinity states for substrates, respectively. ATP hydrolysis, binding and release of unfolded polypeptides from DnaK are regulated by cochaperone DnaJ (Hsp40) and the nucleotide exchange factor GrpE (a homodimer of 20kD subunits) (Hartl and Hayer-Hartl, 2002) (Figure 2A). Repeated cycles of substrate binding and release promote the folding and assembly of proteins in *E. coli* cells (Agashe and Hartl, 2000; Ellis and Hartl, 1999). DnaK, DnaJ and GrpE occur in a molar ratio of ~10:1:3 (Slepenkov and Witt, 2002).

DnaK protects nascent polypeptides from premature misfolding and aggregation (Deuerling et al., 1999; Teter et al., 1999). DnaK exhibits a 'buffer' function by efficiently preventing the aggregation of unfolded proteins. In cooperation with ClpB (Hsp100) (Haslberger et al., 2007; Zietkiewicz et al., 2004), Hsp31 (Mujacic et al.,

2004) or the small heat shock proteins IbpA and IbpB (Mujacic et al., 2004). At elevated temperatures, DnaK prevents protein aggregation. DnaK is present at approximately 10,000 copies per cell under normal conditions at 37^0C and it doubles following a shift to 42^0C (Genevaux et al., 2007). DnaK is also known to protect proteins against oxidative damages (Echave et al., 2002; Fredriksson et al., 2005). DnaK/DnaJ/GrpE has been shown to be involved in the protein secretion in *E.coli* (Wild et al., 1996). DnaK negatively controls the rate of its own synthesis by binding to the stress σ32 transcription factor and facilitating its degradation by proteases (Straus et al., 1990).

DnaK is composed of an N-terminal ATPase domain (residues 1-385) and a C-terminal substrate binding domain (residues 393-537) that is capped by a lid domain (residues 538-638). DnaK binds ATP tightly but hydrolyzes it very slowly (Russell et al., 1998). The binding motif that DnaK prefers includes approximately seven residues consisting of hydrophobic amino acids enriched in the middle, including leucine and isoleucine. These residues are usually exposed only in non-native proteins (Hartl and Hayer-Hartl, 2002).

DnaK reaction cycle:

In most cases, the unfolded substrate is transferred to ATP-bound DnaK by its cofactor DnaJ, which recognizes hydrophobic peptides via its C-terminal domain (Langer et al., 1992; Rudiger et al., 2001; Sha et al., 2000) and binds to Dnak with its N-terminus (Figure 3A). At this point the alpha helical latch of DnaK over the peptide binding cleft is in the open conformation. ATP hydrolysis by DnaK is accelerated by the binding of DnaJ, which results in the closure of binding site. DnaJ will then disassociate and the nucleotide exchange factor GrpE induces the release of ADP from DnaK (Erbse et al., 2004; Harrison et al., 1997; Mayer et al., 2000; Young et al., 2004). Upon rebinding of ATP the protein is released (Figure 3A). This leaves the substrate binding cavity free and open to receive another peptide.

2.4.4   GimC/Prefoldin

Prefoldin (PFD) is found in the cytosol of eukaryotes and archaea. It is a heterohexameric complex (consisting of six subunits in the range of 14-23 kD). Prefoldin has a jellyfish-like structure with six long coiled tentacles and a large central cavity (Martin-Benito et al., 2002; Martin-Benito et al., 2007; Ohtaki et al., 2008; Siegert et al., 2000b) (Figure 3B). Prefoldin is the mammalian orthologue of yeast Gim protein complex identified through a set of Genes Involved in Microtubule biogenesis in yeast (Geissler et al., 1998). Prefoldin stabilizes nonnative proteins and subsequently delivers them to a group II chaperonin (e.g., thermosome, TRiC (TCP-1 Ring Complex, also called CCT, chaperonin containing TCP-1)) in eukaryotes, for final folding (Siegers et al., 1999; Vainberg et al., 1998; Zako et al., 2006) (Figure 2C).

In addition to Hsp70, which has been shown in several studies to interact with nascent chains of actin, prefoldin has been implicated in the folding of actin and tubulin (Vainberg et al., 1998). Prefoldin binds to actin cotranslationally and remaines bound until substrate transfer to TRiC. On the other hand, Prefoldin interacts with tubulin following the synthesis of at least 250 residues (Hansen et al., 1999c; Vainberg et al., 1998). Recently, the X-ray structure and simulation studies suggested that the β-subunit binds to substrate and the α-subunit interacts with TRiC and modulates the shape and width of the central prefoldin cavity (Ohtaki et al., 2008).
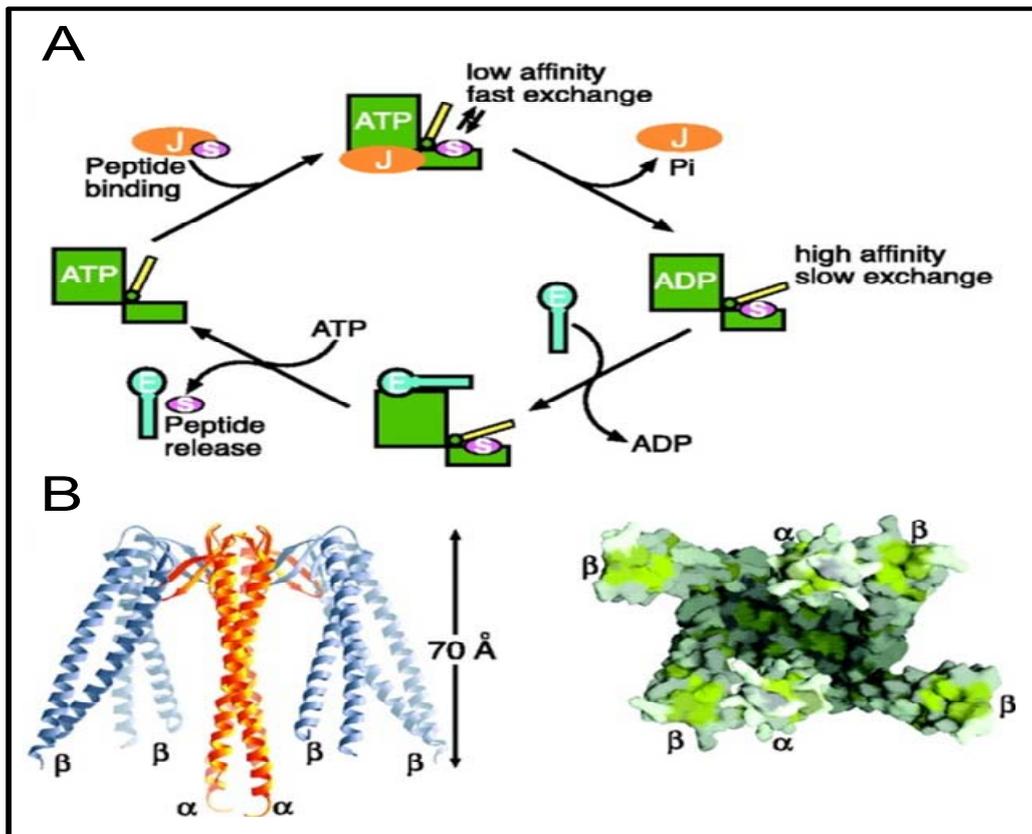
Figure 3. Structure and function of chaperones with the ability to bind nascent chains.

(A) Simplified reaction cycle of the DnaK system. Not all substrates are presented to DnaK by DnaJ. The intermediate DnaK-DnaJ-substrate-ATP is probably very transient, as this is the fastest step of the cycle.

(B) (Left) Side view and dimension of the structure of archaeal PFD with the two α subunits shown in gold and the four β subunits in gray. (Right) Bottom view of the PFD complex showing the central space enclosed by the six coiled-coil segments. Surface representation is shown with hydrophobic patches in yellow and hydrophilic regions in gray (Hartl and Hayer-Hartl, 2002; Siegert et al., 2000a)

2.4.5 Chaperonins

Chaperonins are large, cylindrical complexes that enhance the efficiency of protein-folding reactions by sequestering protein substrates in their central cavities. They are ubiquitously distributed in nature and are indispensable for cell survival and growth.

Chaperonins are divided into two groups: group I is found in bacteria, mitochondria, and plastids (Hartl and Hayer-Hartl, 2002; Ranson et al., 1998) and the group II in the eukaryotic cytoplasm and in archaea (Kim et al., 1994). In both cases protein binding and release is ATP regulated.

The best studied group I chaperonin system is the *E. coli* GroE system (Figure 4). GroEL is a large cylindrical protein that recognizes the hydrophobic surfaces of the substrates by using its multiple subunits. GroEL is composed of two heptameric rings of identical 57-kD subunits stacked back to back. Unfolded proteins spend their most of time in the hydrophilic, non-interacting cage where the unproductive non-native interactions are prevented (Ranson et al., 1997). The co-chaperone GroES is made up of identical 10-kD seven subunits and it binds transiently to the ends of GroEL cylinder. Non-native proteins up to around 60kD can be encapsulated by the GroEL-GroES cage (Brinker et al., 2001; Ellis, 2001b; Mayhew et al., 1996; Weissman et al., 1996)**.**

GroEL/ES reaction cycle:

The chaperonin reaction begins by binding of non-native substrates to the free end of GroEL-GroES complex (*trans* ring) (Figure 4B). Seven molecules of ATP and GroES then bind to the same ring, leading to the displacement of the substrate into the GroEL cavity capped by GroES. This causes dissociation of seven ADP molecules and GroES from the former *cis* complex. Binding of GroES causes large conformational changes in GroEL. The apical domains of GroEL, which bind substrate, make a clockwise rotation and upward movement upon GroES binding. This results in expansion of the volume of the cavity and also results in changing the inner surface of the *cis* cavity to hydrophilic that favors folding (Roseman et al., 1996; Xu et al., 1997; Xu and Sigler, 1998).
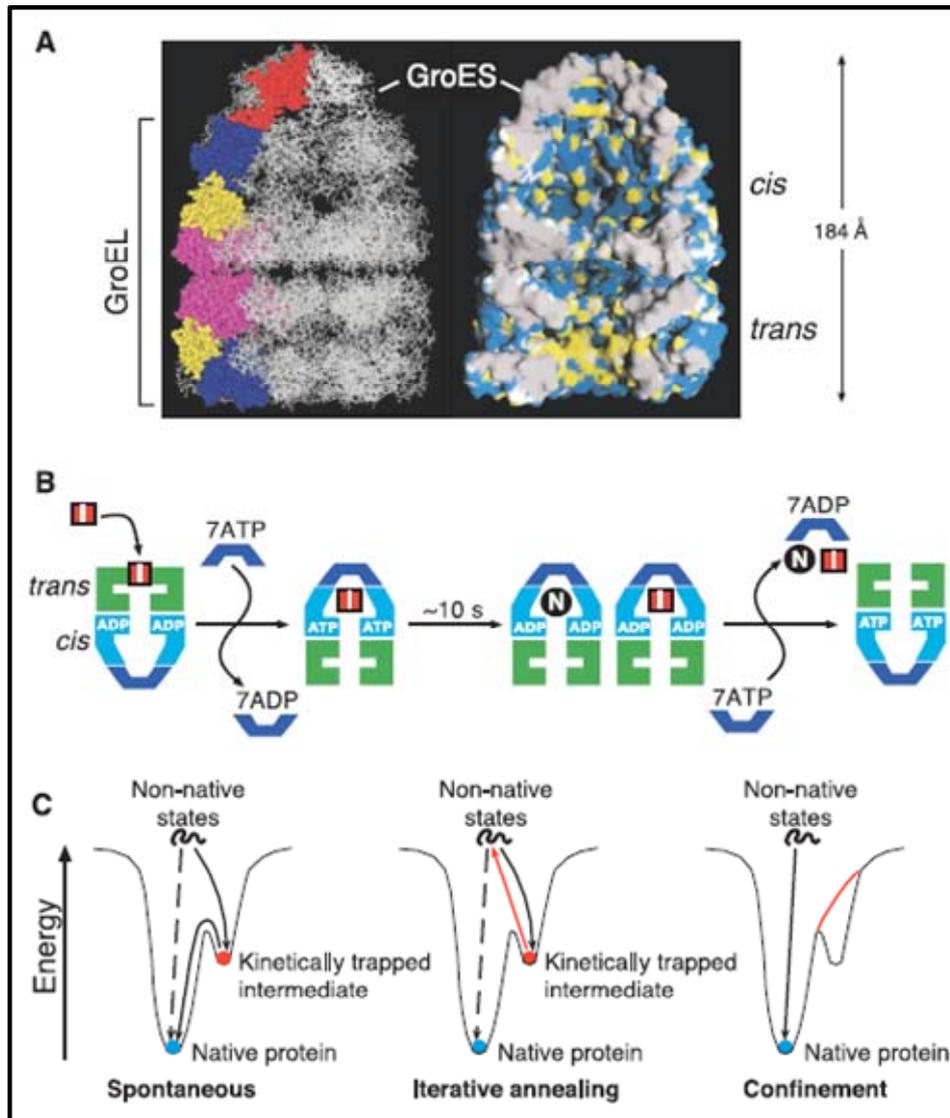
Figure 4. The GroEL-GroES chaperonin system.

(**A**) (Left) View of the asymmetric GroEL-GroES-(ADP)$_7$ complex. The equatorial, intermediate, and apical domains of one subunit each in the *cis* and *trans*-ring of GroEL are colored in pink, yellow, and dark blue, respectively, and one subunit of GroES is colored red. (Right) The accessible surface of the central cavity of the GroEL-GroES complex. Polar and charged side-chain atoms, blue; hydrophobic side-chain atoms, yellow; backbone atoms, white; and solvent-excluded surfaces at subunit interfaces, gray (Hartl and Hayer-Hartl, 2002; Xu et al., 1997).

(**B**) Simplified reaction of protein folding in the GroEL-GroES cage. I, folding intermediate bound by the apical domains of GroEL; N, native protein folded inside the cage. For a typical GroEL substrate, multiple rounds of chaperonin action are required for folding; both I and N accumulate after a single reaction cycle and exit the cage upon GroES dissociation. I is then rapidly re-bound by GroEL.

(**C**) Mechanisms of accelerated folding. Simple energy diagrams are shown for a protein that forms a kinetically trapped intermediate during spontaneous folding (left). In the iterative annealing model,

this intermediate is thought to be actively unfolded by GroEL/GroES (Shtilerman et al., 1999) and allowed to repartition (middle), whereas confinement of nonnative protein in the narrow, hydrophilic environment of the GroEL-GroES cage is suggested to result in a smoothing of the energy landscape (right), such that formation of certain trapped intermediates is avoided (Brinker et al., 2001). Both proposed mechanisms would result in accelerated folding.

GroEL substrates:

Formation of β-sheet structure is expected to be the most difficult step in the folding process and thus the αβ-domain proteins exhibit slow folding rates and misfolding or kinetic trapping which might occur though improper packing of helices and sheets within one domain or between domains. GroEL substrates, identified by co-immunoprecipitation and mass spectrometry (Houry et al., 1999; Kerner et al., 2005), preferentially contain αβ-domains. Several known stringent substrates such as rhodanese, RuBisCO and malate dehydrogenese belong to this group (Lin and Rye, 2006). These proteins fold efficiently only when they are enclosed in the GroEL-GroES cage (Brinker et al., 2001). Typical GroEL substrates are proposed to consist of preferentially two or more αβ-domains (Kerner et al., 2005). The estimated number of proteins with multiple αβ-domains in the *E. coli* cytoplasm is between 200 to 600 and GroEL may assist in folding and conformational maintenance of at least a subset of these proteins (Houry et al., 1999; Kerner et al., 2005).

The Group II chaperonins, TRiC of the eukaryotic cytoplasm and thermosome of archaea share the same overall architecture, yet are substantially different from bacterial chaperonins (Figure 5 & 2C) (Sigler et al., 1998; Spiess et al., 2004). TRiC shares the double-ring structure of group I chaperonins, but is composed of eight different gene products that form an eight-membered ring (Spiess et al., 2004; Valpuesta et al., 2002). All subunits have essentially the same ATPase domain, whereas the polypeptide binding domains are diverged (Kim et al., 1994). Group II chaperonins don't have a GroES-like cochaperone which works as lid for

GroEL chaperonin, instead a segment of their apical domains act as lid (Gutsche et al., 1999; Leroux and Hartl, 2000; Meyer et al., 2003).
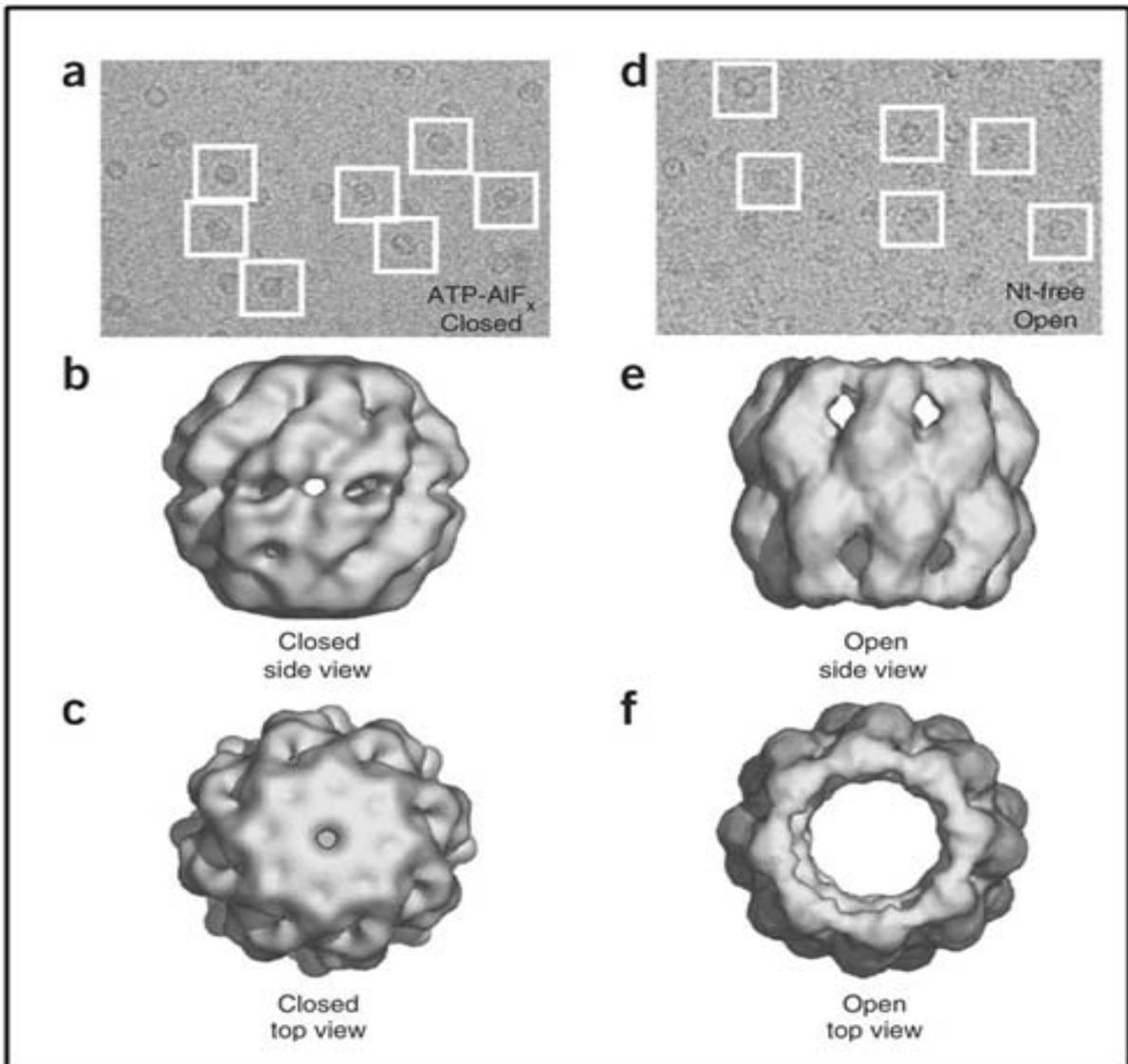
Figure 5. TRiC structure

(**a**) & (**d**) CCD captured images of ice embedded TriC in closed (nucleotide bound) and open states (nucleotide free).

(**b**,**c**) Side and top views of a single-particle reconstruction of TRiC in its closed state, carried out with 7,129 particles to ~15-Å resolution.

(**e**,**f**) Side and top views of a single-particle reconstruction of TRiC in its open state, carried out with 13,287 particles to ~18-Å resolution (Booth et al., 2008).

Recent studies using single-particle cryo-EM with comparative protein structure modeling studies showed that ATP-induced movements leading to lid closure change the properties of the central chamber facing the substrate. Moreover, the conformational change in the ATPase domain of chaperonin is conserved across chaperonins including GroEL but drives a different motion in the TRiC apical domains from that it produces in GroEL (Booth et al., 2008) (Figure 5).

The main TRiC substrates include the two cytoskeletal proteins, actin and tubulin. Earlier studies have shown that TRiC binds to unfolded polypeptides, thereby preventing their aggregation. It mediates ATP dependent renaturation of unfolded firefly luciferase and tubulin (Frydman et al., 1992). TRiC also has a large number of additional protein substrates, including myosin, enzymes, cell-cycle controllers, etc. (Dunn et al., 2001) α-transducin and VHL tumor suppressor protein (Hartl and Hayer-Hartl, 2002; McClellan et al., 2005a).

2.4.6 Chaperonins of plant chloroplast and cyanobacteria

Chloroplast chaperonin:

The chloroplast version of GroEL is Ch-Cpn60 (Figure 6), which was termed 'RuBisCO binding protein' because of its transient association with newly synthesized RuBisCO subunits (Ellis et al., 1988; Gatenby et al., 1988; Hemmingsen, 1988; Roy et al., 1988). The ch-cpn60, like other GroEL homologs, consists of two stacked heptameric rings (Pushkin, 1982). In contrast to chaperonins of bacteria and mitochondria, the native chloroplast Cpn60 contains two divergent subunits, cpn60α and cpn60β, each having the size of around 60 kDa (Musgrove et al., 1987) that are no more similar to each other than they are to GroEL (Martel et al., 1990).

In recent studies scientists have been able to assemble Ch-Cpn60 of pea (*Pissum sativum*), α and β-subunits together or β subunits alone (in the absence of α-subunit) to form tetradecamers *in vitro* (Cloney et al., 1992b), but it has not been possible to assemble α subunits into tetradecamers in the absence of β-subunit (Cloney et al., 1992b). The formation of $(\alpha\beta)_{14}$ and β_{14} requires adenine nucleotides. Both the reconstituted α/β 14-mers and the β14-mers can assist in protein folding under non-permissive conditions (Dickson et al., 2000).

It is unclear if the existence of distinct chaperonin subunits implies distinct substrate specificities. Also it is still unclear if one Cpn60 subunit is able to compensate for the function of the other *in vivo*. It is proposed that both Cpn60 subunit proteins might possess indispensable roles for acting cooperatively as a chaperonin oligomer (Nishio et al., 1999). This is evident from the finding that the assembly of *Brassica napus* stromal Cpn60 α-proteins into Cpn60 oligomers in *E.coli* cells requires coexpression of Cpn60 β proteins (Cloney et al., 1992a; Cloney et al., 1992b). It remains unclear if there are distinct α and β ring structures *in vivo* or if there are structures composed of varying proportions of α and β subunits.

Cyanobacterial chaperonin:

GroEL and GroES are encoded by a single, bicistronic *groESL* operon in most eubacterial strains, but several examples are known in which two, three or even five *groEL*-homologous genes and proteins exist in the same organism (Fischer, 1999; Tanaka et al., 1997; van der Vies, 1996). Among the photosynthetic prokaryotes, the completely sequenced *Synechocystis* PCC 6083 was reported to contain a groEL-like gene, named cpn60 (Chitnis and Nelson, 1991) in addition to a traditional groESL operon (Lehel et al., 1993). However the expression of both the chaperonins (GroEL and Cpn60) is regulated differently under light–dark transition during heat stress (Glatz et al., 1997). These findings support the hypothesis that multiple chaperonins play different physiological roles under stress

conditions. It was found that the GroEL showed the high chaperon activity compared to that of Cpn60 (Kovacs et al., 2001).
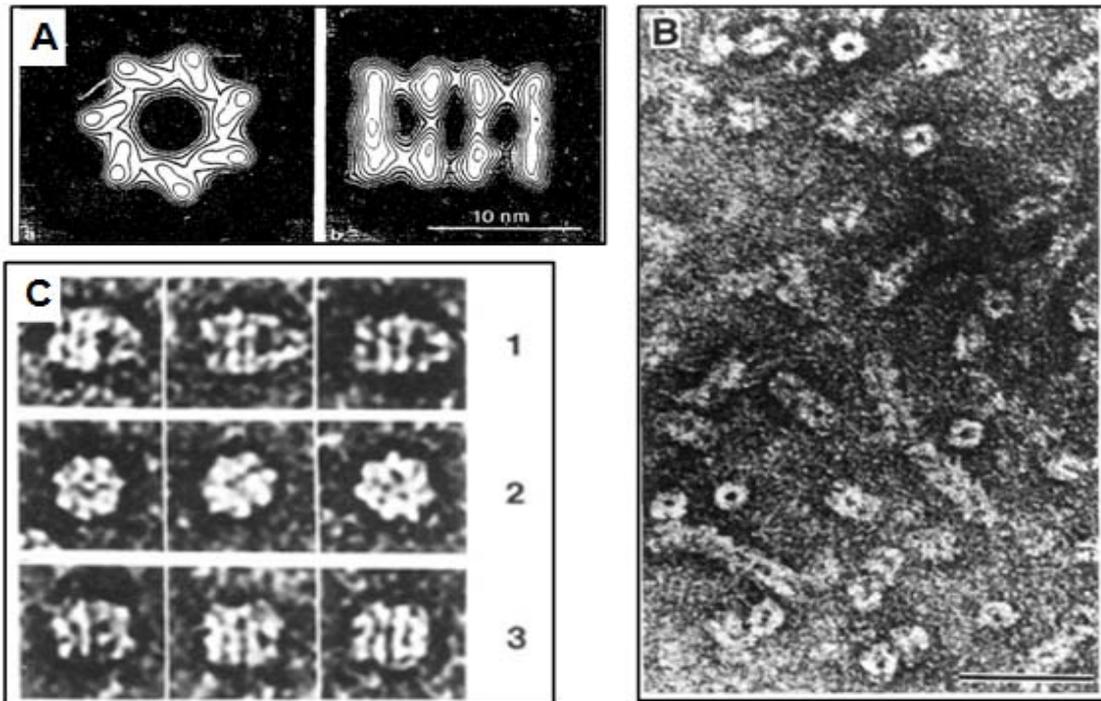
Figure 6. Chloroplast chaperones

(A) Molecular projections of Pea chloroplast Cpn60. a. Correlation averaging of images of top view with seven fold symmetry imposed b. The average of images of the side view (Tsuprun et al., 1991).
(B) Electron micrographs of recombinantly purified spinach Chloroplast Cpn20 negatively stained with uranyl acetate (Baneyx et al., 1995).
(C) Electron micrograph of asymmetric chloroplast chaperonin complexes formed in the presence of ADP. 1, side views of asymmetric complexes; 2, end views of asymmetric complexes and/or unliganded ch-cpn60; 3, side views of unliganded ch-cpn60 (Viitanen, 1995).

Two kinds of GroEL like proteins are found also in a thermophilic cyanobacterium, *Synechocuccus vulcanus* (Tanaka et al., 1997). This is of particular interest in terms of the endosymbiont hypothesis for the origin of chloroplasts which postulates that the chloroplasts are originated from a cyanobacterium. It is proposed that during the genesis of chloroplasts, gene flow of *groEL2* (and *groES)* occurred from the symbiont into the host nucleus. Gene duplication of the

cyanobacterial *groEL2* gene resulted in several genes for Cpn60α and Cpn60β, while the chloroplastic *groEL1* was most likely eliminated (Wastl et al., 1999).

2.4.7   Co-chaperones of chloroplast Cpn60

The chloroplast Cpn20 (Ch-Cpn20) has nearly twice the size of other known cpn10s. It consists of two complete GroES-like peptides that are fused head-to-tail through a short intervening linker (Baneyx et al., 1995; Bertsch et al., 1992). Both halves are highly conserved at a number of residues that are encoded in many GroES genes, suggesting that each domain could be functional (Baneyx et al., 1995). Each domain contains a "mobile loop" region analogous to that of GroES (Koonin and van der Vies, 1995). In *A. thaliana,* Ch-Cpn20 is encoded by a single gene in the nuclear genome and contains a transit sequence which directs the protein into the chloroplast stroma, where it is processed and assembled (Hill and Hemmingsen, 2001). Native gel electrophoresis and gel filtration analysis estimated the mass of 70-80 kD for Ch-Cpn20 (Hirohashi et al., 1999; Koumoto et al., 1999).

In vitro, Ch-Cpn20 was capable of assisting GroEL and chloroplast Cpn60 in refolding denatured bacterial RuBisCO (Bertsch et al., 1992; Dickson et al., 2000; Viitanen et al., 1995). Given the unusual binary organization of the Ch-Cpn20, it was anticipated that Ch-Cpn20 would be an obligate co-chaperone for Ch-Cpn60. However, both GroES and mouse mitochondrial-cpn10 were fully interchangeable with the chloroplast homologue (Viitanen 1995). Ch-Cpn20 from *A. thaliana* is 61% identical to that of spinach. Two Cpn10 (GroES)-like domains of *Arabidopsis* cpn20 are 46% identical to each other (Hirohashi et al., 1999). The functional Ch-Cpn20 seems to form a homotetramer with a ring-like toroidal structure (Figure 6B) (Baneyx et al., 1995; Hirohashi et al., 1999; Koumoto et al., 1999; Sharkia et al., 2003). Ch-Cpn20 forms assymetric complexes with Ch-Cpn60 similar to bacterial chaperonins (Figure 6C) (Viitanen, 1995). *A. thaliana* Ch-Cpn20 was shown to be heat inducible, defining it as a heat shock protein (Koumoto et al., 1999). It was

observed that Ch-Cpn20 can bind in a calcium-dependent manner to Calmodulin, (a ubiquitous Ca^{2+} signal transducer in eukaryotes) (Yang and Poovaiah, 2000). However, little is known about the structural and functional details of Ch-Cpn20 in plants.

Until recently, it was thought that chloroplasts contain only the double-domain co-chaperones. However, recent research provided the information that *Arabidopsis thaliana* chloroplasts may contain at least one normal-sized Ch-Cpn10, in addition to Ch-Cpn20. Ch-Cpn10 seems to be expressed in later stages of plant development than Ch-Cpn20 and its expression is less stimulated by light in comparison to Ch-Cpn20. Northern blot analysis revealed that Ch-Cpn10 is present in leaves and stems, but not in roots (Yasuko Koumoto, 1999). The requirement of these two different co-chaperones in the different stages of plant growth and development is yet to be understood.

Chloroplast chaperonin 60 (Ch-Cpn60) and co-chaperone 10 (Ch-Cpn10) function together in an ATP-dependent manner to facilitate the folding, assembly and translocation of numerous other proteins (Boston et al., 1996; Hartl, 1996a)

2.4.8   Assembly chaperones

In addition to chaperones that assist in protein folding, chaperones assisting in the assembly of protein complexes have been reported in recent times (Ellis, 1987; Ellis, 2006). These include the chaperones in the assembly of nucleosomes and proteosomes.

2.5   Photosynthesis

Photosynthesis is the physico-chemical process by which plants, algae and photosynthetic bacteria convert light energy into the chemical energy in the form of sugars and other organic compounds (Gest, 1993). In plants, algae and certain

types of bacteria, photosynthesis results in the release of molecular oxygen and the removal of carbon dioxide from the atmosphere which will be used to synthesize carbohydrates (oxygenic photosynthesis). Other types of bacteria use light energy to create organic compounds, but do not produce oxygen (anoxygenic photosynthesis).

The photosynthetic process consists of a series of chemical reactions that require carbon dioxide (CO_2) and water (H_2O). Oxygen (O_2) is a byproduct of photosynthesis and is released into the atmosphere (Figure 7A). The following equation summarizes photosynthesis:

$$6\ CO_2 + 6\ H_2O \rightarrow 6(CH_2O)\ (sugar) + 6\ O_2$$

Photosynthesis transfers electrons from water to energy-poor CO_2 molecules, forming energy-rich sugar molecules. This electron transfer is an example of an oxidation-reduction process: the water is oxidized (loses electrons) and the CO_2 is reduced (gains electrons).

2.5.1　Oxygenic photosynthetic organisms

The photosynthetic process in all plants and algae as well as in certain types of photosynthetic bacteria involves the reduction of CO_2 to carbohydrates and removal of electrons from H_2O, which results in the release of O_2. In this process, known as oxygenic photosynthesis, water is oxidized by the photosystem II reaction center. Oxygenic photosynthesis involves about 100 proteins that are highly ordered within the photosynthetic membranes of the cell (Allen and Martin, 2007). The structure and function of photosystem II is similar in plants, algae and certain bacteria.

2.5.2   Anoxygenic photosynthetic organisms

Some photosynthetic bacteria can use light energy to extract electrons from molecules other than water. These organisms are of ancient origin, presumed to have evolved before oxygenic photosynthetic organisms. Anoxygenic photosynthetic organisms occur in the domain Bacteria and have representatives in four phyla - Purple Bacteria, Green Sulfur Bacteria, Green Gliding Bacteria, and Gram Positive Bacteria.

2.5.3   Photosynthetic process

When light is absorbed by a green plant, a small portion of that energy is converted into chemical energy in the process of photosynthesis. Leaf structure is closely associated with its photosynthetic function. Leaves must permit carbon dioxide access to the photosynthetic cells but impede water from diffusing out. The oxygen which is a waste product of photosynthesis must be allowed to escape from the leaf.   Mesophyll cells contain chloroplasts, the organelles that perform photosynthesis and so are specialized for photosynthesis (Figure 7A). These double-membrane bound organelles enclose additional membranes called thylakoids. The disc-shaped thylakoids possess an interior space. The thylakoids are stacked to form grana, which are suspended in the stroma of the chloroplasts (Staehelin, 1986) (Figure 7B).

2.5.4   Photosynthetic pigments

Pigments are molecules that absorb light. When a photon of light strikes a photosynthetic pigment, an electron in an atom contained within the molecule becomes excited. Energized electrons move further from the nucleus of the atom. The excited (energized) molecule can pass the energy to another molecule or release it in the form of light or heat.

Chlorophyll molecules embedded in the thylakoid membrane absorb light energy. These molecules are the most important pigments for absorbing the light energy used in photosynthesis. A chlorophyll molecule has a hydrophobic "tail" that embeds the molecule into the thylakoid membrane. The "head" of a chlorophyll molecule is a ring called a porphyrin, which has a magnesium atom at its center, and is the part of a chlorophyll molecule that absorbs light energy (Figure 7C).

Chlorophyll-a is the main photosynthetic pigment in all organisms, except bacteria. Other pigments called accessory pigments absorb slightly different wavelengths of light. The combination of all of the pigments increases the range of light that plants can use in photosynthesis. Accessory pigments include chlorophyll-b and carotenoids. They do not participate in photosynthetic reactions directly but are able to pass their energy to chlorophyll-a.

2.5.5   Photosystems

The closely packed pigment molecules and the reaction center form a unit referred to as an antenna complex. Photons of light that are picked up by any of the pigment molecules in the antenna pass their energy to nearby pigment molecules until it is eventually passed on to a special molecule of chlorophyll-a called the reaction center. The reaction center molecule becomes ionized and it loses its electron to an electron acceptor. This electron will need to be replaced.

The antenna, the reaction center, and the electron transport molecules make up a photosystem. There are two kinds of photosystems in eukaryotes. The reaction center chlorophyll molecule of photosystem I absorbs 700 nm light best and is therefore called P_{700}. The reaction center of photosystem II absorbs 680 nm light best and is called P_{680} (Figure 8B). Photosystem I evolved very early; photosystem II evolved later.
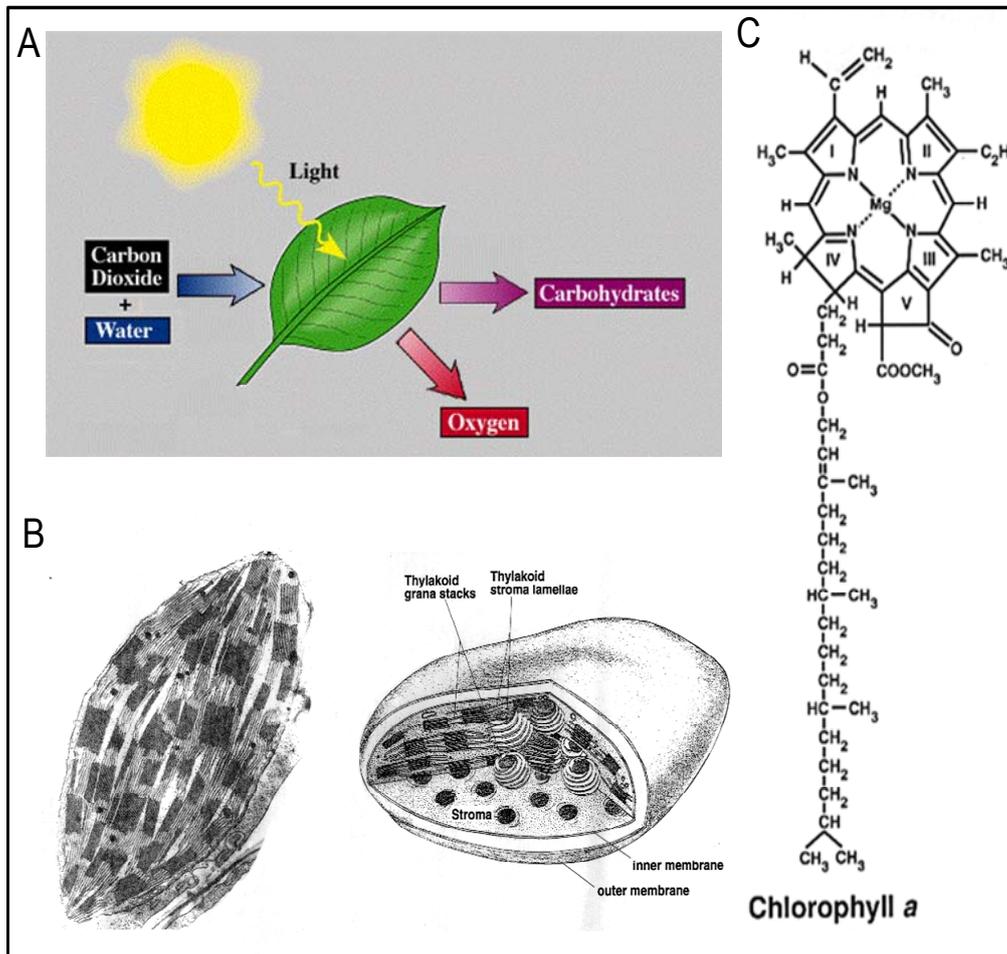
 Figure 7. Photosynthesis

(A) Photosynthesis in plants. In photosynthesis, light energy is converted into chemical energy in the form of sugars and other organic compounds.
(B) Electron micrograph showing structure of chloroplast (Ort, 1994). The chloroplast is about 6 microns long. Inside the chloroplast is the photosynthetic membrane, which is organized into stacked and unstacked regions. The stacked regions are linked by unstacked membranes. (C). Chemical structure of chlorophyll-a molecule.

Photosysyem I and II transform sunlight into electrical current with the help of chlorophyll (Figure 8B). Photosystem II generates an electrochemical potential of +1.1 volts, enough to remove two electrons from each of two water molecules, making a molecule of O_2 at a cost of four photons — one for each electron moved. Photosystem II performs this remarkable feat only when photosystem I is present

to dispose of the electrons. Photosystem I grabs the four electrons and uses four more photons to deposit them, in two pairs, on an electron carrier called $NADP^+$. Ultimately $NADP^+$ transfers the electrons to carbon dioxide, thereby providing the energy to make carbon-based sugars and the other molecules (Allen and Martin, 2007) (Figure 8).
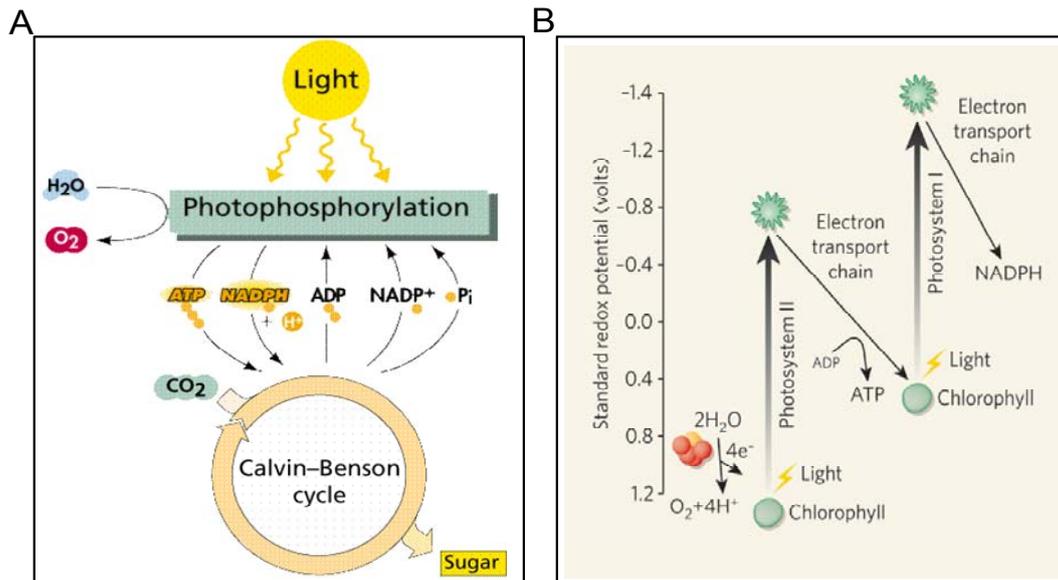
Figure 8. Photosynthesis and Photosystems

(A) Overview of photosynthesis showing integration of the light reactions (light harvesting, water splitting, electron transport and photophosphorylation) with the dark reactions (fixation of CO_2 and reduction to sugar). (B). The Z scheme of photosynthetic electron transport: Light energy captured by chlorophyll molecules actually involves the interplay of both photosystem I and photosystem II that are linked together by redox reactions. Photosystems I and II absorb light energy, convert it into electrochemical potential, and are connected in series electrically. These two 'light reactions' of photosynthesis form links in a chain of electron (e^-) transfers that is coupled, by means of proton pumping, to synthesis of the energy-storage molecule adenosine triphosphate (ATP). The electron transport chain of photosynthesis, also known as the Hill and Bendall Z-scheme, ends with photosystem I delivering electrons to $NADP^+$, making NADPH. ATP and NADPH drive the 'dark reactions' that transfer the electrons to CO_2 so as to provide the energy to make sugars and the other molecules of life. The chain begins when water is oxidized to oxygen by the very high electrochemical potential of photosystem II. The catalyst of water oxidation is shown here as a cluster of four red spheres and one yellow one (Allen and Martin, 2007).

2.5.6   Light reactions and the Calvin cycle

Photosynthesis depends on an interaction between two sets of reactions: the light reactions and the Calvin cycle (also called as Calvin-Benson-Bassham cycle) (Calvin, 1989). Chlorophyll and the other molecules responsible for the light reactions are located in thylakoid membranes. The enzymes that catalyze the Calvin cycle are located in the stroma. Beginning with the absorption of light by chlorophyll, the light reactions convert light energy into chemical energy in the form of ATP and NADPH. The ATP provides the energy, and the NADPH supplies the electrons for the Calvin cycle, which converts carbon dioxide to sugar (Calvin, 1989) (Figure 9 & 8A). The ADP and $NADP^+$ that result from the Calvin cycle shuttle back to the light reactions, which regenerate ATP and NADPH (Figure 8A).

All plants and algae remove CO_2 from the environment and reduce it to carbohydrates by the Calvin cycle (Figure 9). The process is a sequence of biochemical reactions that reduce carbon and rearrange bonds to produce carbohydrate from CO_2 molecules. The first step is the addition of CO_2 to a five-carbon compound (ribulose 1,5-bisphosphate). The resulting six-carbon compound is then split to give two molecules of a three-carbon compound (3-phosphoglycerate). This key reaction is catalyzed by RuBisCO (Ribulose-1,5-bisphosphate carboxylase/oxygenase), a large water soluble protein complex. The main energy input in the Calvin cycle is the phosphorylation by ATP and subsequent reduction by NADPH of the initial three-carbon compound forming a three-carbon sugar, triosephosphate. Some of the triosephosphate is exported from the chloroplast and provides the building block for synthesizing more complex molecules. In a process known as regeneration, the Calvin cycle uses some of the triosephosphate molecules to synthesize the energy rich ribulose 1,5-bisphosphate needed for the initial carboxylation reaction. This reaction requires the input of energy in the form of one ATP (Figure 9).

Overall, thirteen enzymes are required to catalyze the reactions in the Calvin cycle. The energy conversion efficiency of the Calvin cycle is approximately 90%. The reactions do not involve energy transduction, but rather the rearrangement of chemical energy. Each molecule of CO_2 reduced to a sugar $[CH_2O]_n$ requires 2 molecules of NADPH and 3 molecules of ATP.

RuBisCO is a bifunctional enzyme that, in addition to binding CO_2 to ribulose bisphosphate, can also bind O_2. This oxygenation reaction produces 3-phosphoglycerate that is used in the Calvin cycle and a two-carbon compound (2-phosphoglycolate) which is not useful for the plant. In response, a complicated set of reactions known as photorespiration is initiated which utilizes the fixed carbon. Thus RuBisCO oxygenation reaction appears to serve no useful purpose for the plant.
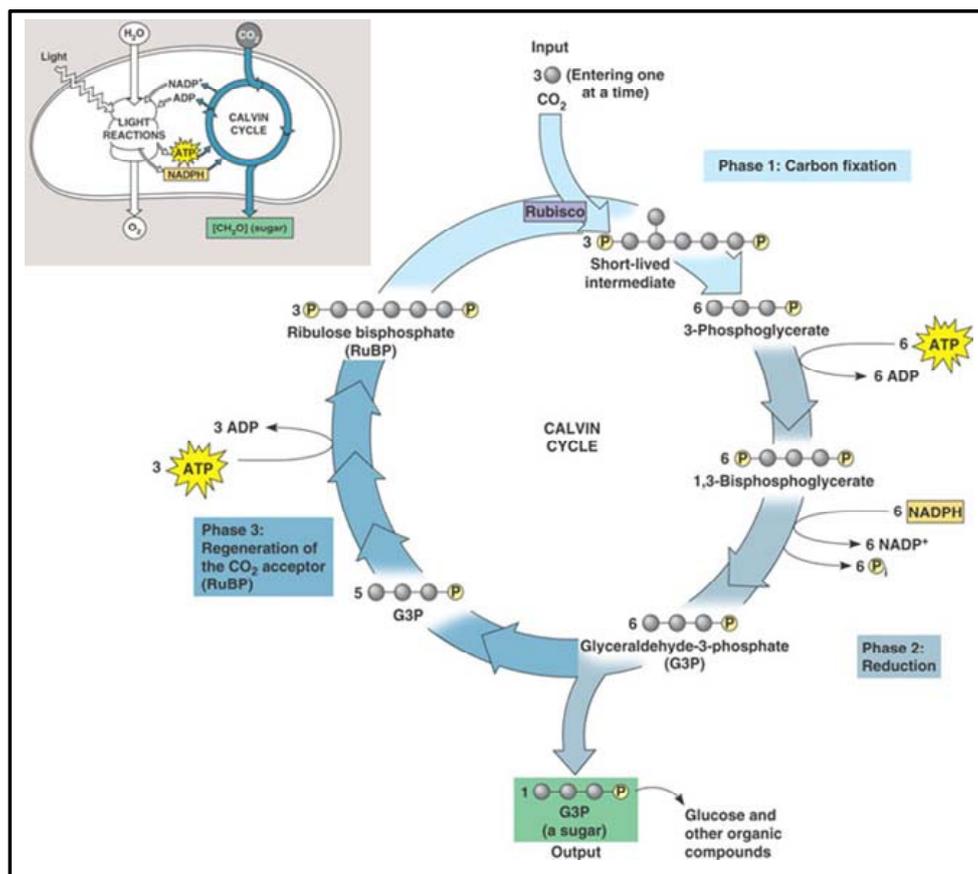
Figure 9. Scheme showing reduction of carbon dioxide by the Calvin cycle.

The first step is carboxylation, in which Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) catalyzes the addition of CO2 to the five-carbon compound, ribulose 1,5-bisphosphate, which is subsequently split into two molecules of the three-carbon compound, 3-phosphoglycerate. Next are reduction and phosphorylation reactions that form the carbohydrate, triose phosphate. Some of the triose phosphate molecules are used to form the products of photosynthesis, sucrose and starch, while the rest is used to regenerate ribulose 1,5-bisphosphate needed for the continuation of the cycle (Calvin, 1989).

2.5.7   C4 plants and the CAM Pathway

Some plants have evolved specialized structures and biochemical pathways that concentrate CO2 near RuBisCO. These pathways (C4 and CAM), serve to decrease the fraction of oxygenation reactions. Plants like crabgrass, corn, and sugar cane belong to this group. These plants capture CO_2 in a different way: they do an extra step first, before performing the Calvin cycle. These plants have a special enzyme that can work better, even at very low CO_2 levels, to capture CO2 and turn it first into oxaloacetate, which contains four carbons. Thus, these plants are called C_4 plants. The CO_2 is then released from the oxaloacetate and put into the Calvin cycle.

There is yet another strategy to cope with very hot, dry, desert weather and conserve water. Some plants (for example, cacti and pineapple, *Sedum*) that live in extremely hot, dry areas like deserts, can only safely open their stomata at night when the weather is cool. Thus, there is no chance for them to get the CO_2 needed for the dark reaction during the daytime. At night when they can open their stomata and take in CO_2, these plants incorporate the CO_2 into various organic compounds to store it. In the daytime, when the light reaction is occurring and ATP is available (but the stomata must remain closed), they take the CO_2 from these organic compounds and put it into the Calvin cycle. These plants are called CAM plants, which stands for crassulacean acid metabolism.

2.5.8 Diversity among photosynthetic organisms

Microbes are the predominant photosynthetic organisms in most aquatic environments. In aerobic conditions, (e.g. shallow water), algae, diatoms and cyanobacteria are predominant. In anaerobic conditions (polluted or eutrophic waters), other photosynthetic bacteria are dominant.

Cyanobacteria such as *Spirulina* are also photoautotrophs, able to use CO_2 as their sole carbon source and light as their energy source and are closer than any other bacteria to plant-style photosynthesis. It is believed that plants evolved directly from cyanobacteria. Other photosynthetic bacteria contain different pigments referred to as bacteriochlorophylls, and they do not produce oxygen.

Photosynthesis in *Rhodospirillales* (phototrophic bacteria):

Purple sulphur bacteria must fix CO_2 to live, whereas non-sulphur purple bacteria can grow aerobically in the dark by respiration on an organic carbon source.

Rhodospirillaceae: The purple non-sulphur bacteria, e.g. *Rhodospirillum* cells, contain bacteriochlorophyll-a or b. They are not able to use elemental sulphur as electron donor and typically use an organic electron donor, such as succinate or malate, but can also use hydrogen gas.

Chromatiaceae: Purple sulphur bacteria, e.g. *Chromatium*. They are able to use sulphur and sulphide as the sole photosynthetic electron donor and sulphur can be oxidized to sulphate. These bacteria use an inorganic sulphur compound, such as hydrogen sulfide as an electron donor.

Chlorobiaceae: Green sulphur bacteria. These cells contain bacteriochlorophyll c or d located in chlorobium vesicles attached to the cytoplasmic membrane.

2.6 RuBisCO

As mentioned earlier, RuBisCO (Ribulose-1,5-bisphosphate carboxylase/ oxygenase), catalyzes the carboxylation of ribulose bisphosphate in the Calvin cycle, a cycle through which most of the carbon fixed on Earth is processed. RuBisCO catalyzes the primary photosynthetic CO_2 reduction reaction, the fixation of atmospheric CO_2 to ribulose-1, 5-bisphosphate (RuBP) to form two molecules of 3-phosphoglycerate (3PGA) (Figure 10). RuBisCO is found in most autotrophic organisms, including photosynthetic bacteria, cyanobacteria, algae and plants. RuBisCO is the most abundant protein in the biosphere (Ellis, 1979); it constitutes up to 50% of the soluble protein in the leaf of C3 plants (Spreitzer and Salvucci, 2002) and ~ 30% in C4 plants (Sugiyama et al., 1984).

But, RuBisCO is the slowest metabolic enzyme in the contemporary biosphere (kcat ~3/s) (Mann, 1999; Parikh et al., 2006)**.** Unlike most enzymes, RuBisCO is not substrate specific - it also has an oxygenase function (Photorespiration). Photorespiration is an apparently wasteful reaction since it leads to a loss of 20-50% of fixed carbon (Mann, 1999) and it consumes energy (Figure 10). Net photosynthetic yield and in turn, yield (agronomic performance) of plants depends on these competing reactions. RuBisCO is a poor catalyst, has low affinity for carbon. It has a small turn-over number that is among the lowest for any known biological catalyst (Cleland et al., 1998; Tabita, 1999). This means that relatively large amount of enzyme must be present to sustain sufficient rates of photosynthesis. Despite major advances towards a comprehensive understanding of the structure, mechanism, physiology and molecular biology of RuBisCO, the feasibility of designing a more efficient enzyme remains problematic.
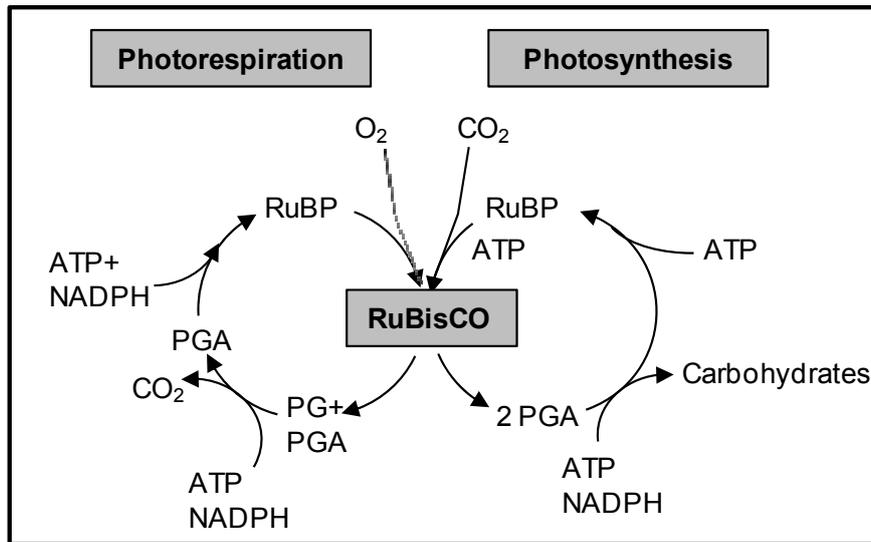
Figure 10. Photorespiration and Photosynthesis
Photosynthesis occurs when RuBP is carboxylated by RuBisCO and two molecules of phosphoglyceric acid (PGA) are processed into carbohydrates and used to regenerate RuBP in the reaction sequences requiring ATP and NADPH. Photorespiration begins with the oxygenation of RuBP to form one phosphoglycolate (PG) and PGA, in a side reaction catalyzed by RuBisCO. Processing the PG to PGS and eventually RuBP requires ATP and NADPH.

2.6.1   Molecular forms of RuBisCO

Based on phylogenetic analyses, RuBisCO proteins have been classified into form I, II and III, all of which catalyze the same reactions (Tabita et al., 2008). In addition, a homologue of RuBisCO called the RuBisCO-like protein (RLP) has been reported very recently, belonging to form IV. RLPs cannot catalyze RuBP carboxylation or oxygenation, mainly because of key substitutions of many essential active-site residues (Hanson and Tabita, 2001).

Form I RuBisCO:

Form I, the predominant RuBisCO, exists in higher plants, eukaryotic algae, cyanobacteria and most phototrophic and chemolithoautotrophic proteobacteria (Tabita, 1999; Tabita et al., 2007). Form I RuBisCO has a hexadecameric structure, being composed of eight large subunits (~55 kD each) and eight small

subunits (~15 kD each) (Figure 11i, B and C). In higher plants the large subunits are encoded by the plastome (*rbcL* gene) and the small subunits are encoded by the nuclear genome (*rbcS* gene). RuBisCO assembly takes place in the chloroplast.

Form I RuBisCO consists of a core of four $RbcL_2$ dimers arranged around a 4-fold axis, capped at each end by four small subunits (Knight et al., 1990) (Figure 11i, B and C). A further classification of form I enzymes include green-type enzymes (forms IA and IB from cyanobacteria, eukaryotic algae, and higher plants) and red-type enzymes (forms IC and ID from non-green algae and phototropic bacteria) (Delwiche and Palmer, 1996; Tabita, 1999). The crystal structures have been determined for hexadecameric RuBisCO from tobacco (Curmi et al., 1992; Schreuder et al., 1993), spinach (Knight et al., 1990), from a cyanobacterium-*Synechococcus* PCC6301 (Newman and Gutteridge, 1993) (Figure 12A, 12B & 14), *Chlamydomonas reinhardtii* (Mizohata et al., 2002; Taylor et al., 2001), *Ralstonia eutropha* and *Galdieria partita* (Hansen et al., 1999a; Sugawara et al., 1999).

Form II RuBisCO:

Form II RuBisCO was originally found in the purple non sulphur bacterium *Rhodospirillum rubrum* but has also been found in several chemoautotropic bacteria (Shively et al., 1998) and in eukaryotic dinoflagellates (Morse et al., 1995; Rowan et al., 1996; Whitney et al., 1995). It is a dimer of large subunits $(RbcL_2)_n$ and shares only 25–30% identity to large subunits from form I RuBisCO. The structure of form II RuBisCO closely resembles that of the basic dimer of the form I RuBisCO despite some differences in catalytic properties (Schneider et al., 1990) (Figure 11i, D and A).
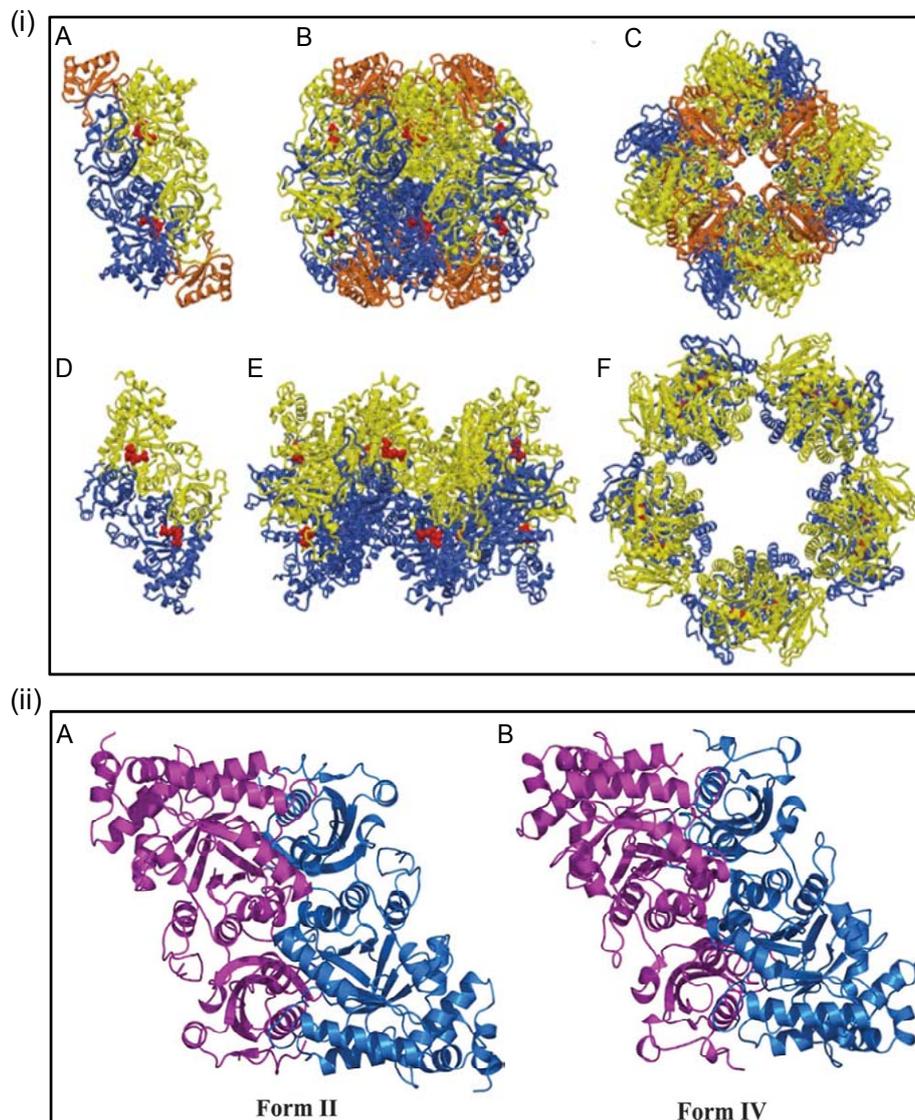
Figure 11. Structures of different forms of RuBisCO.
All forms are comprised of dimers of catalytic large subunits.
(i) (A) $RbcL_2S_2$ unit of Form I RuBisCO from spinach viewed down the two-fold symmetry axis; (B and C) Entire $RbcL_8S_8$ hexadecamer viewed down the two-fold and four-fold axes. (D) Dimeric Form II RuBisCO from *Rhodospirillum rubrum* showing the two-fold symmetry; (E and F) $RbcL_{10}$ RuBisCO from *Thermococcus kodakaraensis* viewed down the two-fold and five-fold axes, respectively (Andersson and Taylor, 2003).
(ii) Comparison of Form II and Form IV RuBisCO. (A) Form II is comprised of dimers of large subunits, ranging from $RbcL_2$–$RbcL_8$ depending on the source. (B) Form IV (the RuBisCO-like Protein or RLP) appears to have an RbcL structure (Li et al., 2005; Tabita et al., 2007; Tabita et al., 2008).

Rhodobacter sphaeroides, R. capsulatus, Thiobacillus sp., and *Hydrogenovibrio marinus* which belong to non-sulphur phototropic bacteria posses both form I and form II enzymes (Gibson and Tabita, 1977; Hayashi et al., 1997).

Form III RuBisCO:

Archaeal RuBisCOs have been classified into form III. They comprise $RbcL_2$, $RbcL_8$, and $RbcL_{10}$ enzymes and differ in the quaternary structure from other forms. The crystal structure of RuBisCO from *Thermococcus kodakaraensis* (Kitano et al., 2001) shows a pentagonal structure composed of five $RbcL_2$-units (Figure 11i, E and F) and that of *Pyrococcus horikoshii* consists of an octamer of large subunits, $RbcL_8$ (Andersson, 2008).

Form IV RuBisCO:

RuBisCO like proteins (RLPs), reported very recently, belong to form IV. *Chlorobium tepidum* is autotrophic but does not assimilate CO_2 via the Calvin cycle and *B. subtilis* does not use CO_2 as a carbon source at all (Tabita et al., 2007). RLPs lack several of the substrate binding and catalytic residues and do not catalyze RuBP-dependent CO_2 fixation *in vitro*. The overall structure of the RLP of *C. tepidum* is similar to the structures of the three other forms of RuBisCO (Figure 11ii, A and B); however, the active site is distinct from those of bona fide RuBisCOs and suggests that the RLP is possibly capable of catalyzing enolization but not carboxylation (Li et al., 2005). RLP from *C. tepidum* shows about 35% amino acid sequence identity with the other forms of RuBisCOs (Li et al., 2005). It was found that RLP of *C. tepidum* has a role in sulfur metabolism (thiosulfate oxidation) (Li et al., 2005) and RLP of *B. subtilis* participates in a methionine salvage pathway and catalyzes the enolization of the RuBP analog 2,3-diketo-5-methylthiopentyl-1-P (Imker et al., 2007; Tabita et al., 2007).
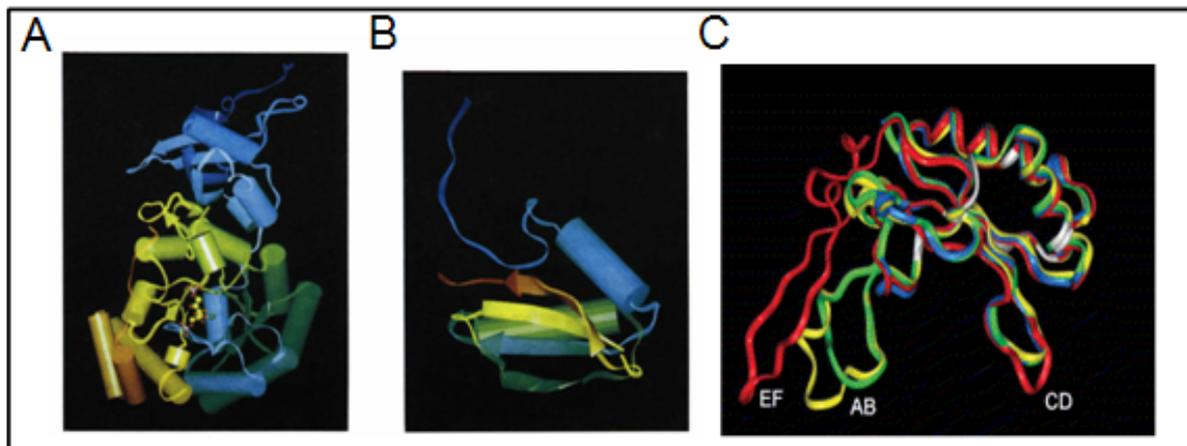
Figure 12. *Synechococcus* sp. PCC6301 RuBisCO large and small subunits
(A) The *Synechococcus* sp. PCC6301 RuBisCO large subunit has two domains; N-terminal domain consists of a five-stranded mixed B-sheet with two helices covering one side of the sheet. A third small helix makes up part of the connection between the domains. The C-terminal domain is an eight-stranded β/α-barrel with an extension containing two short β-strands between $\alpha6$ and $\beta37$. There are further two helices at the C terminus. The N terminus is shown in blue and the C terminus in red colors. The position of the active site is shown by the sugar molecule (Newman et al., 1993).
(b) The small subunit is a four-stranded antiparallel β-sheet covered on one side by two helices. N-terminus is shown in blue and the C-terminus in red (Newman et al., 1993)
(C) Comparison of Stereo images of the X-ray crystal structures of RuBisCO small-subunits from *Chlamydomonas* (yellow) (Taylor et al., 2001), spinach (green) (Andersson, 1996), *Synechococcus* (blue) (Newman and Gutteridge, 1993), and *Galdieria partita* (red) (Sugawara et al., 1999). Loops are labeled relative to their flanking β strands. Residues that are more than 95% conserved among all known small-subunit sequences are colored white.

2.6.2 Active site

As mentioned above, the large subunits of all the four forms display 25-30% sequence identity (Andersson, 2008). The active site is located at the intra-dimer interface between the carboxy-terminal domain of one large subunit and the amino-terminal domain of the second large subunit in the $RbcL_2$ dimer (Figure 13A). In fact, there are eight active sites in one $RbcL_8S_8$ molecule on the external part of the cube-like structure. In the hexadecameric RuBisCO, the dimers are arranged such that the eight active sites face the outside solvent (Figure 11i, B) (Knight et al., 1990).
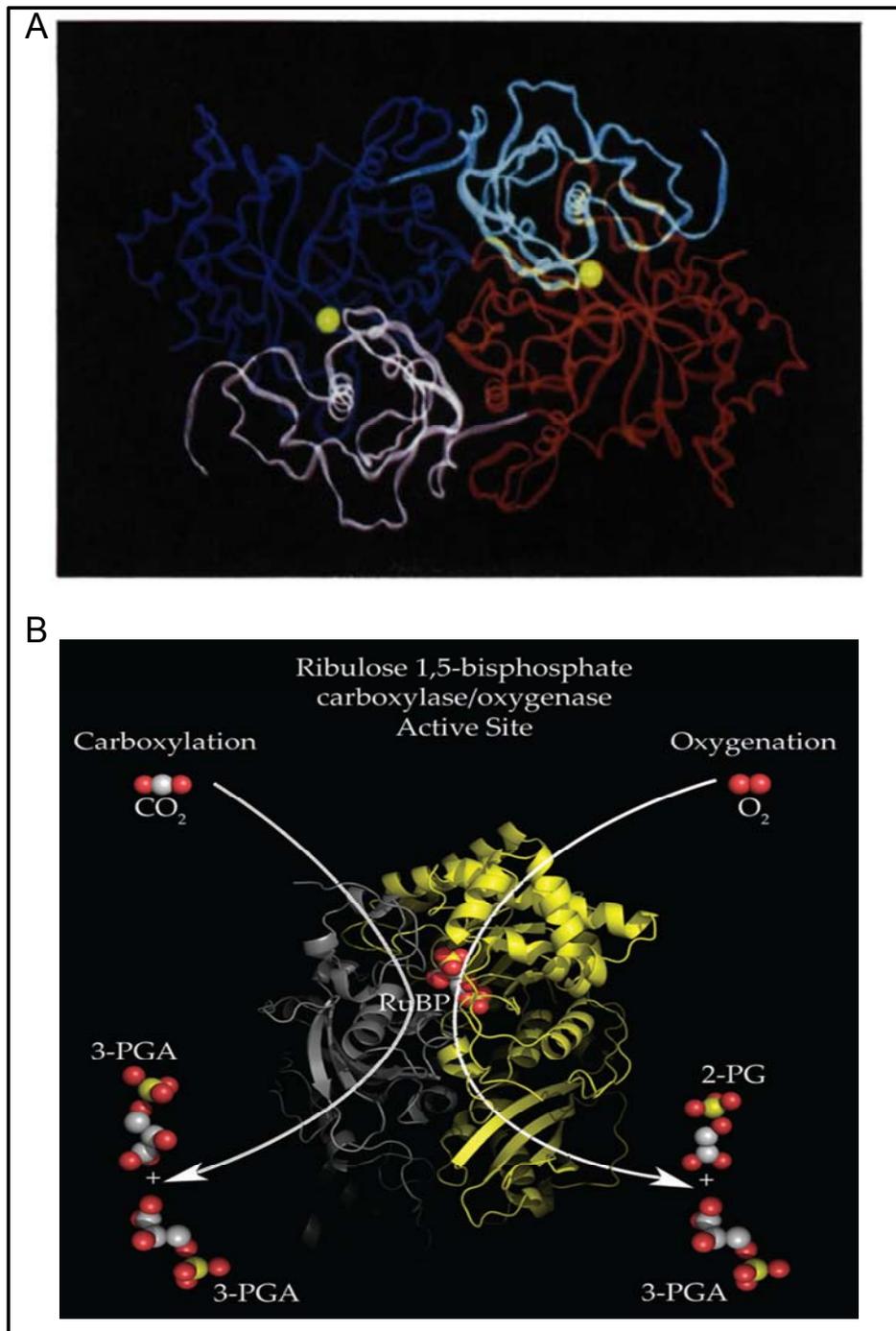
Figure 13. Structure of RuBisCO large subunits.

(A) Ribbon diagram of a RuBisCO dimer of large subunits. The barrel of one large subunit is colored in blue, and the N-terminal domain of the same subunit in cyan. The barrel domain of the second large subunit is in red, and its N-terminal domain is in magenta. The yellow spheres are Mg^{2+} ions, and show the positions of the active sites (Newman et al., 1993)**.**

(B) The fundamental dimeric unit of all forms of RuBisCO. The basic catalytic unit of all RuBisCO molecules is a dimer of large catalytic subunits in which each monomer interacts in a head to tail arrangement such that the C-terminus of one monomer interacts with the N-terminus of the second monomer. The active site thus contains residues primarily from the C-terminus of one monomer, but also includes residues from the N-terminus of the second monomer, conferring two active sites per dimer (Tabita et al., 2008).

Loops 1, 2, and 5–8 (connecting β-strand 1 with α-helix 1, etc) at the carboxy-terminal end of the β-strands contribute residues involved in catalysis and substrate binding. Two loop regions in the amino-terminal domain of the second large subunit in the dimer contribute additional residues to the active site. Thus, the functional unit of RuBisCO is an $RbcL_2$ dimer of large subunits containing two active sites (Figure 13B). The substrate binds in an extended conformation across the opening of the α/β barrel and is anchored at two distinct phosphate-binding sites at opposite sides of the α/β-barrel and in the middle at the magnesium-binding site.

2.6.3 Information from spinach RuBisCO structure

The large subunit of RuBisCO from spinach possesses two clearly separated domains (N and C) (Figure 14B). The N-domain (1–150 residues) is folded into a central mixed five-stranded β-sheet with two α-helices on one side of the sheet. The C-terminal domain is an α/β barrel (157–475 residues) and was found in the structures of many enzymes. The barrel has 8 β-strands which form the core of a barrel surrounded by the 8 α-helices. Loops of aminoacid sequences connect the β-strands to the α-helices (Figure 14B). The majority of aminoacid residues involved in catalysis and substrate binding are situated in the C-terminal loops (connecting the C-termini of the β-strands to the N-termini of the α-helices). The N-terminal loops are involved in subunit interactions in the $RbcL_8S_8$ molecule (Knight et al., 1990). As mentioned above, the active site is located at the intra-dimer interface between the C-terminal domain of one large subunit and the N-terminal domain of the second large subunit (Figure 14B & 14C) and so there are eight

active sites in one RbcL$_8$S$_8$ molecule on the external part of the cube-like structure (Figure 14A).

The sites of activation, of substrate binding and of some amino acids participating directly in catalytic reaction have been determined. In the α/β-barrel of the C-domain of large subunit, in loop 2, the active site of the enzyme containing Lys, Asp and Glu has been established. Arg which binds one of the phosphate groups of the substrate, and His, which is involved in catalysis, have been found in loop 5. Loop 6 is flexible, has an open and closed conformation and is very important for enzyme activation (Figure 14B & 14C). Ser binding RuBP and the product of the reaction as well as carboxyarabinitol bisphosphate (CABP) is located in loop 7. Loop 8 (Gly-rich helix) is the second phosphate binding site. In the N-domain there are sites which participate in substrate-enzyme binding (Knight et al., 1990) (Figure 14B).

2.6.4 RuBisCO small subunits

Small subunits are more divergent than large subunits (Spreitzer and Salvucci, 2002). Whereas land-plant and green-algal small subunits generally have larger loops between β strands A and B, some prokaryotes and all nongreen algae have longer C-terminal extensions (Figure 12C). Small subunits are not necessary for the assembly of the RbcL$_8$ core. But the small subunits maintain the structural integrity of holoenzyme by establishing dominant contacts between the four RbcL$_2$ dimers (Andrews, 1988; Andrews and Ballment, 1983; Curmi et al., 1992; Lee et al., 1991; Spreitzer and Salvucci, 2002).

The large subunit of form-II RuBisCO contains an insertion of 5 residues in the loop connecting α helix 3 to β strand 4 at the bottom of the α/β-barrel (I. Andersson, 1989). These residues are absent in the large subunits of form-I enzymes (in plants and green algae). But this has been complemented by the loop between β strands A and B of the small subunit (Cleland et al., 1998; Mizohata et al., 2002; Schreuder

et al., 1993). Among the form I RuBisCOs the small subunits differ in their βA–βB loop (Figure 12C) (Andersson, 1996; Newman and Gutteridge, 1993; Taylor et al., 2001). This loop in plants consists of 22 residues, in green algae of 28 residues and in prokaryotes and nongreen algae of only 10 residues (Figure 12C). However, the small subunits of all nongreen algae, dinoflagellates, and some prokaryotes have two additional β strands (E and F) formed by a long carboxy terminus. The resulting βE–βF loop resides in the central-solvent channel of the holoenzyme, filling in the space in a neighboring small subunit that arises from the smaller βA– βB loop (Hansen et al., 1999b; Sugawara et al., 1999). All small subunits that can form βE–βF loops also have several fewer residues at the N-terminus and two extra residues in the βC–βD loop (Figure 12C).

In higher plants, small subunits are encoded by the rbcS gene/s in the nuclear genome. They possess a transit peptide and so after translation on cytosolic ribosomes, the small subunit precursor proteins have to be imported into the chloroplasts, where they are processed and folded to the native state (Gatenby and Ellis, 1990). RuBisCO assembly takes place in the chloroplast. In prokaryotes rbcL and rbcS genes exist on the same operon.
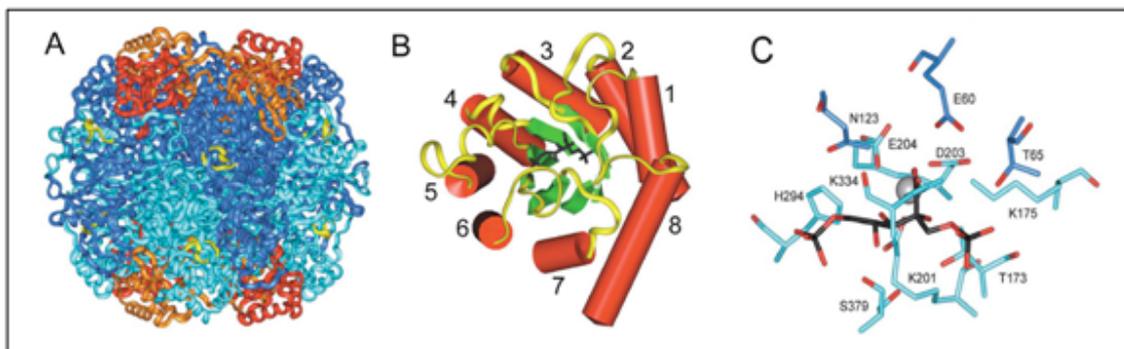
Figure 14. Stereo images of the X-ray crystal structure of spinach RuBisCO in relation to catalysis
(A) The holoenzyme is composed of eight large subunits (dark blue, light blue) and eight small subunits (red, orange). Active sites that form between two neighboring large subunits are denoted by loop 6 (yellow).

(B) The C-terminal domain of each large subunit forms an α/β barrel. Loops (yellow) between β strands (green) and α helices (red) contain residues that interact with the transition-state analog CABP (black).

(C) C-terminal domain residues (light blue) from one large subunit and N-terminal-domain residues (dark blue) from a neighboring large subunit interact with CABP (black). Mg2+ is denoted as a gray sphere. Oxygen atoms are colored red (Spreitzer and Salvucci, 2002).

2.6.5 RuBisCO folding/refolding attempts-role of chaperonins

The first evidence for chaperonin-interaction with the RuBisCO large subunits was found over 25 years ago when studies revealed that newly synthesized large subunits were transiently associated with a large multimeric complex (Milos, 1984). It is now well established that plant RuBisCO large subunits interact with chloroplast homologs of the GroEL and GroES for their folding (Drzymalla et al., 1996; Gatenby et al., 1990; Goloubinoff, 1989; Gutteridge, 1995; Lubben et al., 1989). The Form II RuBisCO (from *Rhodospirillum rubrum*) has been successfully reconstituted with the help of bacterial GroEL/GroES (Goloubinoff, 1989). After their folding in the chaperonin cage and release into solution, the subunits dimerize spontaneously to yield active dimer (Brinker et al., 2001).

The proposed maturation pathway for the assembly of $RbcL_8S_8$, form I RuBisCO, involves the folding of large subunits and dimerization of the folded monomers to yield $RbcL_2$ (dimer), then tetramerization of the dimers to yield a core of 8 large subunits ($RbcL_8$) to which 8 individual small subunits are spontaneously added to yield a $RbcL_8S_8$ holoenzyme (Andrews and Lorimer, 1985; Fitchen et al., 1990; Goloubinoff et al., 1989; Lee and Tabita, 1990; Paul et al., 1991). Among the Form I RuBisCO, the expression of RuBisCO from a cyanobacterium, *Synechococcus sp.* PCC6301 in *E. coli* was reported to result in the chaperonin dependent production of enzymatically active holoenzyme (Goloubinoff et al., 1989). Formation of dimers of RuBisCO large subunits ($RbcL_2$) and octomer core ($RbcL_8$) during the expression of the cyanobacterium *Anacystis nidulans* large subunits in *E. coli* has been reported (Lee and Tabita, 1990).

2.6.6 RbcX and its role

Recently it has been reported that the product of the *rbcX* gene, which is present in the intergenic space between the *rbcL* and *rbcS* genes in several cyanobacteria (Larimer and Soper, 1993) enhances the production of enzymatically active RuBisCO upon coexpression with *rbcL* and *rbcS* in *E. coli* (Emlyn-Jones, 2006; Li and Tabita, 1997; Onizuka et al., 2004).

When RuBisCO from *Anabaena* sp. CA was produced in *E. coli*, co-expression of GroEL/GroES ameliorated the recovery of active enzyme. Deletion of *rbcX* from the operon of *rbcLXS* of *Anabaena* sp. PCC 7120 decreased the production of active RuBisCO. RuBisCO activity increased about ~40 fold when Cpn60 and Cpn10 were cotransformed (Larimer and Soper, 1993). This means that large excess of chaperonins were needed for complementing RbcX in the formation of RbcL$_8$S$_8$, suggesting a chaperone-like role of RbcX (Larimer and Soper, 1993; Li and Tabita, 1997). Partial inactivation of the rbcX gene in *Synechococcus* sp. PCC7002 resulted in a substantial reduction in RuBisCO solubility and activity (Onizuka et al., 2004). However, the *rbcX* gene was reported to be nonessential in *Synechococcus* sp. PCC7942 (Emlyn-Jones, 2006).

The clustering of genes (example *rbcL*, *rbcX* and *rbcS*) is thought to assist in coordinating the synthesis and assembly of the various complex components (Shi et al., 2005; Tabita, 1999). Unlike the products of *rbcL* and *rbcS*, the *rbcX* product (~15.5 kDa) is not part of the final RuBisCO complex. Unlike RbcL and RbcS whose sequences are highly conserved by functional constraints, the RbcX sequence is highly variable amongst cyanobacterial species (Rudi et al., 1998). Importantly, the juxtaposition of *rbcX* within an *rbcLXS* operon is highly conserved in many cyanobacteria, suggesting that the RbcX product may function in a role associated with *rbcL* and *rbcS* consistent with the above mentioned results (Larimer and Soper, 1993; Li and Tabita, 1997; Onizuka et al., 2004).

Reconstitution *in vitro* using Form I RuBisCO has not been successful so far. *in vivo* expression studies involving Form I RuBisCO from many cyanobacteria except a few (e.g., *Synechococcus* sp. PCC6301), and any of the higher plants have yielded no fruitful results (Andrews and Lorimer, 1985; Chaudhari and Roy, 1989; van der Vies et al., 1986; Voordouw et al., 1984). Higher plant RuBisCO has never been produced as an active enzyme in *E. coli*, even in the presence of co-expressed chloroplast chaperonins (Cloney et al., 1993; Gatenby, 1984; Gatenby et al., 1987). The question now is why higher plant RuBisCO cannot assemble in *E. coli* into active enzyme, while cyanobacterial *Synechococcus* sp. PCC6301 RuBisCO can (Gurevitz et al., 1985; van der Vies et al., 1986).

3 Aim of the study

The efficiency of plant growth depends on photosynthetic yield. RuBisCO is a key factor for photosynthesis. As mentioned above, RuBisCO is inefficient as an enzyme as it participates in the side reaction of photorespiration (Mann, 1999; Tabita, 1999). The efficiency of RuBisCO varies among photosynthetic organisms; red algae RuBisCOs have higher specificity for CO_2 than those of higher plants (Andersson, 2008). One of the attempts to improve RuBisCO is through *in vitro* mutagenesis. However, a prerequisite for such attempts is the possibility to produce functional Form I RuBisCO upon recombinant expression in a bacterial host such as *E. coli*. The classical GroEL (Ch-Cpn60) and GroES (Ch-Cpn20/10) chaperones have been shown to assist in the folding of newly translated RbcL protein of Form II RuBisCO (Brinker et al., 2001; Goloubinoff, 1989). Among the Form I RuBisCO, the coexpression of *Synechococcus sp*. PCC6301 RuBisCO along with GroEL/GroES in *E. coli* was reported to result in the chaperonin dependent production of enzymatically active holoenzyme (Gatenby and Ellis, 1990; Gatenby et al., 1990; Goloubinoff et al., 1989). But these chaperones are apparently not sufficient for the reconstitution of RuBisCO from other cyanobaceria and higher plants (Andrews and Lorimer, 1985; Chaudhari and Roy, 1989; Cloney et al., 1993; van der Vies et al., 1986).

In some cyanobacteria, the *rbcX* gene is present in the intergenic space between *rbcL* and *rbcS*, while in others it does not exist in the same operon. The goal of this study was to analyze the assembly of Form I RuBisCO, based on recent findings that protein RbcX allows the improved production of cyanobacterial Form I RuBisCO (Emlyn-Jones, 2006; Larimer and Soper, 1993; Li and Tabita, 1997; Onizuka et al., 2004). However, the structure and mechanism of RbcX function had remained unclear. It was of interest, therefore, to determine the step at which RbcX is involved in RuBisCO biogenesis.

An extensive functional analysis of RbcX in the maturation of Form I RuBisCO was

carried out using *in vitro* transcription/translation assays, *E. coli* coexpression experiments and a range of biochemical techniques.

4   Materials and methods

4.1   Laboratory equipment

Abimed (Langenfeld, Germany): Gilson Pipetman (2 to 1000 µl).

Amersham Pharmacia Biotech (Freiburg, Germany): ÄKTA Explorer; SMART-System; chromatography columns: Mono-Q, HiTrap-Heparin, Superdex 200, Superose 6; resins: Source 30-Q; electrophoresis power supplies: EPS200, EPS600.

Amicon (Beverly, MA, USA): concentration chambers: Centriprep, Centricon.

Avestin (Mannheim, Germany): EmulsiFlex C5 homogenizer.

Beckmann Coulter GmbH (Krefeld, Germany): centrifuges (J6-MI, GS-6R, Avanti 30, Avanti J-25I, Optima LE-80K ultracentrifuge), spectrophotometers (DU640, DU800), LS 6500 multi-purpose scintillation counter.

Biometra (Göttingen, Germany): PCR-Thermocycler.

Bio-Rad (Munich, Germany): electrophoresis chambers: MiniProtean 2 and 3, electrophoresis power supply Power PAC 300.

Branson (Connecticut, USA): Sonifier cell disruptor B15.

Eppendorf (Hamburg, Germany): centrifuges: 5415C, 5417R; Thermomixer Comfort.

Forma Scientific (Marietta / OH, USA): Orbital Shaker 4581

Fuji (Tokyo, Japan): Phosphoimager FLA-2000; ImageReader LAS-3000.

Getinge (Getinge, Sweden): autoclave.

Hoefer Scientific Instruments (San Francisco, USA): SemiPhore blotting transfer unit.

HORIBA Jobin Yvon GmbH (München, Germany): Spex Fluorolog 3.

Mettler Toledo (Giessen, Germany): balances: AE160, AG285, PB602.

Millipore (Eschborn, Germany): deionization system MilliQ plus PF; Millex-HA filters (0.22 µm); vacuum filtration unit (0.22 µm).

MWG BiotechAG (Göttingen, Germany): gel documentation system BioCapt.

New Brunswick Scientific (Nürtingen, Germany): orbital shaker and incubator Innova 4430.

Raytest (Straubenhardt, Germany): AIDA gel imaging software version 3.5.

Savant/Thermoquest (Engelsback, Germany): Stacked Gel Dryer SGD300.

WTW (Weilheim, Germany): pH-Meter.

Wyatt Technology (Santa Barbara/CA, USA): FFF-MALS system, software ASTRA.

4.2 Materials

4.2.1 Chemicals

Chemicals and biochemicals used in this work were of *pro analysi* grade and purchased from Fluka (Deisenhofen, Germany), Calbiochem (Bad Soden, Germany), Merck (Darmstadt, Germany), Sigma-Aldrich (Steinheim, Germany), Roth (Karlsruhe, Germany), and Roche (Mannheim, Germany) unless stated otherwise.

Amersham Pharmacia Biotech (Freiburg, Germany): western blotting detection systems: ECL, chromatographic resins, [^{35}S] -Methionine, $NaH^{14}CO_3$.

BioMol (Hamburg, Germany): IPTG, HEPES.

BioRad (Munich, Germany): ethidiumbromide; Bradford Protein-Assay.

Cambrex Bio Science (Rockland / ME, USA): Sea Kem LE Agarose.

Difco (Heidelberg, Germany): Bacto tryptone, Bacto yeast extract, Bacto agar.

Fermentas (St. Leon-Rot, Germany): GeneRuler 1kb DNA Ladder.

Fluka (Deisenhofen, Germany): acetic acid, acetone, Bis-Tris, H_2O_2, luminol, PEG.

Hampton Research (Aliso Viejo / CA, USA): Crystallization screens.

Invitrogen (Karlsruhe, Germany): protein marker for SDS PAGE, dNTP set.

J.M. Gabler Saliter GmbH & Co. KG (Obergünzburg, Germany): skim milk powder.

New England Biolabs (Frankfurt a. Main, Germany): restriction enzymes; calf intestinal alkaline phosphatase (CIP); T4 DNA ligase; prestained protein marker for SDS PAGE.

Promega (Mannheim, Germany): Wizard SV Gel and PCR Clean-Up System, Wizard Plus SV Miniprep DNA Purification System, PureYield Plasmid Midiprep System.

Qiagen (Hilden, Germany): QIAprep Plasmid Mini and Midi kits; QIAquick PCR purification and gel extraction kit; Ni-NTA agarose.

Roche (Mannheim, Germany): RTS *in vitro* translation systems: RTS 100 *E. coli* HY Kit, Complete protease inhibitor cocktail, DTT.

Roth (Karlsruhe, Germany): ampicillin, glycin, scintillation fluid.

Schleicher & Schuell (Dassel, Germany): protran nitrocellulose transfer membrane; fluted paper filter 595-1/2 (270 mm).

Serva (Heidelberg, Germany): Acrylamide-Bis, BSA, Coomassie blue G/R, PMSF, SDS.

USB (Cleveland, USA): ammoniumsulfate, chloramphenicol, EDTA, MES, tricine, urea.

4.2.2 Strains

Novagen (Darmstadt, Germany): *E. coli* DH5α
Stratagene (Heidelberg, Germany): *E. coli* BL21(DE3)

4.2.3 Plasmids, DNA and oligonucleotides

Novagen (Darmstadt, Germany): pET11a, pET15b, pET28b, pET30b. Plasmid pHUE was a kind gift from Dr. Spencer Whitney, ANU, Canberra, Australia.

Metabion (Martinsried, Germany): oligonucleotides (primers).

Plasmids generated during this study will be described below.

4.2.4 Enzymes, proteins, peptides and antibodies

AgriSera (Vännäs, Sweden): anti-RbcL antibody.

Amersham Bioscienses (Freiburg, Germany): porcine RNAguard ribonuclease inhibitor, Protein A sepharose beads.

Merck (Darmstadt, Germany): Benzonase.

New England Biolabs (NEB, Frankfurt Am Main, Germany): restriction endonucleases, T4 DNA ligase.

Fermentas (St. Leon-Rot, Germany): restriction endonucleases.

JPT Peptide Technologies GmbH (Berlin, Germany): peptide array.

MPI for Biochemistry, Department Cellular Biochemistry (Martinsried, Germany): purified protein stocks of GroEL, GroEL-SR1 (R452E, E461A, S463A, V464A), GroES, DnaK, DnaJ, GrpE, Gp31, yeast mt-Hsp10, Rr-RbcL, Rr-RbcL(K168E), TEV-protease.

MPI for Biochemistry, peptide synthesis service (Martinsried, Germany): oligopeptides.

MPI for Biochemistry (Martinsried, Germany): antisera against purified Syn7002-RbcX and E. coli GroEL (produced in rabbits).

Pineda-Antikörper-Service (Berlin, Germany): anti-Cpn60α antibody, anti Cpn60β-antibody (peptide antibodies, produced in rabbits).

Promega (Mannheim, Germany): Pfu DNA polymerase.

Roche (Basel, Switzerland): ProteinaseK, shrimp alkaline phosphatase, hexokinase, MDH.

Sigma-Aldrich (Steinheim, Germany): BSA, lysozyme, 3x FLAG peptide, mouse monoclonal anti-FLAG M2 antibody, EZview Red ANTI-FLAG M2 Affinity Gel, HRP-conjugated secondary antibodies.

4.2.5 Media

LB medium: 10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, (+15 g/l agar for solid medium).

4.2.6 Buffers and stock solutions

PBS: 137 mM NaCl, 2.7 mM KCl, 8.4 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.4.

PonceauS: 0.2% PonceauS, 3% trichloracetic acid.

TAE: 40 mM Tris-Acetate, 1 mM EDTA, pH 8.0.

TBS: 50 mM Tris, 150 mM NaCl, pH 7.5.

TBST: 0.1% Tween-20 in TBS.

All other buffers and solutions were prepared as convenient stock solutions and either autoclaved or filter sterilized before usage, if applicable.

4.3 Molecular biology methods

All experimental methods used were performed according to "Molecular Cloning" (Sambrook, 1989) unless stated otherwise.

4.3.1 Plasmid purification

A single *E. coli* colony containing the plasmid of interest was inoculated in LB medium (supplemented with the appropriate antibiotic) and shaken overnight at 230 RPM at 37°C. Plasmid DNA was purified using the QIAprep Plasmid kits or Wizard Plus SV Miniprep DNA Purification System, according to the manufacturer`s protocol.

4.3.2 DNA analytical methods

DNA concentrations were measured by UV absorption spectroscopy at λ=260 nm. An optical density of OD_{260}=1 corresponds to approximately 50 µg/ml double stranded DNA. The absorbance ratio 260/280 nm for pure DNA should be approximately 1.85. Deviations from this value are indicating quality deficiencies caused by impurities, such as RNA or protein (Sambrook, 1989).

Agarose gel electrophoresis was performed with 0.8-2% TAE-buffered agarose gels supplemented with 10 µg/ml ethidiumbromide. Size fractionation of DNA fragments was carried out at 80-100 V in TAE buffer. Prior to electrophoresis, loading buffer (6x loading buffer: 60% glycerol, 0.25% bromphenol blue, 0.25% xylene cyanol FF) was added to the DNA samples to a 1x concentration.

Primers were purchased from Metabion (Martinsried, Germany); DNA sequencing was performed by Medigenomix GmbH (Martinsried, Germany) or Sequiserve (Vaterstetten, Germany) or MPI sequencing service.

4.3.3 PCR amplification

DNA was amplified using PCR (polymerase chain reaction) according to the standard protocol mentioned below. Modifications in the reaction setup and the running conditions were made when necessary.

Typical PCR reaction

DNA template	100-200 ng
Primers	20 pmol each
dNTPs	200 µM each
Pfu DNA Polymerase	3 U/50 µl
Polymerase buffer	1x

Typical PCR cycling conditions (30 cyles):

Initial denaturation	95°C for 2 min
Cycle denaturation	95°C for 30-60 s
Annealing	50-58°C for 30-60 s
Extension	72°C for 2-6 min (~40 s /1.0 kb of DNA)
Final extension	72°C for 10 min
Pause	4°C indefinite

4.3.4 DNA restriction digestions and ligations

DNA restriction digestions were performed according to the product instructions of the respective enzymes. Typically, a 60 µl preparative reaction containing the purified PCR product or plasmid DNA: restriction enzyme = 1µg :10U, 0.1 mg/ml BSA (if necessary) in the appropriate reaction buffer was used. In order to avoid religation, dephosphorylation of vector DNA cohesive ends was carried out using calf intestinal alkaline phosphatase (CIP) following subsequent purification with QIAquick PCR purification and gel extraction kit or Wizard SV Gel and PCR Clean-Up System according to the manufacturer's instructions.

DNA ligations were performed in the presence of T4 DNA ligase. Typically, for a 10 µl reaction containing 1 µg DNA consisting of dephosphorylated vector DNA and an insert fragment in a molar ratio between 1:3 and 1:10, 1 µl T4 DNA ligase (400 U/µl) and 1x ligase buffer was used. The ligation was carried out for 1 h at 25°C or overnight at 16°C for increased efficiency. The complete reaction was used to transform chemically competent *E. coli* DH5α cells.

For processing of constructs resulting from whole plasmid PCR (for point mutations or for inserting FLAG or His6 tags) a subsequent digestion with *Dpn*I for 1 h at 37°C was performed to destroy methylated original template DNA (Weiner et al., 1994). After enzyme inactivation for 20 min at 80°C and DNA-purification, ligation was carried out in 10 µl reactions containing 8 µl DNA, T4 DNA ligase reaction buffer and 400 U of T4 DNA ligase. After incubation for at least 2 h at RT, reactions were transformed into chemically competent *E. coli* DH5α cells.

4.3.5 Preparation and transformation of competent *E. coli* cells

Chemically competent *E. coli* BL21(DE3) or DH5α cells were prepared by the Rubidium Chloride method (Hanahan, 1983). 50 ml LB medium was inoculated (1:100) with an overnight culture of the respective strain and grown at 37°C to an

OD_{600} of 0.4-0.6. The culture was chilled on ice for 15 min before it was centrifuged for 15 min at 2500 rpm and 4^oC. Cells were resuspended in 16 ml buffer I (100 mM RbCl, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$, 15% (v/v) glycerol, pH 5.8 adjusted with acetic acid) and incubated on ice for 15 min. Then, cells were centrifuged again and resuspended in 4 ml buffer II (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15% (v/v) glycerol, pH 6.8 adjusted with NaOH). After incubation on ice for 15 min, aliquots of 50 µl were frozen on dry ice and stored at -80^oC.

For transformations, competent cells were thawed on ice, immediately mixed with 1 µl plasmid DNA or 10 µl of a ligation reaction and incubated on ice for 30 min. Cells were then subjected to a heat shock at 42°C for 45 s and immediately chilled on ice for 2 min. 800 µl of prewarmed LB medium was added and upon phenotypical expression for 60 min at 37°C, the transformation reaction was plated on selective LB-agar plates and incubated at 37°C overnight.

4.3.6 Cloning strategies

A cDNA-library of *Arabidopsis thaliana* was used to amplify the genes of At-*rbcL* (aa 1-479, AtCg00490), At-*rbcS1A* (aa 56-180, AT1G67090), At-*rbcS3B* (aa 56-181, AT5G38410), and At-*rbcX* (aa 83-203, AT5G19855). The transit peptide for At-RbcS1A, At-RbcS3B and At-RbcX was omitted based on ChloroP 1.1 Server (expasy.org) prediction and by multialigning (MultAlin) (Corpet, 1988) with cyanobacterial RbcS and RbcX respectively. Since four genes encode four types of RbcS in *A. thaliana*, those genes with highest expression levels in planta were chosen for this study (Yoon et al., 2001). The genes were cloned between the NdeI/BamHI sites of plasmid pET11a (Amp^R) resulting in At-*rbcL*-pET11a, At-*rbcS*-pET11a and At-*rbcX*-pET11a.

The reverse primer for At-*rbcL* contained an additional NheI site between the stop codon and the BamHI site, the At-*rbcL* expression cassette was extracted from At-*rbcL*-pET11a with SphI/NheI and introduced between the SphI/XbaI sites of At-

rbcX-pET11a, in front of the ribosomal binding site of At-*rbcX*. In the resulting co-expression plasmid At-rbcLX-CoEx-pET11a, both At-*rbcL* and At-*rbcX* share the same T7-promotor for co-transcription, but have their own ribosomal binding sites for translation.

pHUE vector has His6+Ubiquitin and a MCS site (Catanzariti et al., 2004). At-*rbcX* was subcloned from At-*rbcX*-pET11a into pHUE between the Sac1/BamH1 sites. The FLAG-tagged construct At-*rbcX*$_\text{N-FLAG}$-pHUE was obtained by introducing the sequences for FLAG-tag (MDYKDDDDK) and a connecting spacer (AG) after the sequence encoding Ubiquitin (before At-*rbcX*) via whole plasmid PCR (Weiner et al., 1994). Point mutations in At-*rbcX* were introduced by the same method using mutation site-specific primers.

The plasmids that were generated in this study were confirmed by sequencing either by Medigenomix GmbH (Martinsried, Germany) or by the MPI sequencing service. Comparative sequence analysis was performed using BLAST (www.ncbi.nlm.nih.gov) and MultAlin analysis program (Corpet, 1988) (http://bioinfo.genotoul.fr/multalin/mult alin.html).

The constructs for cyanobacterial RbcX and RuBisCO components, At-ch-cpn60, At-ch-cpn20 and At-ch-cpn10 were made by Sandra Saschenbrecker.

4.4 Protein biochemical methods

4.4.1 Protein analytical methods

4.4.1.1 Determination of protein concentrations

Concentrations of purified proteins were determined spectrophotometrically on the basis of the Beer-Lambert law and their theoretical extinction coefficients at λ=280 nm (Gill and von Hippel, 1989), as calculated by the ProtParam tool at the ExPASy proteomics server (http://www.expasy.org) unless mentioned otherwise.

Molar concentrations of proteins present as complexes in their native sate (e.g. chaperones or RuBisCO) are expressed for the native state oligomers.

Protein concentrations of complex protein mixtures and cell lysates were determined spectrophotometrically at λ=595 nm using the Bio-Rad protein assay reagent according to the manufacturer`s recommendations (Ausubel, 1997; Bradford, 1976).

4.4.1.2   Sodium-dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using a discontinuous buffer system under denaturing and reducing conditions (Laemmli, 1970). Typically, gels were poured with a 5% polyacrylamide stacking gel on top of a 8-16% polyacrylamide separating gel, depending on the required resolution (Table I). SDS sample buffer was added to the protein samples to a 1x concentration. Prior to loading, samples were boiled at 95°C for 5 min. Electrophoresis was carried out in Mini-Protean electrophoresis chambers in running buffer at 150 V.

Chemicals (For 4 mini-gels)	Separating gel			Stacking gel
	8 %	12.5 %	16 %	5 %
30 % Acryalmide (0.8% bis)	4.5 ml	6.8 ml	9.0 ml	1.66ml
1.875 M Tris-HCl, pH 8.8	3.5 ml	3.5 ml	3.5 ml	-
0.6 M Tris-HCl, pH 6.8	-	-	-	1.00 ml
ddH$_2$O	8.7 ml	6.4 ml	4.2 ml	7.20 ml
10 % (w/v) SDS	167 µl	167 µl	167 µl	100 µl
10 % (w/v) APS	100 µl	100 µl	100 µl	50 µl
TEMED	10 µl	10 µl	10 µl	10 µl

Table I. Composition for SDS-PAGE gels.

2x SDS sample buffer: 14 mM Tris-HCl, pH 6.8, 5% SDS, 20% (v/v) glycerol, 0.1% (w/v) bromphenolblue, 2% (v/v) β-mercaptoethanol

Running buffer: 25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS

4.4.1.3 Native-PAGE

In Native-PAGE the separation of proteins is based on the charge and the hydrodynamic size. Gels were cast as mentioned in Table II and samples were taken up in 2x native loading dye (50% (v/v) glycerol, 0.25% (w/v) bromphenolblue, in native electrophoresis buffer). Electrophoresis was performed in Mini-Protean electrophoresis chambers in native electrophoresis buffer (50 mM Tris, 38 mM glycin, pH not adjusted) at 4^0C, employing a constant voltage of 100 V for the first 30 min and 200 V till the end of the run.

Chemicals	Separating gel		Stacking gel
(For 2 mini-gels)	6 %	10 %	5 %
Acrylamide/Bis (37.5:1, 30 %)	1.67 ml	2.83 ml	0.83 ml
1.875 M Tris-HCl, pH 8.8	1.75 ml	1.75 ml	-
0.6 M Tris-HCl, pH 6.8	-	-	0.50 ml
ddH$_2$O	5.00 ml	3.84 ml	3.65 ml
10 % (w/v) APS	50 µl	50 µl	25 µl
TEMED	5 µl	5 µl	5 µl

Table II. Composition for Native-PAGE gels.

4.4.1.4 Tricine-PAGE

Tricine gels were used for analyzing small proteins (5-20 kDa). Protein samples were mixed with 3x sample buffer (200 mM Tris-HCl, pH 6.8, 2% SDS, 40% (v/v) glycerol, 0.04% (w/v) Coomassie blue G-250, 2 % (v/v) β-mercaptoethanol) and boiled for 5 min at 95^oC, before they were applied to the gels prepared according to Table III. Gels were run in Mini-Protean electrophoresis chambers with separate Cathode-buffer (100 mM Tris, 100 mM Tricine, 0.1% (w/v) SDS) and Anode-buffer (200 mM Tris-HCl, pH 8.9) at a constant voltage of ≤ 130 V.

Chemicals (For 2 mini-gels)	Separating gel	Stacking gel
	16 %	4 %
30 % Acryalmide (0.8% bis)	5.35 ml	0.65 ml
3 M Tris-HCl, pH 8.45	3.35 ml	1.24 ml
ddH$_2$O	200 µl	3.10 ml
10 % (w/v) SDS	100 µl	50 µl
Glycerol	1.00 ml	-
10 % (w/v) APS	100 µl	50 µl
TEMED	10µl	5 µl

Table III. Composition for Tricine-PAGE gels.

4.4.1.5 Bis-Tris Native PAGE

In the Bis-Tris Native PAGE system, protein migration correlates with molecular masses of proteins and was performed according to described procedures (Hansen et al., 1999c; Schagger and von Jagow, 1991). Gels were prepared as mentioned in Table IV. Samples were mixed with loading dye to final concentrations of 50 mM Bis-Tris, pH 7.0, 5% glycerol and 0.1% bromphenolblue. Electrophoresis was performed in cold room or on ice in Mini-Protean electrophoresis chambers. The cathode buffer consisted of 50 mM Tricine and 15 mM Bis-Tris, pH 7.0, whereas the anode buffer contained 50 mM Bis-Tris, pH 7.0. Initially the gels were run at 100 V for 40 minutes and further at 240 V.

Chemicals (For 4 mini-gels)	Separating gel	Stacking gel
	12.5 %	5 %
30 % Acryalmide (0.8% bis)	4.16 ml	2.20 ml
3x gel buffer (150 mM Bis-Tris, pH 7.0, 1.5 M ε-amino n-caproic acid)	3.34 ml	4.00 ml
Glycerol	1.58 ml	-
ddH$_2$O	0.88 ml	5.68 ml
10 % (w/v) APS	33.4 µl	96 µl
TEMED	3.4 µl	9.6 µl

Table IV. Composition for Bis-Tris Native PAGE gels.

4.4.1.6 Coomassie blue staining of polyacrylamide gels

Coomassie blue staining was carried out to detect protein amounts of ≥ 0.5 µg on SDS-PAGEs,.To fix and to stain the proteins, gels were incubated in staining solution (0.16% (w/v) Coomassie brilliant blue R-250, 40% (v/v) ethanol, 8% (v/v) acetic acid) followed by multiple washes in destaining solution (20% (v/v) ethanol, 7% (v/v) acetic acid) to remove background staining.

4.4.1.7 Silver staining of polyacrylamide gels

After electrophoresis, gels were incubated for 30 min in fixing solution (40% (v/v) ethanol, 10% (v/v) acetic acid), resulting in precipitation of the proteins and diffusion of SDS. Subsequently, gels were placed into incubation solution (300 ml ethanol, 68 g/l sodium acetate x 3 H_2O, 2 g/l sodium thiosulphate x 5 H_2O) for 30 min to oxidize the proteins. Gels were then washed in water three times for 5 min and transferred into silver solution (1 g/l silver nitrate, 0.025% (v/v) formaldehyde added before use) for 40 min. Thereafter, proteins were visualized by replacement of the silver solution with developing solution (26 g/l sodium carbonate, 0.0125% (v/v) formaldehyde added before use). The sodium carbonate of the latter solution reduces the silver nitrate attached to the proteins and thus the proteins adopt a brown color. As soon as the desired staining intensity was achieved, the reaction was stopped by addition of stop solution (20% (v/v) ethanol, 7% (v/v) acetic acid), which was replaced by water after 20 min.

4.4.1.8 Autoradiography

Autoradiography was used to detect radioactively (^{35}S) labeled proteins. Polyacrylamide gels were Coomassie stained, destained, rinsed in water and vacuum dried (Stacked Gel Dryer SGD300) on whatman paper. Migration distances of standard proteins were marked on the whatman paper with a

radioactive dye. The radioactive products from the gel was transferred to a phosphoimaging plate using the phosphoimaging cassette (exposure time 3-8 hours) which was read by phosphoimager (Image Reader Fuji-FLA2000) and processed using AIDA software.

4.4.1.9   Western blotting and immunodetection

Western blotting was carried out in a semi dry blotting unit. Upon separation by SDS-PAGE, proteins were transferred to nitrocellulose membranes by applying a constant current of ~1 mA/cm^2 gel size in 25 mM Tris, 192 mM glycine, 20% methanol for 1.5 h. Prior to blocking the membranes with 5% skim milk powder in TBS (2.1.3.2) for 1 h, transfer efficiency was verified by PonceauS (2.1.3.2) staining. The membranes were then incubated with primary antibodies (diluted to a suitable concentration in 5% milk TBS) for 1 h at RT or overnight at 4 °C, followed by the incubation with HRP conjugated secondary antibodies (diluted 1:1000 in milk TBS) for 1 h at RT. Extensive washing between the incubation steps was performed with TBST. For immunodetection, ECL chemiluminescence solution was freshly prepared by mixing equal amounts of ECL solution I (100 mM Tris-HCl, pH 8.5, 2.5 mM luminol (3-aminophtalhydrazid), 400 uM p-coumaric acid) and ECL solution II (100 mM Tris-HCl, pH 8.5, 5.4 mM H$_2$O$_2$). Membranes were incubated in the resulting solution and protein bands were detected and documented with the Fuji-LAS3000 luminescence and densitometry system.

4.4.1.10   TCA precipitation

TCA was added to the protein samples at a final concentration of 20% (v/v). After incubation on ice for 15 min and centrifugation (15 min, 20800 x g, 4°C), 200 µl of chilled acetone were added. After another centrifugation, the supernatant was removed and pelleted samples incubated at RT until all residual acetone was evaporated. Pellets were dissolved in 20 µl 1 x SDS-loading dye (7 mM Tris-HCl, pH 6.8, 2.5% SDS, 10% (v/v) glycerol, 0.05% (w/v) bromphenolblue, 1% (v/v) β-

mercaptoethanol), mixed with 1 µl 2 M Tris-base, boiled for 5 min at 95°C and subjected to SDS-PAGE analysis.

4.4.1.11 FFF-MALS (Field Flow Fractionation - Multiangle Light Scattering)

Mass Determination of proteins by FFF-MALS was performed by Dr. Manajit Hayer-Hartl. Protein complexes (80 µg) were analyzed by field flow fractionation (FFF) using a 490 nm spacer and 30 kDa MWCO membrane (Wyatt Technology) with elution and crossflow of 1 ml/min (Roessner, 1994). The FFF was online with DAWN EOS multi-angle light scattering (Wyatt Technology, 690 nm laser), variable wavelength UV absorbance set at 280 nm (Agilent 1100 series) and Optilab DSP refractive index (Wyatt Technology, 690 nm) detectors (Wyatt, 1993). Masses were calculated using the ASTRA software (Wyatt Technology) with a value set to for dn/dc for protein of 0.185 ml/g. Alternatively, gel filtration (TSK Super 3000 SW, 4.5 mm x 30 cm) could precede MALS-analysis instead of FFF.

4.4.1.12 N-terminal sequencing of proteins

For N-terminal sequencing, proteins were separated by SDS-PAGE and blotted onto a PVDF membrane, which was soaked in methanol and equilibrated with Western blot buffer prior to protein transfer. The membrane was stained (0.1% (w/v) Coomassie brilliant blue R-250, 10% (v/v) acetic acid, 40 % (v/v) methanol) for 2-5 min and subsequently destained (10% (v/v) acetic acid, 30% (v/v) methanol), before it was rinsed in H$_2$O and air-dried. Protein bands of interests were cut and analyzed *via* Edman degradation by the *MPI* protein sequencing service.

4.4.1.13 Mass spectrometry LC/MSMS

This work was performed in collaboration with Dr. Frank Siedler, MPI, Martinsried.

Sample preparation

For LC-MS/MS analysis, slices were cut from the gel as annotated on the corresponding gel images. Slices were cut into pieces of circa 1 mm^3 and transferred to 0.5 ml reaction tubes. Samples from silver-stained gels were destained by oxidation of the silver before digestion (Gharahdaghi et al., 1999). The gel pieces were shaken in 30 µl destaining solution (15mM potassium hexacyanoferrate III, 50mM sodium thiosulfate) until they were colorless. Subsequently, they were washed three times for five minutes with 100 µl H_2O. Coomassie stained gels were used directly for digestion. Tryptic in-gel digestion was performed by a protocol modified from (Shevchenko et al., 1996). Gel pieces were washed alternately in 50 µl 50% acetonitrile and 50 µl 50mM ammonium bicarbonate, each three times for 10 min. After this, they were incubated in 50 µl 10mM dithiothreitol (DTT), 50mM ammonium bicarbonate for 45 min at 56 °C and in 50 µl 55mM iodacetamide, 50mM ammonium bicarbonate for 30 min at room temperature in the dark. The gel pieces were washed again alternately in 50 µl 50% acetonitrile and 50 µl 50mM ammonium bicarbonate, each three times for 10 min. To digest proteins, the gel pieces were incubated overnight in 25 µl trypsin solution (20 µg trypsin solved in 20 µl storage buffer and diluted with 5.2 ml 50mM ammonium bicarbonate) at 37°C on a shaker. The supernatant was transferred to a 0.5 ml reaction tube and remaining peptides were eluted in three steps by incubation in 50 µl H_2O, 50 µl 50% acetonitrile, and 50 µl 50% acetonitrile, 0.1% TFA on a shaker for 20 min each time. The supernatant of the digestion and all elution steps was pooled, frozen in liquid nitrogen, and dried down in a vacuum concentrator.

Desalting

Salts remaining from the tryptic digestion (mainly ammonium bicarbonate) need to be removed from the sample prior to mass spectrometric measurements. Samples for Nano-LC MS/MS were desalted by reversed-phase (RP) chromatography using self-packed Stage tips (STop And Go Extraction) (Rappsilber et al., 2003). A small disk (app. 0.5 mm diameter) was punched out of Teflon embedded C18 material (C18 Empore™ Extraction Disk, 3M) and placed in a GELoader® pipette tip. Solutions were pressed through this column by applying pressure with a 1 ml syringe or by centrifugation (1500 x g for binding peptides, 3000 x g for all other steps). The C18 material was washed with 10 µl isopropanol and equilibrated with 10 µl 10% formic acid. The peptides were dissolved in 1 µl formic acid, diluted with 9 µl H_2O, and passed through the column. The tips were washed twice with 10 µl 10% formic acid and peptides eluted with 5 µl 80% methanol, 10% TFA into a 0.5 ml reaction tube. Eluted peptides were dried down in a vacuum concentrator.

Nano-LC MS/MS (Q-TOF)

Peptides were chromatographically separated on a CapLC system (Waters) and the eluate directly injected into a Q-TOF ultimate mass spectrometer (Waters). The dried peptides were dissolved in 20 µl 5% formic acid, and 1-6 µl (depending on the amount of protein estimated by the intensity of the Coomassie stained gel) were loaded into the CapLC using the systems auto sampler unit. They were bound to the precolumn (self-packed, 100 µmx 25 mm ReproSil-Pur 200 C18-AQ, 5 µm, Dr. Maisch GmbH, Ammerbuch- Entringen, Germany) with a flow of 2 µl/min and analysed on the main column (selfpacked, 75 µmx 150 mm ReproSil-Pur 200 C18-AQ, 3 µm) with a flow of 200 nl/min. Bound peptides were eluted by a stepwise 100 min gradient of 0-100% between buffer A (0.5% formic acid in water) and buffer B (0.5% formic acid in acetonitrile). Mass spectrometric analysis was performed in the *positive ion mode* with a capillary voltage of 2.3 kV. The mass window was set to 300-2000Da in MS mode and 50-2000Da in MS/MS mode. Survey scans were acquired for 1.5 s. From each survey scan up to two peptides were chosen for fragmentation by CID; selection criteria were the signal intensity

and the charge state (at least two-fold). CID was performed with a collision voltage between 16 and 40 kV (depending on peptide mass and charge) and helium as collision gas.

Processing of Nano-LC MS/MS data

Peak lists were extracted from the raw data with Mascot Distiller (vers. 2.1.1, default parameters, MatrixScience, London, UK) and submitted to an InHouse Mascot server (vers. 2.1, MatrixScience, London, UK) for search against the SwissProt, or non-redundant NCBI protein sequence database. Carbamidomethylation of cysteine was set as a required modification and oxidation of methionine as variable modification. Up to one missed cleavage site was allowed. Mass tolerance was set to 1.5 Da for MS and 0.6 Da for MS/MS. If several samples had to be analysed, the process of peak list generation and search submission was automated by use of the Mascot Daemon.

4.4.1.14   Sequence alignments

Multiple sequence alignment with hierarchical clustering was performed using MultAlin (Corpet, 1988) (http://bioinfo.genotoul.fr/multalin/multalin.html).

4.4.2   Protein expression and purification

All protein purification steps were performed at $4^{\circ}C$ unless stated otherwise. Purification of At-ch-cpn60αβ, Syn7002-RbcX, Syn7002-RbcS, Syn6301-RbcL$_8$S$_8$ and Syn6301-RbcS were standardized by Sandra Saschenbrecker and the protocol is as described below.

4.4.2.1   At-ch-cpn60αβ (*Arabidopsis thaliana* ch-cpn60α$_7$β$_7$)

E. coli BL21(DE3), transformed with at-ch-cpn$\alpha\beta$-CoEx-pET11a, were grown at $37^{\circ}C$ in LB medium. After reaching mid-log phase, expression of At-ch-cpn60αβ

was induced with 1 mM IPTG for 3.5 h. Cells were harvested by centrifugation (25 min at 4200 rpm), resuspended and incubated for 1 h in buffer A (20 mM Tris-HCl, pH 7.5, 20 mM NaCl), supplemented with 1 mM EDTA, 0.5 mg/ml lysozyme, 10 U/ml Benzonase and Complete protease inhibitor cocktail. Cells were disrupted by freeze-thawing as well as ultrasonication and cell debris was removed by ultracentrifugation (35 min at 40000 rpm). The lysate supernatant was applied to a SourceQ column, equilibrated with buffer A/1 mM EDTA and eluted with a linear salt gradient from 0.02 to 1 M NaCl. Throughout the purification, fractions were analyzed both by SDS-PAGE and by Native PAGE to distinguish complexes of At-ch-cpn60$\alpha_7\beta_7$ from GroEL and from At-ch-cpn60-monomers. Fractions containing At-ch-cpn60$\alpha\beta$ were pooled, dialyzed against 20 mM Tris-HCl, pH 7.5 and applied to an equilibrated MonoQ ion exchange column (Amersham Biosciences). Elution was performed with a linear gradient from 0 to 0.95 M NaCl and At-ch-cpn$\alpha\beta$ containing fractions were subsequently applied to a Hi-Trap Heparin Sepharose column, equilibrated with 20 mM Tris-HCl, pH 7.5. At-ch-cpn$\alpha\beta$ did not bind to the latter column and was collected in the flow through, which was concentrated using Amicon Ultra MWCO 100 kDa and applied to Superdex 200 gel filtration chromatography in buffer B (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 5% (v/v) glycerol). Fractions containing At-ch-cpn60$\alpha\beta$ were pooled, concentrated (MWCO 100 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Complex concentration was determined spectrophotometrically at 280 nm (192150 M^{-1} complex At-ch-cpn60$\alpha_7\beta_7$). Native PAGE, light scattering and functional assays confirmed purification of active At-ch-cpn60$\alpha_7\beta_7$-complexes. ESI-MS and SDS-PAGE (8% resolution) verified the absence of GroEL/GroES and the presence of equal amounts of α- and β-subunits in the purified complexes.

4.4.2.2 At-ch-cpn20, At-ch-cpn10 (*Arabidopsis thaliana* ch-cpn20 and ch-cpn10)

E. coli BL21(DE3) cells, transformed with at-ch-cpn*20*-pET11a, at-ch-cpn*10*-pET11a were grown to mid-log phase at 37^{0}C and induced with 1 mM IPTG for 3 h.

Harvested cells were incubated for 1 h in buffer C (50 mM Tris-HCl, pH 7.5, 20 mM NaCl), containing 1 mM EDTA, 0.5 mg/ml lysozyme, 10 U/ml Benzonase and Complete protease inhibitor cocktail. After ultrasonication, cell debris was removed by ultracentrifugation (35 min, 40,000 rpm). The supernatant was applied to a DE52 ion exchange column, equilibrated with buffer C, and eluted with a linear salt gradient from 0.02 to 1 M NaCl. Fractions containing Cpn20 or Cpn10 were pooled, dialyzed against 50 mM Tris-HCl, pH 8, applied to an equilibrated SourceQ column and eluted with a linear gradient from 0 to 1 M NaCl. The resulting protein pool was concentrated (MWCO 30 kDa) and subjected to Sephacryl S-200 gel filtration chromatography in buffer C. Fractions containing Cpn20 or Cpn10 were applied to a Hi-Trap Heparin Sepharose column, equilibrated in the same buffer. Cpn20 or Cpn10, eluting in the flow through, were concentrated (MWCO 30 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined spectrophotometrically at 280 nm (30720 M^{-1} tetramer Cpn20, 57750 M^{-1} heptamer Cpn10). N-terminal sequencing confirmed that the proteolysis product of At-ch-cpn20, which appeared throughout the purification, was due to C-terminal proteolysis. Light scattering confirmed the tetrameric nature of Cpn20 and the heptameric complex of Cpn10.

4.4.2.3   At-ch-cpn20$_{N-His6}$

For *E. coli* BL21(DE3) cells harboring plasmid At-ch-cpn*20$_{N-His6}$*-pProEx, protein expression and cell lysis was performed as described for At-ch-cpn60 with buffer E (20 mM Tris-HCl, pH 7.5, 500 mM NaCl), supplemented with 0.5 mg/ml lysozyme, 10 U/ml Benzonase and Complete protease inhibitor cocktail. Soluble cell lysate was applied to a Ni-NTA-agarose column, equilibrated with buffer E. After stepwise washing with buffer E containing 10 mM, 50 mM and 100 mM imidazole, the majority of At-ch-cpn20$_{N-His6}$ eluted with 250 mM imidazole. This protein pool was dialyzed against buffer A and applied to a MonoQ column, which was equilibrated with buffer A and developed with a linear salt gradient from 0 to 1 M NaCl. Fractions containing At-ch-cpn20$_{N-His6}$ were concentrated (MWCO 10 kDa) and

applied to Superdex 200 gel filtration chromatography in buffer A. Fractions containing the desired protein were concentrated, flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined spectrophotometrically at 280 nm.

If the N-terminal His-tag had to be removed, gel filtration was preceded by overnight digestion with TEV-protease (1 mg TEV per 100 mg tagged protein) at 4°C, followed once more by Ni-NTA affinity chromatography with Cpn20 eluting in the flow through.

4.4.2.4 Syn6301-RbcL$_8$S$_8$ (*Synechococcus* sp. PCC6301 RbcL$_8$S$_8$)

To increase the amount of soluble Syn6301-RbcL$_8$S$_8$ in the *E. coli* lysate, overexpression of GroEL/ES preceded the expression of RbcLS. For this purpose, *E. coli* BL21 (DE3) cells were transformed with Syn6301-*rbcLS*-pET11a and pBAD33*ES/EL* and grown to OD$_{600}$ ~0.6 at 30°C. Then expression of GroEL/ES was induced with 0.4% (w/v) arabinose for 1.5 h, before cells were shifted to fresh LB medium (w/o arabinose) containing 1 mM IPTG for expression of RbcLS for 3 h at 30°C. Cells were harvested by centrifugation, incubated for 1 h in buffer C (supplemented with 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 0.5 mg/ml lysozyme, 10 U/ml Benzonase, Complete protease inhibitor cocktail), freeze-thawed and passed through a high pressure cell disruptor. Cell debris was removed by ultracentrifugation and the lysate supernatant was fractionated on a DE52 column with a linear salt gradient from 0.02 to 1 M NaCl in buffer C / 1 mM DTT. Fractions were analyzed for Syn6301-RbcL$_8$S$_8$ by SDS-PAGE and immunoblotting against RbcL. The Syn6301-RbcL$_8$S$_8$ pool was supplemented with (NH$_4$)$_2$SO$_4$ (20 % saturation) and applied to a Phenyl-Sepharose column, equilibrated with 50 mM Tris-HCl, pH 7.5, 20% saturation (NH$_4$)$_2$SO$_4$, 0.5 mM DTT. Elution was performed with a linear gradient from 20 to 0% saturation (NH$_4$)$_2$SO$_4$. Syn6301-RbcL$_8$S$_8$ fractions were dialyzed against 20 mM Imidazol, pH 6.2. After ultracentrifugation, the supernatant was loaded onto an equilibrated MonoQ column and the protein

was eluted with a linear gradient from 0 to 0.7 M NaCl. Fractions containing Syn6301-RbcL$_8$S$_8$ were concentrated (MWCO 100 kDa) and passed over a Superose 6 gel filtration column in buffer B. Eluted Syn6301-RbcL$_8$S$_8$ was concentrated (MWCO 100 kDa), snap-frozen in liquid nitrogen and stored at -80°C. Complex concentration was determined spectrophotometrically at 280 nM (705520 M^{-1} complex RbcL$_8$S$_8$). Light scattering, Native PAGE and carboxylation acitvity confirmed the complex nature and activity of the purified Syn6301-RbcL$_8$S$_8$.

4.4.2.5 Syn6301-RbcL$_8$ (*Synechococcus* sp. PCC6301 RbcL$_8$)

E. coli BL21(DE3) cells, harboring plasmids Syn6301-*rbcL*-pET11a and pG-KJE8, were grown to mid-log phase at 30°C. Expression of DnaK/DnaJ/GrpE was induced with 0.4% (w/v) arabinose for 2 h, before cells were shifted for ~3 h to fresh LB medium (w/o arabinose) supplemented with 1 mM IPTG as inducer for Syn6301-RbcL expression. Note that since purification of RbcL$_8$ was facilitated in the absence of elevated GroEL levels, coexpression of DnaK/DnaJ/GrpE (instead of GroEL/ES) was carried out to increase the amount of soluble RbcL in *E. coli*. As described for At-ch-cpn60, cells were lysed and fractionated in 25 mM Tris-HCl, pH 8, 1 mM EDTA, 0.5 mg/ml lysozyme, 10 U/ml Benzonase, Complete protease inhibitor cocktail. The lysate supernatant was applied to a SourceQ ion exchange column, eqilibrated with buffer F (50 mM Tris-HCl, pH 8, 50 mM NaHCO$_3$, 10 mM MgCl$_2$) / 1 mM EDTA, 0.5 mM DTT. Proteins were eluted with a linear salt gradient from 0 to 1 M NaCl. Fractions were analyzed for the presence of RbcL$_8$ by SDS-PAGE and Native PAGE, followed by immunoblotting, as well as by measurement of carboxylation activity upon addition of RbcS. Fractions with highest activity and most enriched RbcL$_8$ were pooled and dialyzed against 20 mM Imidazol, pH 6.5, 50 mM NaHCO$_3$, 10 mM MgCl$_2$, resulting in a white precipitate, which was pelleted and dissolved in 50 mM Tris-HCl, pH 8. After filtration (0.22 µM) and dialysis against buffer F, the protein solution was applied to a MonoQ column and eluted with a linear salt gradient from 0 to 0.7 M NaCl in buffer F. The Syn6301-RbcL$_8$ containing fractions were concentrated (MWCO 100 kDa) and subjected to

Superdex 200 gel filtration chromatography in buffer F. Fractions containing Syn6301-RbcL$_8$ were pooled, complemented with 10% (v/v) glycerol, concentrated (MWCO 100 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Complex concentration was determined spectrophotometrically at 280 nM (553040 M^{-1} complex RbcL$_8$). Light scattering and Native PAGE confirmed that the purified protein was assembled Syn6301-RbcL$_8$, which showed carboxylation acitvity upon addition of RbcS.

4.4.2.6 Syn6301-RbcS and Syn7002-RbcS$_{FLAG}$ from inclusion bodies

RbcS was purified from inclusion bodies by modification of previously described methods (Coligan, 2000; Somerville, 1986). *E. coli* BL21(DE3) cells, transformed with Syn6301-*RbcS*-pET11a or Syn7002-*RbcS$_{FLAG}$*-pET11a, were grown to mid-log phase at 37 ^{o}C, before induction of RbcS expression with 1 mM IPTG for 3.5 h. The majority of RbcS was found in inclusion bodies. Harvested cells were incubated for 1 h in lysis buffer (40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA, 5% (v/v) Triton X-100, 0.5 mg/ml lysozyme, 10 U/ml Benzonase, Complete protease inhibitor cocktail). After freeze-thawing, ultrasonication and centrifugation, the pellet was resuspended and washed in 40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA, 5% (v/v) Triton X-100, 2 M urea. Subsequent to centrifugation, washing of pellets was repeated in 40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA. Pellets were finally dissolved in denaturation buffer (50 mM Tris-HCl, pH 7.5, 6 M GdnHCl, 1 mM EDTA, 5 mM DTT). The denatured RbcS was refolded by dialysis against 50 mM Tris-HCl, pH 8, 1 mM EDTA, 0.1 mM GSH, 0.01 mM GSSG at a concentration of ca. 0.5 mg/ml and finally frozen in liquid Nitrogen for storage at -80°C. The recovery of refolded protein was ca. 65 to 85%. Protein concentration was determined spectrophotometrically at 280 nM (19060 M^{-1} monomer Syn6301-RbcS, 24410 M^{-1} monomer Syn7002-RbcS$_{FLAG}$).

4.4.2.7 Syn7002-RbcX, AnaCA-RbcX (WT/mutant/FLAG-tagged RbcX)

E. coli BL21(DE3) cells, transformed with the respective RbcX plasmids, were used for expression of RbcX upon induction with 1 mM IPTG for 3.5 h at 37°C. Harvested cells were incubated for 1 h in lysis buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM DTT, 0.5 mg/ml lysozyme, 10 U/ml Benzonase, Complete protease inhibitor cocktail) and disrupted by ultrasonication. After removal of cell debris by centrifugation, the supernatant was applied to a SourceQ column, equilibrated with 50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, and eluted with a linear gradient from 0 to 1 M NaCl. Fractions containing RbcX were dialyzed against 20 mM imidazole, pH 6.4, 20 mM NaCl, 0.5 mM DTT, applied to an equilibrated MonoQ column, and eluted with a linear salt gradient up to 0.7 M NaCl. Fractions containing RbcX were dialyzed against buffer G (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM MgCl$_2$) and applied to a HiTrap Heparin Sepharose column. RbcX eluted in the flow through, which was then concentrated (MWCO 10 kDa) and subjected to Superdex 200 gel filtration chromatography in buffer G. Fractions containing RbcX were supplemented with 10% glycerol, concentrated (MWCO 10 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined spectrophotometrically at 280 nM (19060 M^{-1} dimer Syn7002-RbcX, 30440 M^{-1} dimer AnaCA-RbcX).

Syn7002-RbcX mutant proteins were purified in close collaboration with Bharathi Vasudeva Rao.

4.4.2.8   Syn7002-RbcLX$_{N-His6}$ (*Synechococcus* sp. PCC7002 RbcLX$_{N-His6}$ complex)

E. coli BL21(DE3) cells were transformed with plasmid Syn7002-*rbcLX$_{N-His6}$*CoEx-pET28b. Cells were grown at 30°C in to mid-log phase and LX-expression was induced for 3 h with 1 mM IPTG. Cells were lysed and subjected to Ni-NTA-agarose affinity chromatography as described for At-ch-cpn20$_{N-His6}$. The complex of Syn7002-RbcLX$_{N-His6}$ as well as free dimers of Syn7002-RbcX$_{N-His6}$ eluted from the

Ni-NTA-agarose at 250 mM imidazole. This protein pool was dialyzed against 20 mM Tris-HCl, pH 9.2 and its volume reduced (MWCO 10 kDa). Note that the basic pH condition supported complex stability and solubility. Subsequently, size-exclusion chromatography was performed in the same buffer on a Superdex 200 gel filtration column. The initially eluting peak contained Syn7002-LX$_{N-His6}$-complexes, whereas a second peak contained only free Syn7002-X$_{N-His6}$ dimers. Fractions containing the complexes were concentrated (MWCO 10 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined by Bradford assay.

4.4.2.9 Syn6301-RbcL/AnaCA-RbcX$_{N-His6}$ (complex of Syn6301 RbcL and AnaCA-RbcX$_{N-His6}$)

E. coli BL21(DE3) cells, transformed with pG-*KJE8*, Syn6301-*rbcL*-pET11a and AnaCA-*rbcX*$_{N-His6}$-pET28b, were grown to mid-log phase at 30°C. Expression of GroEL/ES and DnaK, DnaJ, GrpE was induced with 20 ng/ml tetracycline and 0.4% (w/v) arabinose, respectively. After 2 h, cells were shifted to fresh medium (w/o tetracycline and arabinose) and expression of Syn6301-RbcL as well as AnaCA-RbcX$_{N-His6}$ was induced with 1 mM IPTG for ~3 h. Soluble cell lysate was produced and applied to Ni-NTA-Agarose as described above for Syn7002-RbcLX$_{N-His6}$. The resulting pool of Syn6301-RbcL/AnaCA-RbcX$_{N-His6}$-complexes and free dimers of AnaCA-RbcX$_{N-His6}$ (eluted at 250 mM imidazole), was dialyzed against 20 mM Tris-HCl, pH 9, 50 mM NaHCO$_3$, 10 mM MgCl$_2$ and applied to a MonoQ column, equilibrated with the same buffer and eluted with a linear salt gradient from 0 to 0.7 M NaCl. Fractions containing the complex were concentrated (MWCO 10 kDa) and applied to a Superdex 200 gel filtration column in 20 mM Tris-HCl, pH 9. The initially eluting peak contained Syn6301-RbcL/AnaCA-RbcX$_{N-His6}$-complexes as verified by Native PAGE. These fractions were concentrated (MWCO 10 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined by Bradford assay. This purification was performed by Sandra Saschenbrecker.

4.4.2.10   At-RbcX$_{N-His6}$ (*Arabidopsis thaliana* RbcX$_{N-His6}$)

At-RbcX$_{N-His6}$ was purified by Bharathi Vasudeva Rao.

E. coli BL21(DE3) cells transformed with At-rbcX$_{N-His6}$-pET11a plasmid, were used for expression of At-rbcX upon induction with 1 mM IPTG for 7 h at 18°C. Harvested cells were incubated for 1 h in buffer E (20 mM BisTris-HCl, pH 6, 500 mM NaCl, 10mM Imidazole), supplemented with 0.5 mg/ml lysozyme, 10 U/ml Benzonase and Complete protease inhibitor cocktail and disrupted by ultrasonication. Soluble cell lysate was applied to a Ni-NTA-agarose column, equilibrated with buffer E. After stepwise washing with buffer E containing 10 mM, 50 mM and 100 mM imidazole, the majority of At-RbcX$_{N-His6}$ eluted with 250 mM imidazole. This protein pool was dialyzed against buffer A (20mM BisTris-HCl, pH6, 100mM NaCl, 8% Glycerol). The protein was concentrated in centricons (MWCO 10 kDa). Concentration was determined spectrophotometrically at 280 nm and the protein was flash-frozen in liquid nitrogen and stored at -80°C.

4.4.2.11   At-RbcX$_{N-His6+Ub}$, At-RbcX$_{N-His6+Ub-FLAG}$ and AtRbcX (Q29A) $_{N-His6+Ub-FLAG}$

These proteins were purified in collaboration with Bharathi Vasudeva Rao.

E. coli BL21(DE3) cells, transformed with either At-rbcX-pHUE or At-rbcX$_{N-FLAG}$–pHUE plasmids, were used for expression of At-rbcX upon induction with 1 mM IPTG for 5 h at 30°C. Harvested cells were resuspended and incubated for 1 h in buffer E (20 mM BisTris-HCl, pH 6, 500 mM NaCl, 10mM Imidazole), supplemented with 0.5 mg/ml lysozyme, 10 U/ml Benzonase and Complete protease inhibitor cocktail and disrupted by ultrasonication. Further the purification was performed using the modification of the protocol published in (Catanzariti et al., 2004). Soluble cell lysate was applied to a Ni-NTA-agarose column, equilibrated with buffer E. After stepwise washing with buffer E containing 10 mM, 50 mM and 100 mM imidazole, the majority of His+Ubiquitin tagged At-RbcX eluted with 250 mM imidazole. The fractions containing At-RbcX were pooled and dialyzed against buffer containing 20mM BisTris-HCl, pH6, 100mM NaCl, 8%

Glycerol for 8 hours. After dialysis, the protein was concentrated (MWCO 10 kDa), flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined spectrophotometrically at 280 nm (29910 M^{-1} momoner for His6+Ub+At-RbcX and 31400 M^{-1} monomer for His6+Ub+Flag+At-RbcX).

To remove the N-terminal His+Ub-tag, the samples from the first Nickel NTA containing At-RbcX was incubated with the ubiquitin protease, Usp2-cc at 1:100 protease to protein molar ratio at 16^{0}C, overnight. 3% Glycerol and 1.5M mercaptoethanol were added to the cleavage reaction. After incubation this mixture was dialysed against buffer containing 20mM BisTris-HCl, pH9.2, 100mM NaCl, 10% Glycerol and then subjected to Ni-NTA chromatography. All the cleaved ubiquitin and the protease which are His-tagged bound to the Ni-NTA beads and At-RbcX was in the flow-through. The flow-through was concentrated in a 10kDa cut off centricon and applied to an equilibrated MonoQ ion exchange column (Amersham Biosciences). Elution was performed with a linear gradient from 0 to 1M NaCl and At-RbcX containing fractions were subsequently pooled and concentrated. The concentrated samples were flash-frozen in liquid nitrogen and stored at -80°C. Protein concentration was determined spectrophotometrically at 280 nm (29910 M^{-1} momoner for At-RbcX$_{N-FLAG}$ and 28420 M^{-1} momoner for AtRbcX).

4.4.2.12 At-RbcS1A (*Arabidopsis thaliana* RbcS1A) from inclusion bodies

At-RbcS1A was purified from inclusion bodies. *E. coli* BL21(DE3) cells, transformed with At-rbcS1A-pET11a, were grown to mid-log phase at 37°C, before induction of rbcS expression with 1 mM IPTG for 3.5 h. The majority of RbcS was found in inclusion bodies. Harvested cells were incubated for 1 h in lysis buffer (40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA, 5% (v/v) Triton X-100, 0.5 mg/ml lysozyme, 10 U/ml Benzonase, Complete protease inhibitor cocktail). After freeze-thawing, ultrasonication and centrifugation, the pellet was resuspended and washed in 40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA, 5% (v/v) Triton X-

100, 2 M urea. Subsequent to centrifugation, washing of pellets was repeated in 40 mM Tris-HCl, pH 8, 0.25 M sucrose, 10 mM EDTA. Pellets were finally dissolved in denaturation buffer (50 mM Tris-HCl, pH 7.5, 6 M GdnHCl, 1 mM EDTA, 5 mM DTT). The denatured RbcS was refolded by dialysis against 50 mM Tris-HCl, pH 8, 1 mM EDTA, 0.1 mM GSH, 0.01 mM GSSG at a concentration of ca. 0.5 mg/ml and finally frozen in liquid Nitrogen for storage at -80°C. Protein concentration was determined spectrophotometrically at 280 nM.

4.4.3 Functional analyses

4.4.3.1 ATPase activity assay

The ATPase activity of chaperonins was determined as described (Lanzetta et al., 1979). Chaperonins were diluted at a concentration of 0.5 μM into assay-buffer 1 (20 mM MOPS-KOH, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$). If the influence of co-chaperones was analyzed, the latter were added at a final concentration of 1 μM. The reactions were incubated for 5 min at 25°C, before they were initiated by addition of 2 mM ATP. At indicated time points aliquots of 10 μl were withdrawn and the reaction stopped with CDTA (final concentration 20 mM). The resulting samples were mixed with 300 μl color reagent (filtered 3:1 mixture of 0.045 % (w/v) malachite green hydrochloride in H_2O and 4.2% (w/v) ammonium molybdate in 4 N HCl; supplemented before use with 0.1% (v/v) Triton X-100) and 40 μl 37% citric acid. After incubation for 30 min at 25°C, the absorption at 640 nm was measured.

For quantification of ATP hydrolysis, a standard assay was performed, in which 10 μl of solutions containing 0, 10, 25, 50, 100, 250, 500 and 1000 μM K_2HPO_4 were applied to the colorimetric assay described above. The resulting calibration curve displayed the relation between phosphate concentration and the respective absorbance and could thus be used to determine the ATP hydrolysis rate of the chaperonins**.**

4.4.3.2 *in vivo* co-expression of RbcL or RbcLS with chaperones in *E. coli* and RuBisCO carboxylation activity

E. coli BL21(DE3) cells were transformed with pET-vectors for expression of RbcL, RbcL/X, RbcL/S or RbcL/X/S. If necessary, the resulting strains were additionally co-transformed with pG-*KJE8* for GroEL/GroES or/and DnaK/DnaJ/GrpE expression. Single colonies were grown to mid-log phase at 30°C. Expression of the *rbc*-genes from T7-promoters was induced with 1 mM IPTG for ~3.5 h at 30°C with or without prior transient overexpression of GroEL/GroES (induced with 0.4% (w/v) arabinose) or GroEL/GroES+ DnaK/DnaJ/GrpE (induced with 0.4% (w/v) arabinose and 20 ŋg/ml tetracycline) for 2 h at 30°C, followed by a transfer to fresh medium.

Equivalent amounts of cells were pelleted, incubated in lysis buffer (50 mM Tris-HCl, pH 8, 20 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.1% (v/v) Triton X-100, 0.1 mg/ml lysozyme) on ice for 30 min, ultrasonicated and fractionated into soluble and insoluble fractions by centrifugation (20800 x g, 30 min at 4°C). Comparative analysis of total, soluble and insoluble protein was performed by SDS-PAGE. Soluble lysate fractions were analyzed for assembled RbcL by 6% Native PAGE, followed by immunoblotting against RbcL.

Carboxylation activity was determined by diluting aliquots of lysate supernatants into assay-buffer 3 (100 mM Tris-HCl, pH 7.5, 10 mM KCl, 2 mM $Mg(OAc)_2$) containing 1 mM DTT and 2 µM BSA. Samples containing only RbcL were supplemented with 7 µM Syn6301-RbcS or Syn7002-RbcS$_{FLAG}$ or At-RbcS1A and assembly was allowed to proceed for 5 min at RT. Thereafter, a ^{14}C-Mix (in 100 mM Tris-HCl, pH 7.5, 10 mM KCl) was added to give final concentrations of 60 mM $NaHCO_3$, 0.5 µCi $NaH^{14}CO_3$ and 10 mM $MgCl_2$. After incubation for 5 min, carboxylation was initiated by addition of 2.5 mM RuBP and stopped with acetic acid (3 N) after 30 min. The resulting mixes were heated (96°C) until complete evaporation of liquid, the remaining non-volatile components were dissolved in 100 µl water, taken up in 1 ml scintillation fluid and radioactivity of the fixed carbon was

quantified (LS 6500 multi-purpose scintillation counter) (Dickson et al., 2000; Goloubinoff, 1989; Viitanen et al., 1995).

4.4.3.3   *in vitro* translation of RuBisCO, immunodepletion of GroEL from RTS *E. coli* lysate and Pulse-chase assays

T7 promoter constituted plasmid (200-300 ng/µl) for Rr-*rbcL* or Syn6301-*rbcL* or Syn7002-*rbcL* or AnaCA-*rbcX* or At-*rbcL* was translated *in-vitro* in the reconstituted *E. coli* lysate (coupled RTS100 *E. coli* HY transcription/ translation system from a bacterial S30 lysate) in the presence of 0.5 U/µl RNAguard ribonuclease inhibitor, Complete protease inhibitor cocktail (Stock: 1 Complete EDTA free mini tablet/1ml of RTS reconstitution buffer; Final amount used 2 µl/20 µl reaction volume), 0.37 MBq ^{35}S-methionine per 20µl reaction, 50 µM unlabelled methionine. Wherever indicated necessary chaperones and other proteins were added to the reaction in the concentrations mentioned in the figure legends. Translation was carried out at 30°C for 90 minutes and stopped by addition of Chloramphenicol (CAM, 200 µg/ml) on ice. Post-translational addition of proteins was performed after CAM addition, followed by transfer of reactions back to 30°C.

When indicated, immunodepletion of GroEL from the lysate was performed by incubation with polyclonal GroEL antibody bound to Protein A sepharose beads by gentle shaking for 45 min at 4°C and removal of beads by centrifugation. Successful depletion of GroEL was confirmed by SDS-PAGE followed by immunoblotting.

RTS products were separated into soluble and insoluble fractions by centrifugation (20800 x g for 30 min at 4°C). The former was analyzed by native-PAGE and SDS-PAGE and the latter by SDS-PAGE followed by autoradiography.

In pulse-chase experiments, translation was performed in the presence of 0.5 µM GroEL, 1 µM GroES and 40 µM RbcX proteins (as indicated) for 6 min at 30°C, before ^{35}S-methionine was added. After 6 min, reactions were chased by addition

of unlabelled methionine. Samples were collected at indicated time points and reactions stopped by addition of CAM on ice as described above. Soluble and insoluble fractions were analyzed by discontinuous Bis-Tris Native PAGE (13% resolution gel, 6% stacking gel) or 12.5% SDS-PAGE, respectively, followed by autoradiography.

4.4.3.4   Analytical gel filtration of *E. coli* lysate or protein complexes

Gel filtration chromatography of soluble *E. coli* lysate or protein complexes (resulting from co-immunoprecipitation) performed using a Superdex 200 (10/30) gel filtration column, equilibrated in 50 mM Tris-HCl, pH 8, 50 mM NaCl, 5 mM MgCl$_2$. If necessary, the sample volumes were reduced (MWCO 30 kDa) to 200 µl prior to application. The column was eluted at 250 µl/min and fractions of 250 µl were collected. If necessary, fractions were TCA-precipitated. Samples were analyzed for RbcL, RbcX$_{FLAG}$ or GroEL by SDS-PAGE and Coomassie staining or immunoblotting.

Similarly for the analysis of purified proteins, the protein samples were injected into Superdex 200 (3.2/30) column equilibrated in buffer 20 mM Tris-HCl, pH 9.2, 100 mM NaCl and 10% Glycerol.

4.4.3.5   Peptide binding assay

Overlapping peptides of Syn7002-RbcL, which were 12 amino acids in length and N-terminally acetylated, were synthesized and C-terminally covalently bound to a cellulose-PEG-membrane by the JPT Peptide Technologies GmbH. Each of the 231 peptide spots (0.37 cm x 0.37 cm) carried approximately 5 nmol of peptide. Similarly the other peptide array included RbcL C-terminal peptides from different photosynthetic organisms including bacteria, red algae and higher plants. The peptide membrane was equilibrated in TBS buffer (50 mM Tris-HCl, pH 8, 137 mM NaCl, 2.7 mM KCl) and blocked in TBS blocking solution (1% (w/v) milkpowder

in TBS buffer) for 3 h at room temperature. Subsequently, it was rinsed in TBS for 10 min and incubated with 10 µg/ml Syn7002-RbcX or At-RbcX^{NFlag} in TBS blocking solution overnight at RT. Afterwards, the membrane was washed three times for 5 min in TBST buffer (20 mM Tris-HCl pH 7.5, 137 mM NaCl, 0.1% (v/v) Tween 20). For detection of peptide-bound RbcX, the membrane was incubated consecutively with RbcX-specific antiserum or FLAG-antibody and HRP-conjugated secondary antibodies in TBST blocking solution for 45 min with intermediary and final washing in TBST buffer. Peptide binding was visualized using chemiluminescence detection substrates. In order to exclude direct hybridization of either antibody with the peptide spots, a control experiment was performed, in which only the incubation step with RbcX was omitted. For regeneration, the membrane was stripped after each single use according to the manufacturer`s instructions as following.

The membrane was washed three times for 10 min in water, followed by four incubation steps of 30 min at 50°C in regeneration buffer (62.5 mM Tris-HCl, pH 6.7, 2% (w/v) SDS, 100 mM β-mercaptoethanol). The membrane was then incubated in 10x PBS-buffer (92 mM Na_2HPO_4 x 12 H_2O, 16 mM NaH_2PO_4 x H_2O, 1500 mM NaCl, pH 7.2 adjusted with NaOH) and in TBST-buffer, both three times for 20 min at RT.

4.4.3.6 Tryptophan-fluorescence spectroscopy

Tryptophan- fluorescence experiments were done by Bharathi Vasudeva Rao. Tryptophan-fluorescence was measured using Spex Fluorolog 3 with the following parameters: excitation 295 nm, emission scan 315-450 nm, incr. 2 nm, interval time 1 sec, slits 1/5. In order to monitor tryptophan-fluorescence of RbcL in refolding reactions, Rr-RbcL or Syn6301-RbcL$_8$ was denatured in denaturation-buffer (20mM MOPS-KOH, pH 7.5, 6 M GdnHCl, 10 mM KCl, 1 mM EDTA, 10 mM DTT) 1 for 60 min at 25°C. The denatured protein was diluted 100-fold (0.25 µM RbcL monomer) into ice-cold assay-buffer (20 mM MOPS-KOH, pH 7.5, 100 mM KCl, 5 mM Mg(OAc)$_2$) containing 0.5 µM GroEL. After incubation for 10 min on ice,

a tryptophan fluorescence scan was taken. 1 µM GroES was added to this sample and a tryptophan fluorescence scan was taken again. Refolding was started by supplementing the reaction with 2 mM ATP and the kinetics of refolding was measured by monitoring the change in tryptophan fluorescence over time. At the end of kinetics run (after ~35 min), a tryptophan scan was performed. If only native or denatured substrate or merely substrate-binding to GroEL had to be analyzed, reactions were modified accordingly. Background fluorescence of chemically identical reactions lacking RuBisCO was subtracted in each assay.

4.4.3.7 ANS-fluorescence spectroscopy

Analysis of ANS (1-anilino-8-naphthalene-sulphonate) fluorescence was performed by the same method as described for tryptophan fluorescence with the exception that assay-buffer 1 contained 1 µM ANS. Fluorescence was measured with Spex Fluorolog 3 applying the following parameters: excitation 350 nm, emission scan 410-570 nm, incr. 2 nm, slits 1/5. Single ring version of GroEL was used for the refolding reactions and low salt buffer containing 20 mM MOPs/KOH pH 7.5, 10 mM KCl, 5 mM $Mg(OAc)_2$ was used. Data were corrected for background fluorescence.

4.4.3.8 Circular Dichroism spectroscopy

Far-UV CD spectra and melting curves were measured with a Jasco J-715 spectrometer equipped with a Peltier-thermostat using 0.1 cm cuvette. Wavelength scans of Syn7002-RbcX wild type and mutant proteins (Q29A, Y17AY20L, R70A, P87A, R102A), Syn7002-RbcX$_{N-FLAG}$, At-RbcX$_{N-FLAG}$ proteins were recorded at 30°C at a protein concentration of 0.1mg/ml. Syn7002-RbcX wild type and mutant proteins were analyzed in a buffer containing 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$. For At-RbcX$_{N-FLAG}$, 20mM BisTris-HCl, pH 9.2, 100mM NaCl was used.

5 Results

Earlier attempts of *in vivo* and *in vitro* reconstitution of form I RuBisCO from cyanobacterium, *Synechococcus* sp. PCC6301, have shown the necessity of the Hsp60/GroE class of chaperones for the productive folding of RuBisCO (Gutteridge, 1995). The DnaK (Hsp70) system (DnaK/DnaJ/GrpE) is also implicated in assisting *in vivo* folding of RuBisCO large subunits of *R. rubrum* and *Synechococcus* in *E.coli* (Checa and Viale, 1997). But so far, *in vivo* and *in vitro* reconstitutions of cyanobacterial and higher plant RuBisCO have not been possible.

5.1 Optimization of *in vitro* translation for the RuBisCO expression

5.1.1 *in vitro* translation of *Rhodospirillum rubrum* rbcL (Rr-rbcL)

Since the type II RuBisCO enzyme from *Rhodospirillum rubrum* (Rr-RbcL$_2$) (Figure 15A) can be folded and assembled as active enzyme *in vivo* in *E. coli* and *in vitro* upon interaction with GroEL/GroES (Figure 15A) (Brinker et al., 2001; Goloubinoff, 1989), this protein served as a positive control for the *in vitro* translation experiments. S30 translation lysates from *E. coli* support efficient protein synthesis *in vitro* but represent dilute cytosol preparations with low levels of endogenous chaperones, including GroEL and GroES (Agashe et al., 2004).

To determine the amount of plasmid, carrying *R. rubrum rbcL* gene (*Rr-rbcL*), to be used in Rapid Translation System (RTS), a titration was carried out. Though there was a slight increase in the yield of Rr-RbcL protein observed upon increasing the amount of plasmid used, the effect was relatively minor (Figure 15C).

A

SDS-PAGE

B

RbcL →

Autoradiogram SDS-PAGE

C

SDS-PAGE
Immunoblot
(anti-RbcL)
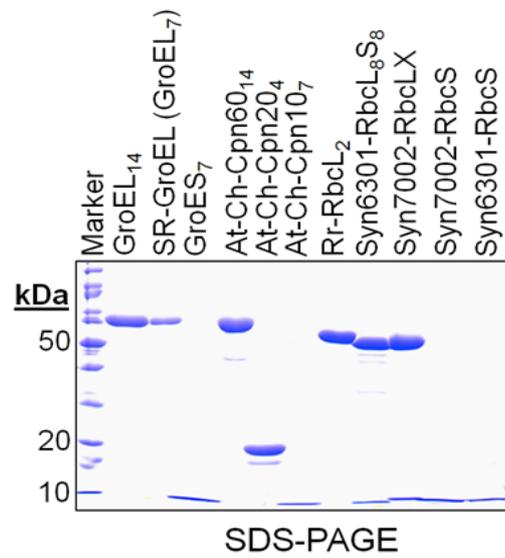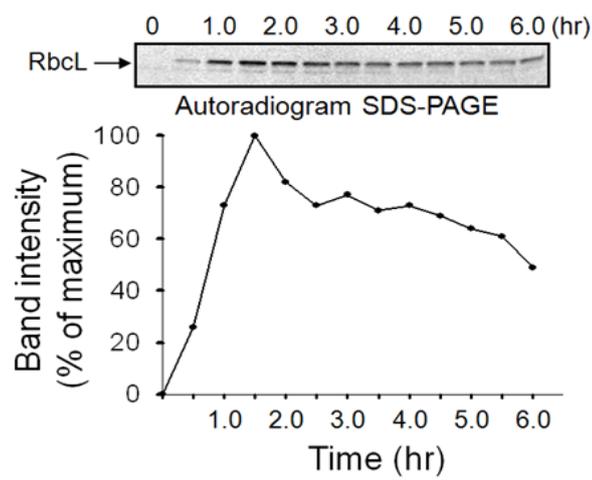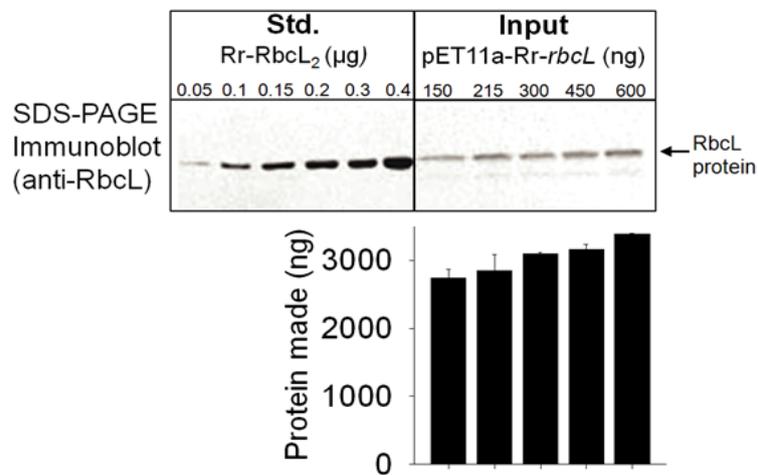
Figure 15. Standardization of *in vitro* translation system for the efficient production of RuBisCO.

(A) Coomassie stained SDS-PAGE of recombinantly purified *E. coli* GroEL, Single-Ring GroEL (SR-GroEL) and GroES, *Arabidopsis thaliana* Ch-pn60, Ch-Cpn20 and Ch-cpn10, *Rhodospirillum rubrum* RbcL$_2$, *Synechococcus* sp. PCC7002-RbcS, *Synechococcus* sp. PCC6301-RbcS.
(B) Autoradiogram of SDS-PAGE showing the total yield of Rr-RbcL protein. T7 driven Rr-*rbcL* was translated in the *E. coli* lysate translation system *in vitro* in 20 µl batch reactions in the presence of ^{35}S-methionine for varying time points at 30^0C and the aliquots were collected and the translation was stopped by the addition of 200 µg/ml chloramphenicol. The products were analyzed by SDS-PAGE followed by autoradiography. Below is the densitometry showing the quantification of bands.
(C) Autoradiogram of SDS-PAGE showing the total yield of Rr-RbcL protein. Different amounts of plasmid carrying Rr-*rbcL* was translated in the translation system *in vitro*, batch reactions were run for 90 min at 30^0C and product was analyzed by SDS-PAGE followed by autoradiography. Below is the quantification of RbcL bands.

To determine the duration needed to obtain optimal yield of RuBisCO protein from *R. rubrum* (Rr-RbcL), the T7 driven Rr-*rbcL* was translated for varying times. Translation products were collected at different time points and translation was inhibited by the addition of chloramphenicol. The highest total protein yield was observed when the translation was run for 90 minutes at 30^0C (Figure 15B). So, for further translation experiments 250-300 ng of plasmid was used and the reactions were incubated for 90 minutes at 30^0C.

5.2   Requirement of GroEL/GroES for efficient folding of RuBisCO

To analyze the requirement of GroEL/GroES system for the folding of RuBisCO large subunits, Rr-*rbcL* encoding RbcL from *R. rubrum* was translated in the *E. coli* lysate in the presence of ^{35}S-Methionine and either in the presence or absence of GroEL/GroES protein. After translation, the products were analysed by SDS-PAGE. A fraction of RbcL protein made in the absence of GroEL/GroES accumulated in the insoluble fraction, whereas RbcL made in the presence of these chaperones was essentially soluble (Figure 16A). As RbcL is the only labeled product, binding of RbcL to GroEL was confirmed by autoradiography.

Further, the products of translation reactions performed in the presence of unlabelled methionine were tested for the enzymatic activity of RuBisCO. Increased specific activity was noticed when RuBisCO was translated in the presence of GroEL/ES (Figure 16B). This suggests a role of GroEL/GroES in folding RbcL. This effect has been previously shown (Goloubinoff, 1989) in *E. coli* and also in reconstitution studies *in vitro* using Form II RuBisCO (Brinker et al., 2001). RuBisCO subunits are thought to fold inside the chaperonin cage, followed by dimerization upon release into solution (Brinker et al., 2001). Note that the minute amounts of endogenous GroEL and GroES in the *E.coli* lysate used for translation has already contributed to the folding of RbcL, which is enhanced further by supplementation with purified GroEL/GroES (Figure 16A and 16B).

5.2.1   Interaction of cyanobacterial RbcL with GroEL

To confirm the interaction of newly synthesized RbcL with GroEL, translation of *rbcL* from a cyanobacterium *Synechococcus* sp. PCC7002 (Syn7002-*rbcL*) was performed. Non-denaturing gel analysis of the translation product in presence of WT GroEL led to the formation of a band comigrating with GroEL, while the translation of RbcL in the presence of SR-GroEL resulted in the band comigrating with SR-GroEL (Figure 16C). These results show that the newly-made RbcL subunit polypeptides were efficiently captured by GroEL, consistent with the findings that GroEL binds to newly-made RbcL protein and thereby prevents the aggregation of RbcL (Checa and Viale, 1997).

A

Rr-*rbcL*

- GroEL/ES   + GroEL/ES

Total | Pellet | Supernatant | Total | Pellet | Supernatant

Autoradiogram of SDS-PAGE

B

Rr-*rbcL*

Specific activity

-GroEL/ES   +GroEL/ES

C

Syn7002-*rbcL*

	1	2
Lysate	+	-
Lysate Δ-GroEL	-	+
SR-GroEL	+	+

RbcL bound to → GroEL

RbcL bound to → SR-GroEL

Autoradiogram Native-PAGE
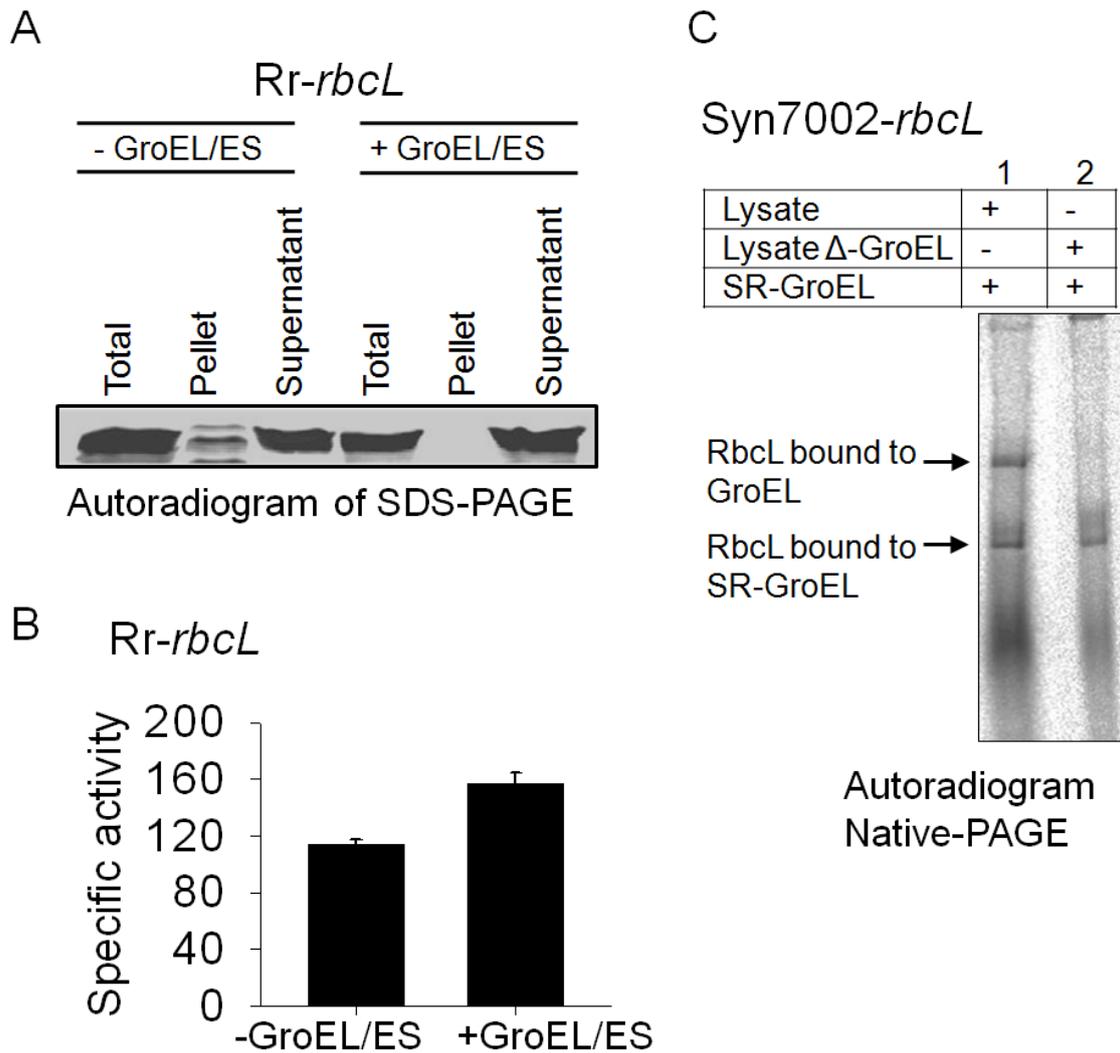
Figure 16. Role of GroEL/ES in folding of newly synthesized RuBisCO large subunit peptides.

(A) Rr-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-Methionine (1.5 hr, 30C). When indicated, GroEL/GroES (1.0 µM/2.0 µM) was added to lysate. The soluble and insoluble products were analyzed by SDS-PAGE, followed by autoradiography. Below is the carboxylation activity of RuBisCO made in the translation. For the carboxylation activity the soluble products of translation reactions carried out in presence of unlabelled methionine were used. Activity was measured; the specific activity of Rr-RuBisCO was measured based on the amount of soluble protein made.

(B) Syn7002-*rbcL* was translated in normal *E.coli* translation lysate supplemented with additional GroEL and single-ring GroEL (SR-GroEL) (Cliff et al., 1999; Weissman et al., 1995) or in the GroEL depleted lysate supplemented with SR-GroEL. The product was analysed by Native-PAGE followed by autoradiography (See Figure 21 for details about the depletion of GroEL in *E. coli* lysate).

5.2.2 *in vitro* synthesis of *Synechococcus* sp. PCC6301 RuBisCO

Expression of cyanobacterial form I RuBisCO from *Synechococcus* sp. PCC6301 (also known as *Anacystis nidulans*) in *E. coli* was reported to result in the chaperonin-dependent production of enzymatically active holoenzyme (Goloubinoff, 1989b). To assess the chaperone requirement for the folding and assembly of RuBisCO, RuBisCO from *Synechococcus* sp. PCC6301 (Syn6301) was translated *in vitro* in an *E. coli* translation lysate and the products were analyzed by Native-PAGE and RuBisCO activity assay. Expression of *rbcL* in the presence of GroEL/GroES enhanced the production of $RbcL_8$ core complex (Figure 17A) and the supplementation of purified RuBisCO small subunits (Syn6301-RbcS) yielded active RuBisCO. Increased activity was observed in the presence of additional GroEL/ES, again demonstrating the role of these chaperones in efficient folding of RbcL (Figure 17A).

Most chaperonin substrates have to pass through several interactive cycles with chaperonin to reach native state. It has been shown that if incompletely folded polypeptide chains are released from GroEL or Cpn60, they are rapidly recaptured by GroEL/Cpn60 for further folding attempts (Hartl and Hayer-Hartl, 2002; Mayhew et al., 1996).

In order to investigate if Syn6301-RbcL recycles to GroEL, Syn6301-*rbcL* was translated in the *E.coli* lysate *in vitro* in presence of varying amounts of GroEL and GroES. The highest yield of $RbcL_8$ and RuBisCO activity upon addition of RbcS was observed when 0.5µM/1.0µM of GroEL/GroES was supplemented (Figure 17B). Increasing the GroEL/GroES concentration further resulted in a reduced amount of $RbcL_8$ and activity, suggesting that released protein is recaptured by GroEL and so prevented from efficient assembly to $RbcL_8$ complex (Figure 17B).
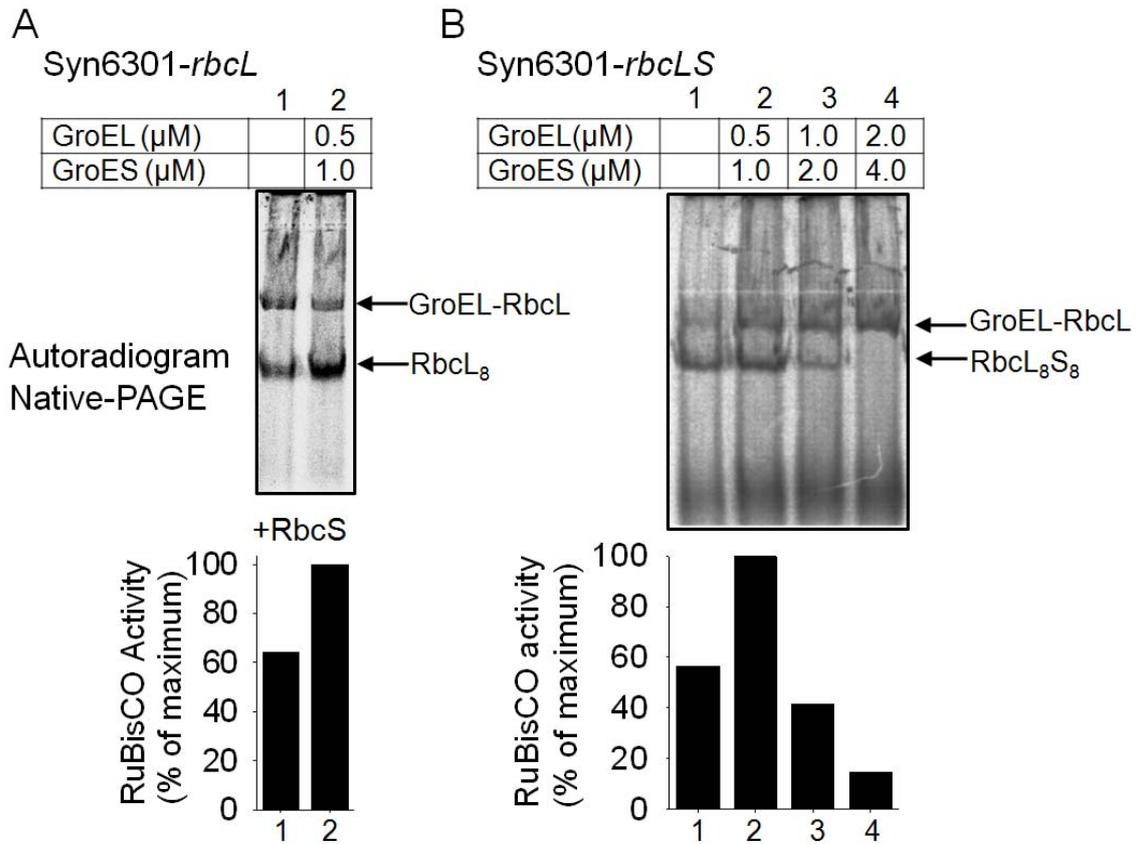
Figure 17. Requirement of the chaperonin system for Syn6301-RuBisCO folding and recycling of Syn6301-RbcL on GroEL.

(A) Syn6301-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine (1.5 hr, 30^{0}C). When indicated, GroEL/GroES (0.5 μM/1.0 μM) was added to lysate. Assembled RbcL$_8$ was analyzed by Native-PAGE, followed by autoradiography. For the carboxylation activity, soluble product of translation reactions carried out in the presence of unlabelled methionine were used and activity was measured upon addition of purified RbcS (7 μM). Activitiy measured upon addition of RbcS to the product of translation in the presence of 0.5 μM/1.0 μM of GroEL/GroES is set to 100%. (B) Syn6301-*rbcLS* was cotranslated in presence of varying amounts of GroEL/ES; assembled RbcL$_8$S$_8$ was analyzed by Native-PAGE. Carboxylation activity was measured in soluble product of translation reaction. Activity measured upon translating *rbcLS* in the presence of 0.5 μM/1.0 μM of GroEL/GroES is set to 100%.

5.2.3 Conformational status of RbcL upon binding and encapsulation by GroEL

To investigate the conformational status of RuBisCO while bound and encapsulated by the GroEL/GroES system, intrinsic tryptophan fluorescence and l-anilinonaphthalenesulfonate (ANS) binding studies were carried out using RuBisCO from sources *Rhodospirillum rubrum* and *Synechococcus* sp. PCC6301.

Changes in tryptophan fluorescence would reflect conformational changes of the polypeptide chain. An increase in fluorescence is commonly associated with a transition from a predominantly solvent-exposed to a more hydrophobic environment, which is mostly brought about by a conformational change typically associated with folding (Schmid, 1990). There are six tryptophan residues in Rr-RbcL, nine in Syn6301-RbcL and none in GroEL and GroES. ANS (1-anilino-8-naphthalene-sulphonate) is commonly used as a probe for exposed apolar binding sites and in turn for alterations in protein tertiary or quaternary structure (Martin et al., 1991). Here, for the ANS experiments, single ring version of GroEL (SR-GroEL) (Cliff et al., 1999; Weissman et al., 1995) was used. SR-GroEL binds GroES in an ATP-dependent manner, but is unable to dissociate it due to the absence of an allosteric signal from the GroEL *trans*-ring. Thus GroES is believed to stably encapsulate protein substrate in the SR-EL/GroES complex without the possibility of active unfolding (Hayer-Hartl et al., 1996; Weissman et al., 1996)

A shift in the maximum emission of tryptophan fluorescence from 335 nm, which is indicative of Native Rr-RbcL (Figure 18Ai, black curve), to 365 nm was observed upon unfolding the protein in 6M guanidinium (Figure 18Ai, pink curve), accompanied by a strong decrease in the intensity of fluorescence. The emission maximum of GroEL-bound Rr-RbcL was at 348 nm (Figure 18Ai, green curve & blue curve), i.e. it shifted ~57 % from the denatured to the folded state. The high ANS-fluorescence suggests strong binding of ANS to the hydrophobic surfaces of

the early folding intermediates that are stabilized by SR-GroEL, presumably in a conformation lacking ordered tertiary structure (molten globule state) (Figure 18Bi).
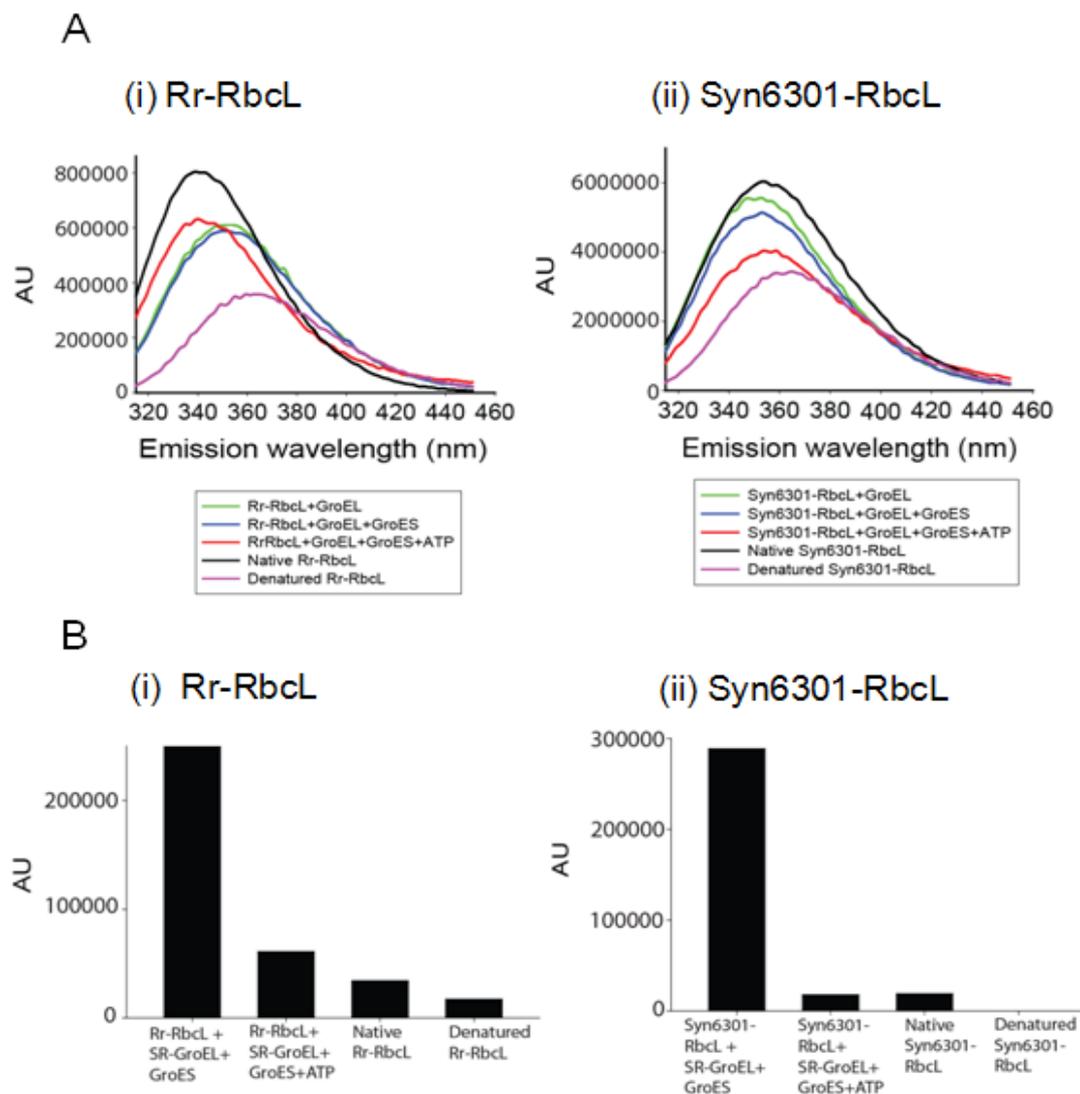
Figure 18. Tryptophan and ANS fluorescence studies of conformational changes in RuBisCO upon interaction with GroEL

A. Tryptophan fluorescence in RbcL refolding: Denatured Rr-RbcL$_2$ (A) or Syn6301-RbcL$_8$ (B) was diluted 100-fold (0.25 µM RbcL monomer) into ice-cold assay-buffer 1 containing 0.5 µM GroEL. After incubation for 10 min on ice, tryptophan fluorescence (excitation 295 nm, emission scan 315-450 nm) was measured. Then 1 µM GroES was added and tryptophan fluorescence was measured again. The reaction was supplemented with 2 mM ATP and refolding kinetics was observed by monitoring the change in tryptophan fluorescence over time. A tryptophan scan at the end of refolding was measured. If only native or denatured substrate had to be analyzed, reactions were modified accordingly. Background fluorescence of chemically identical reactions lacking RbcL was subtracted.

B. ANS-fluorescence in RbcL refolding: Denatured Rr-RbcL$_2$ or Syn6301-RbcL$_8$ were diluted 100-fold (0.25 µM RbcL monomer) into ice-cold assay-buffer 1 containing 1 µM ANS, 1 µM SR-GroEL and 2 µM GroES. After incubation for 10 min on ice, ANS fluorescence (excitation 390 nm, emission scan 420-550 nm) was measured. Refolding was started with the addition of 2 mM ATP and the kinetics was observed by monitoring the change in ANS fluorescence over time. At the end of refolding an ANS fluorescence scan was taken. If only native or denatured substrate had to be analyzed, reactions were modified accordingly. Data were corrected for background fluorescence of chemically identical reactions lacking RbcL and emission at 470 nm is depicted. (Experiment 18 was performed by Bharathi Vasudeva Rao)

The tryptophan fluorescence emission maximum shifted back to 335 nm upon addition of GroEL/GroES and Mg-ATP to the unfolded Rr-RbcL, identical to the emission maximum for the native protein (Figure 18Ai, red curve). This reflects the folding and reconstitution of Rr-RbcL dimer. This was further supported by the decrease in the intensity of ANS fluorescence that occurred during the folding process (Figure 18Bi).

Unfolding of Syn6301-RbcL$_8$ resulted in a redshift of the fluorescence maximum from 355 nm for native protein (Figure 18Aii, black curve) to 365 nm (Figure 18Aii, pink curve), with a decrease in the emission intensity of more than 40%. The emission maximum of GroEL-bound Syn6301-RbcL was 355 nm, identical to that of the native protein, but the fluorescence intensity was slightly lower (Figure 18Aii, green curve). Similar to the Rr-RbcL experiments, the high ANS- fluorescence suggested strong binding of ANS to the hydrophobic surfaces of the early folding intermediates stabilized by SR-GroEL (Figure 18Bii). The tryptophan fluorescence emission maximum was retained at 355 nm upon addition of GroEL/GroES alone or together with Mg-ATP to the unfolded Syn6301-RbcL, but a decrease in fluorescence intensity was observed (Figure 18Aii, blue curve & red curve), which probably reflect the monomeric folded state of RbcL. The decrease in the ANS-fluorescence again suggests a change in the tertiary structure of the enzyme molecule during the refolding reaction (Figure 18Bii).

5.3   Characterization of RbcX from *Synechococcus* sp. PCC7002

5.3.1 RbcX is necessary for the production of *Synechococcus* sp. PCC7002-RbcL$_8$

The chaperone requirement for the folding and assembly of another cyanobacterial form I RuBisCO from *Synechococcus* sp. PCC7002 (Syn7002) was also assessed upon *in vitro* synthesis in *E. coli* translation lysate (Figure19A). In contrast to *Synechococcus* sp. PCC6301-RbcL$_8$ folding and assembly, which requires only GroEL/GroES, Syn7002-RbcL$_8$ was not produced by GroEL/GroES alone, even when supplemented at high concentrations (19A, Lanes 2-4).

In Syn7002, the *rbcX* gene, which encodes the protein RbcX, is present between the *rbcL* and *rbcS* genes in an operon. To understand the mechanism of RbcX action in RuBisCO biogenesis, the functional analysis of Syn7002-RbcX protein was undertaken. The Syn7002-*rbcX* gene was cloned into pET11a vector and expressed in *E. coli* as a soluble protein of 15 kDa on SDS-PAGE (Figure 24A). Indeed, the assembly of Syn7002-RbcL$_8$S$_8$ proved to be critically dependent on coexpression with RbcX (Figure19A, lane 5 & (Onizuka et al., 2004)). Supplementing purified GroEL/GroES along with RbcX caused a more than 2-fold increase in the amount of RbcL$_8$, detected by Native-PAGE (Figure 19A, Lane 6). The active holoenzyme was produced upon addition of RbcS protein to the RbcL$_8$ product (Figure 19A). Thus, RbcX appears to be required predominantly or exclusively for RbcL$_8$ formation.

The requirement of GroEL/ES and RbcX for efficient production of Syn7002-RbcL$_8$ was also tested by the coexpression studies *in vivo* in *E. coli*. Amounts of soluble protein were determined upon fractionation of cell extracts and immunoblotting. Formation of RbcL$_8$S$_8$ holoenzyme (Figure 19C) or RbcL$_8$ complexes (Figure 19D) was monitored by Native-PAGE and RuBisCO activity assay. In wild-type cells, expression of Syn7002-RbcL from an IPTG-inducible plasmid did not result in soluble protein, irrespective of whether RbcS was coexpressed (Figures 19C and 19D, Lane 1).

Overexpression of GroEL/GroES prior to the induction of *rbcLS* or *rbcL* also failed to produce $RbcL_8S_8$ or $RbcL_8$ complex (Figures 19C and 19D, Lane 2). Assembled RuBisCO and activity were essentially absent (Figures 19C and 19D, Lanes 1–2). Importantly, when Syn7002-*rbcL* was expressed in the absence of RbcS, RbcX was required for the formation of $RbcL_8$ core complexes competent for assembly to active enzyme with RbcS (Figure 19D, Lane 3). Overexpression of GroEL/GroES in the presence of rbcX caused a ~3 fold increase in Syn7002 RuBiSCO activity, correlating with an increased amount of $RbcL_8S_8$ or $RbcL_8$ detected by Native-PAGE (Figures 19C and 19D, Lane 4). $RbcL_8$ alone (in the absence of RbcS) does not show any RuBisCO activity. Since the highest yield of $RbcL_8$ complex was obtained when Syn7002-*rbcL* was translated *in vitro* in the presence of 40 µM of Syn7002-RbcX protein (Figure 19B), further translation experiments were generally carried out in the presence of 40 µM of Syn7002-RbcX.

In the case of RuBisCO from Syn6301, where RbcX is encoded outside the *rbcLS* operon, RbcX is not necessary for the formation of $RbcL_8$ core complex (Figure 20, Lane 3), but RbcX strongly increases the yield of $RbcL_8$ and of RuBisCO activity more than two fold upon addition of RbcS (Figure 20, Lane 4). Translating Syn6301-*rbcL* in presence of RbcX from another cyanobacterium, *Anabaena* sp. CA, produced a slower migrating high molecular weight complex, which was not competent to form active enzyme upon addition of RbcS (Figure 20, Lane 5).
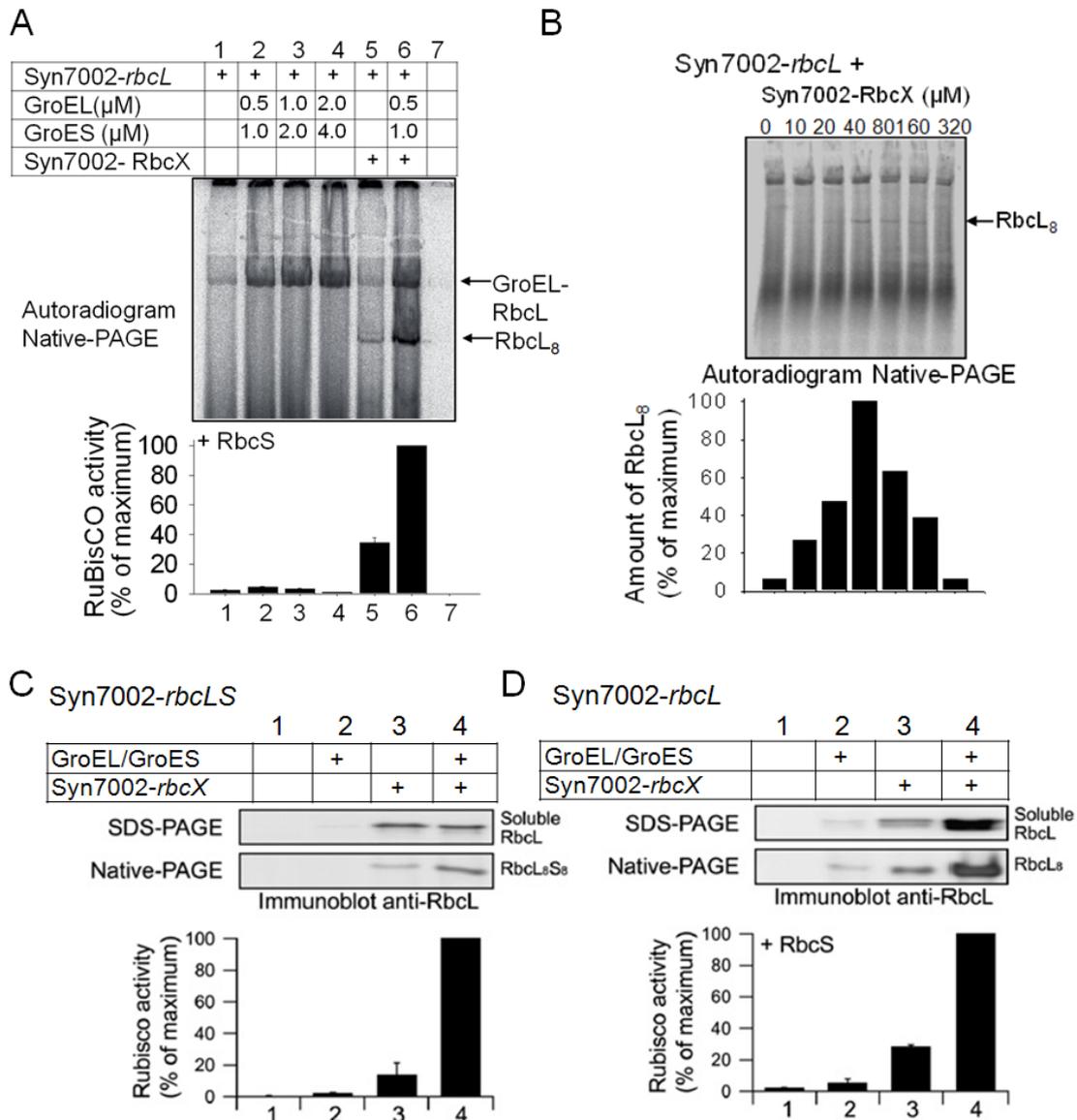
Figure 19. Requirement of the chaperonin system and RbcX for *Synechococcus* sp. PCC7002- $RbcL_8$ complex formation.

(A) Syn7002-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine (1.5 hr, 30[0]C). When indicated, GroEL/GroES (0.5 µM/1.0 µM) and Syn7002-RbcX (40 µM) were added to lysate. Assembled $RbcL_8$ and total RbcL protein were analyzed by Native-PAGE and SDS-PAGE, respectively, followed by autoradiography. For the carboxylation activity, the soluble products of the translation reactions carried out in the presence of unlabelled methionine were used and activity was measured upon addition of purified RbcS (7 µM). Activities measured upon Syn7002-rbcL translation in presence of GroEL/GroES and RbcX are set to 100%. Error bars indicate standard deviation of three independent experiments.

(B) Syn7002-*rbcL* was translated in absence of GroEL/GroES but in presence of varying amounts of Syn7002-RbcX; assembled $RbcL_8$ was analyzed by Native-PAGE followed by autoradiography.

Below is the quantification of $RbcL_8$ bands. Intensity measured upon Syn7002-rbcL translation in presence of 40 µM RbcX is set to 100%.

(C and D) Syn7002-*rbcL* and *rbcS* (C) or Syn7002-*rbcL* (D) was expressed in *E. coli* with or without coexpression of Syn7002-*rbcX* and GroEL/GroES as indicated. RbcL in soluble cell lysates was analyzed by SDS-PAGE and $RbcL_8S_8$ or $RbcL_8$ complexes by Native-PAGE and immunoblotting. Carboxylation activity was measured directly in soluble cell lysates (C) or upon addition of purified RbcS (D). Activities measured upon coexpression with GroEL/GroES and RbcX are set to 100%. Error bars indicate standard deviation of three independent experiments (Experiments in C and D were performed by Sandra Saschenbrecker).
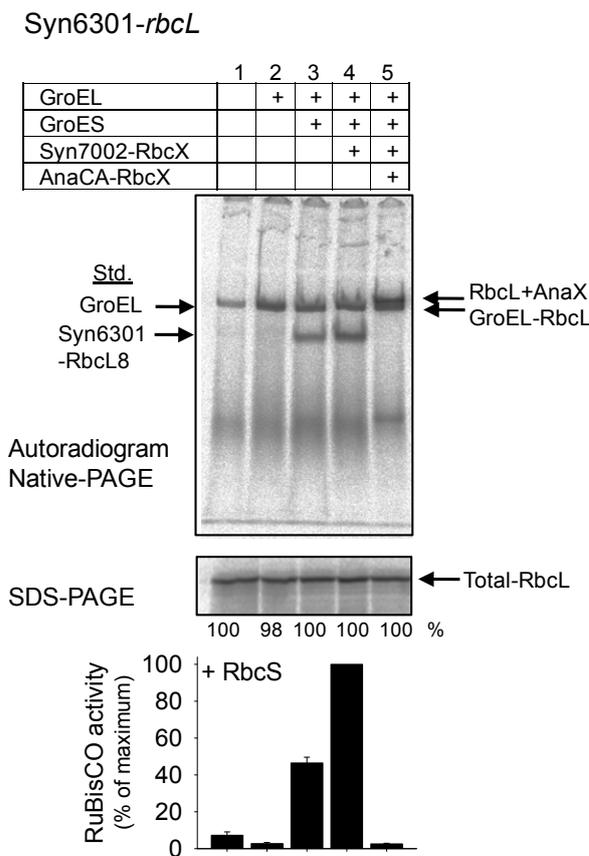
Syn6301-*rbcL*

Figure 20. Role of RbcX in *Synechococcus sp.* PCC6301-RbcL$_8$ formation.

Syn6301-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine (1.5 hr, 30^0C). When indicated, GroEL (0.5 µM) alone or with GroES (1.0 µM) and Syn7002-RbcX (40 µM) or AnaCA-RbcX (40 µM) were added to the lysate. Assembled $RbcL_8$ and total RbcL protein were analyzed by Native-PAGE and SDS-PAGE, respectively, followed by autoradiography. For the carboxylation activity the soluble product of translation done in presence of unlabelled methionine was used and activity was measured upon addition of purified Syn6301-RbcS (7 µM). Activities measured upon Syn6301-*rbcL* translation in presence of GroEL/GroES and Syn7002-RbcX are set to 100%. Error bars indicate standard deviation of three independent experiments.

5.3.2 Sequential action of chaperonin and RbcX in the formation of Syn7002-RbcL$_8$

As mentioned above, *E. coli* translation lyasate contains limited amounts of chaperones, including GroEL/GroES (Agashe et al., 2004). To investigate the stages at which GroEL/GroES and RbcX act, an *in vitro* translation experiment was

carried out in the normal lysate or the GroEL depleted lysate (Figure 21B). Translation of Syn7002-*rbcL* in the normal *E. coli* lysate, supplemented with GroEL/GroES, but not GroES alone, together with RbcX greatly increased the formation of RbcL$_8$ core complexes (Figure 21, Lanes 3 and 4). Translation of *rbcL* in GroEL-immunodepleted lysate (Figure 21B) did not produce RbcL$_8$, even when RbcX was present; demonstrating the requirement of both factors (Figure 21, Lanes 5–7). Importantly, re-addition of GroEL after inhibition of *rbcL* translation with chloramphenicol (CAM) failed to produce RbcL$_8$, despite the presence of RbcX during translation (Figure 21, Lane 8). Note that some RbcL was bound to GroEL under these conditions, but this protein was not productive for folding. Thus, RbcX cannot stabilize RbcL for subsequent productive interaction with GroEL, suggesting that GroEL must interact with RbcL immediately upon synthesis. Together these results indicate that RbcX has a specific chaperone function in preventing RbcL misassembly downstream of chaperonin.
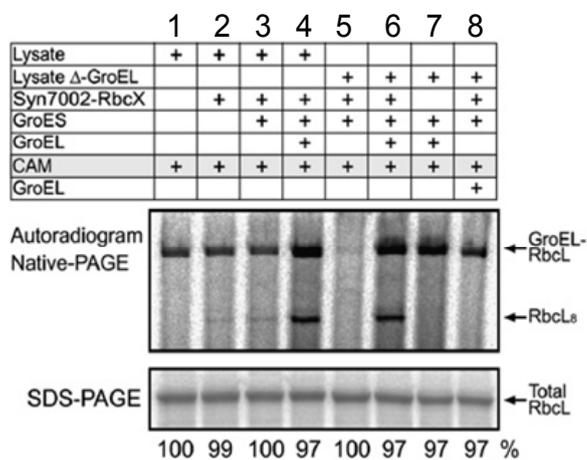
A   Syn7002-*rbcL*   B
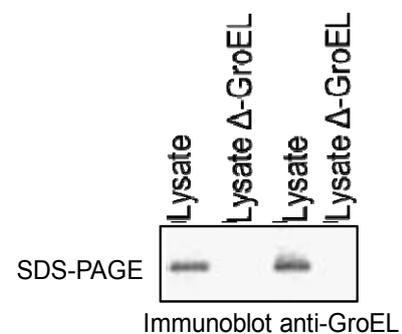
Figure 21. Sequential action of chaperonin and RbcX in assembly of Syn7002-RbcL$_8$. (A) Syn7002-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine (1.5 hr, 30C). When indicated, GroEL/GroES (0.5 µM/1.0 µM) and Syn7002-RbcX (40 µM) were added to normal lysate (Lanes 1–4) or GroEL-depleted lysate (Δ-GroEL, Lanes 5–8). In lane 8, GroEL was added after stopping translation with CAM. Assembled RbcL$_8$ and total RbcL protein were analyzed by Native-PAGE and SDS-PAGE, respectively, followed by autoradiography.

(B) Efficient depletion of GroEL from *E. coli* lysate. GroEL depletion was done by passing the *E. coli* lysate through GroEL antibody-bound Protein A sepharose beads and the product was analyzed by SDS-PAGE followed by immunoblotting using anti-GroEL antibody.

5.3.3 Absence of RbcX results in non-functional aggregates of RbcL subunits

To investigate the affect of absence of RbcX on the RbcL assembly, the following *in vitro* translation experiment was carried out. The production of $RbcL_8$ was most efficient when the lysate was supplemented with GroEL/GroES and RbcX during translation of Syn7002-*rbcL* (Figure 22, Lane 1). Addition of RbcX immediately after blocking translation still mediated $RbcL_8$ formation, albeit at a reduced level (Figure 22, Lane 2). Delayed posttranslational addition of RbcX resulted in the rapid loss of RbcL assembly and resulted in the formation of insoluble RbcL (Figure 22, SDS-PAGE). This result indicates that RbcX has a specific chaperone function in preventing RbcL misassembly downstream of GroEL. In the absence of RbcX, RbcL subunits form non-functional aggregates.

5.3.4 Recycling of Syn7002-RbcL on GroEL

The above results suggest that the folding of RbcL occurs inside the chaperonin cage. As mentioned above, many proteins, including Syn6301-RbcL, must undergo multiple interactive cycles to reach their native state. Incompletely folded polypeptide chains are released from GroEL and are rapidly recaptured for further folding attempts (Hartl and Hayer-Hartl, 2002; Mayhew et al., 1996).

In order to test whether Syn7002-RbcL also recycles to GroEL, an *in vitro* translation in normal *E. coli* lysate supplemented with a constant amount of RbcX and varying amounts of GroEL/ES was carried out. The $RbcL_8$ complex was monitored by Native-PAGE analysis and autoradiography (Figure 23A and 23B). As observed with *Synecococcus sp.* PCC6301 RbcL peptides, the addition of 0.5µM/1.0µM of GroEL/GroES (Figure 23A, Lane 4), but not higher concentrations

(Figure 23A, Lanes 5-7), together with RbcX increased the formation of Syn7002-RbcL$_8$ core complexes. Supplementing higher amounts of GroEL/ES decreased the amount of RbcL$_8$ complex made, in a concentration- dependent manner, suggesting the re-cycling of RbcL to GroEL/ES system (Figure 23A and 17B).
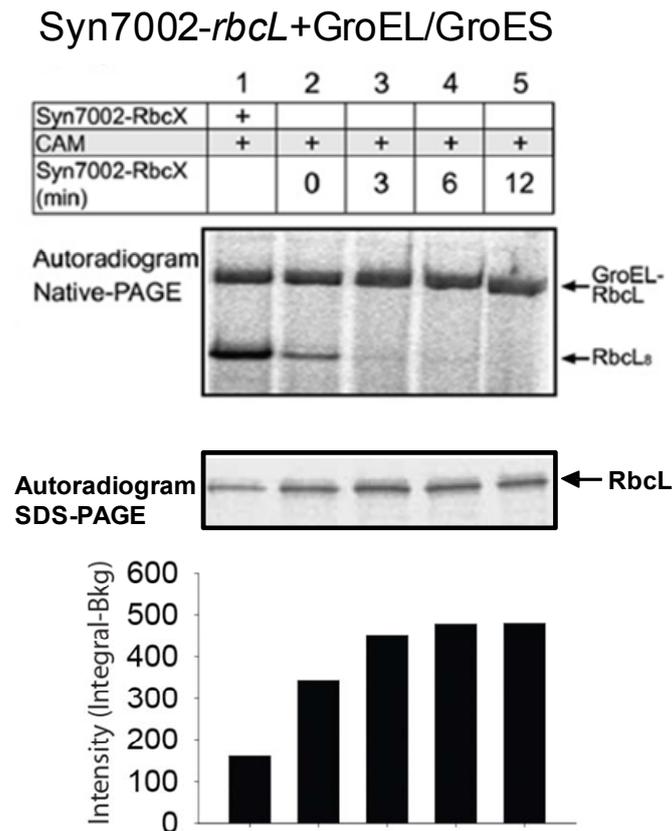
Figure 22. Role of RbcX in *Synechococcus* sp. PCC7002-RbcL$_8$ formation.
Syn7002-*rbcL* was translated in the presence of GroEL/GroES. RbcX was either present during translation (Lane 1) or was added together with CAM (lane 2) or after CAM addition (Lanes 3–5) and assembled RbcL$_8$ and RbcL in the pellet fraction were analyzed by Native-PAGE and SDS-PAGE, respectively, followed by autoradiography. Below is the quantification of bands on SDS-PAGE.

In another experiment, translation of Syn7002-*rbcL* was carried out in the presence of 0.5µM/1.0µM GroEL/ES and the translation was inhibited by the addition of chloramphenicol. Additional GroEL/GroES was supplemented in varying concentrations along with a constant amount of Syn7002-RbcX to the products of

translation and the formation of $RbcL_8$ was monitored by Native-PAGE. This experiment also showed reduced $RbcL_8$ formation upon addition of excess of GroEL/GroES, thereby demonstrating recycling of RbcL protein to the chaperonin system (Figure 23B).

A

Syn7002-*rbcL* + Syn7002-RbcX

	1	2	3	4	5	6	7
GroEL (µM)	-	0.125	0.25	0.5	1.0	2.0	4.0
GroES (µM)	-	0.25	0.5	1.0	2.0	4.0	8.0

GroEL-RbcL→

RbcL₈→

Autoradiogram
Native-PAGE

SDS-PAGE ← Total RbcL

100　98　100　100　100　99　100　%

B

Syn7002-*rbcL* + GroEL/ES

	1	2	3	4	5
CAM	+	+	+	+	+
Syn7002-RbcX	+	+	+	+	+
GroEL/ES (µM)	0/0	0.5/1.0	1.0/2.0	2.0/4.0	4.0/8.0

Autoradiogram
Native-PAGE

GroEL-RbcL
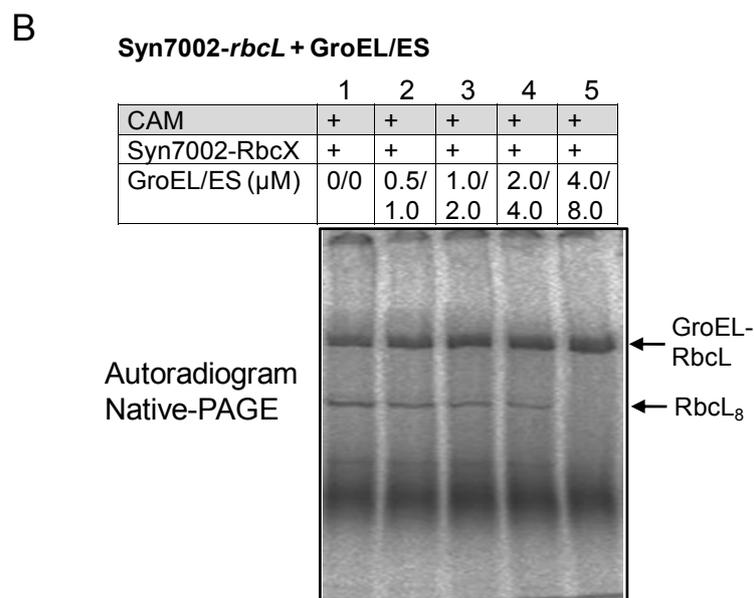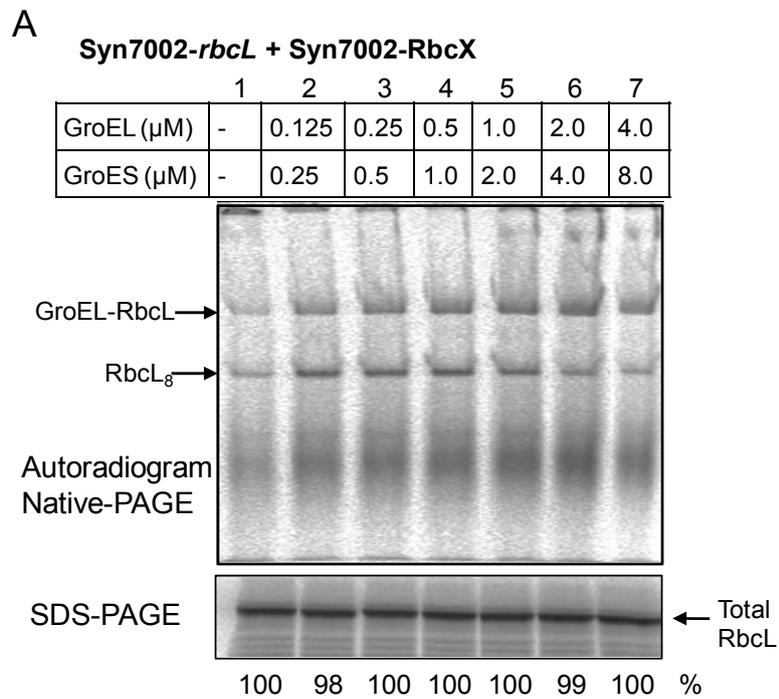
RbcL₈

Figure 23. Syn7002-RbcL recycles on GroEL

(A) Syn7002-*rbcL* was translated for 90 minutes in *E. coli* lysate *in vitro* in the presence of ^{35}S-Methionine, 40 µM of recombinantly purified Syn7002-RbcX protein and varying amounts of GroEL/GroES. The RbcL$_8$ complex and the total RbcL were analyzed by Native-PAGE and SDS-PAGE, respectively, followed by autoradiography.

(B). Syn7002-*rbcL* was translated for 90 minutes in *E. coli* lysate *in vitro* in the presence of ^{35}S-Met, 0.5µM/1.0µM of GroEL/GroES. Translation was inhibited by the addition of chloramphenicol. The products were supplemented with varying amounts of GroEL/ES and 40 µM of Syn7002-RbcX protein and incubated for additional 20 minutes at 30[0]C. The products were analyzed by Native-PAGE, followed by autoradiography.

5.3.5 Crystal structure of *Synechococcus* sp. PCC7002-RbcX

To understand the mechanism of RbcX action in RuBisCO assembly, a structural analysis of Syn7002-RbcX by X-ray crystallography was undertaken by my colleagues (Sandra Sachenbrecker and Dr. Andreas Bracher) (Saschenbrecker et al., 2007). RbcX was expressed in *E. coli* as a soluble protein of 15 kDa on SDS-PAGE (Figure 24A). Analysis by multiangle light scattering indicated that RbcX forms a dimer of 30.2 kDa (Figure 24B).

SDS-PAGE & Native-PAGE
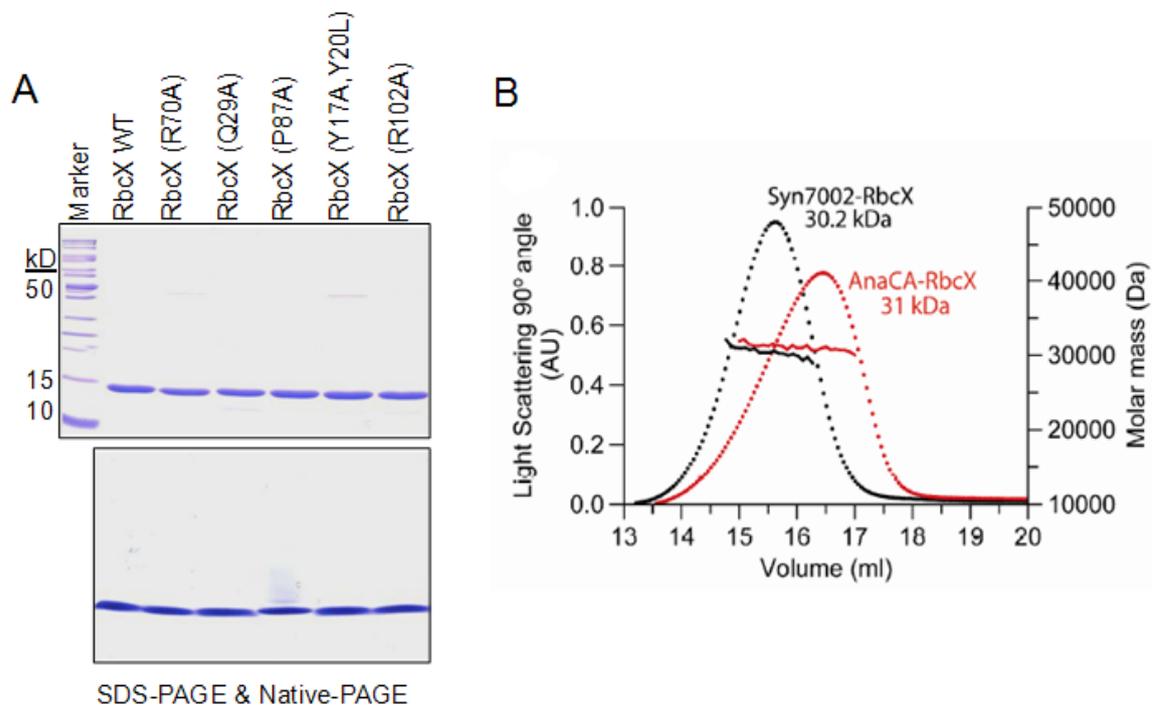
Figure 24. Mass analysis of recombinantly expressed RbcX proteins.
(A) SDS-PAGE and Native-PAGE analysis of recombinantly purified RbcX wild type (WT) and mutant proteins followed by coomassie staining.
 (B) FFF-MALS data showing the analysis of AnaCA-RbcX (31040 Da.) and Syn7002-RbcX (30170 Da.). Horizontal lines across the peaks indicate molar mass and homogeneity of the sample. FFF-MALS was performed in cooperation with Dr. Manajit Hayer-Hartl.

RbcX consists of four α helices per monomer (α 1–4) that form an unusual helix bundle (Figure 25A). α2 (residues 35–48) turns backward relative to α1 (residues 4–33) with a steep interhelical angle of 151^0 and only one residue (N34) forming the turn. A very similar arrangement is observed at the junction between α3 (residues 52–63) and α4 (residues 65–107; interhelix angle 161^0). The short helices α2 and α3 are connected by a five-residue linker. The core of the helical bundle is composed of conserved hydrophobic residues without authentic coiled-coil side-chain packing (Figure 25A and 25B). α4 makes an ~60^0 kink in the vicinity of residue 84 and forms a 35 A° long extension (α4C) pointing away from the helix bundle (Figure 25A) (Saschenbrecker et al., 2007).

The RbcX dimer has overall dimensions of ~64X33X31 A° (length X height X width). A polar network around the conserved residue N98, contacting the amide backbone at position L72 and the side chains of R75 and E76 in the opposing protomer, contributes to dimer stability. The α2-α3 helical hairpins make no direct interchain contacts (Saschenbrecker et al., 2007). Crystallographic analysis of the RbcX from the cyanobacterium *Anabaena* sp. CA (AnaCA) revealed a structure virtually identical to that of Syn7002-RbcX consistent with their high degree of sequence conservation (Figure 26). This suggests that all cyanobacterial RbcX homologs have the same overall architecture (Saschenbrecker et al., 2007).

A

B
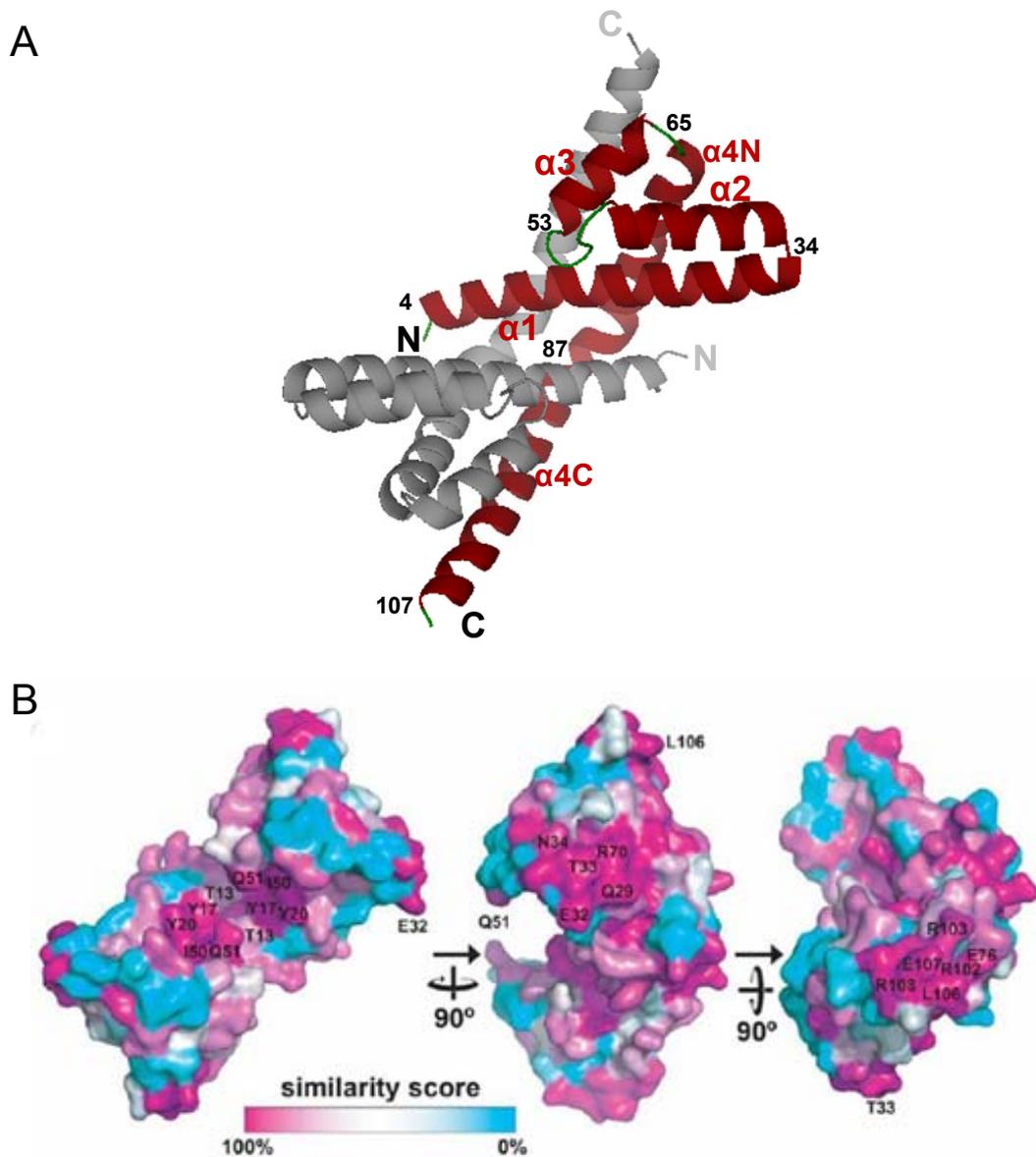
Figure 25. Crystal structure of Syn7002-RbcX

(A) Ribbon representation of the RbcX dimer. The promoters are shown in gray and red. The peptide backbone is depicted from N to C terminus for one of the promoters. For one of the promoters, the secondary structure elements, selected residue numbers, and chain termini are indicated. Alpha helices and loops are shown in red and green respectively.

(B) Surface conservation in RbcX. The similarity score from an alignment of 151 sequences of cyanobacterial RbcX in the PFAM database was plotted onto the accessible surface of the RbcX dimer. Sequence conservation is indicated by a color gradient, indicating highly conserved residues in magenta and variable regions in cyan. The positions of conserved surface residues are indicated (Saschenbrecker et al., 2007) (This work was performed by Sandra Saschenbrecker and Dr. Andreas Bracher).

5.3.6 Conserved regions on Syn7002-RbcX

Two highly conserved regions, representing potential protein-protein interaction sites, were identified on RbcX, the central groove of the dimer and a surface region around the corners of the molecule (Figure 25B). A hydrophobic area comprising the conserved residues Y17, Y20, and I50 of each monomer lines the groove in the center of the molecule (Figure 25B). Central access into the crevice is constricted to a 5.4Å wide opening by the side chains of the highly conserved Q51 residues that is just wide enough to accommodate a polypeptide chain in an extended conformation (Saschenbrecker et al., 2007).

The other conserved region has predominantly polar character and is located at the corners of the RbcX dimer, comprising residues Q29, E32, T33, N34, R70, E76, and R102, R103, L106, E107, R108, respectively (Figure 25C). It is formed by helices 4 of both protomers and by the turn region between the α1 and α2 helices of one subunit and may serve to interact with a polar protein surface. Because of the 2-fold symmetry, this region occurs twice at opposing edges of the dimer.

5.3.7 Mutational analysis of Syn7002-RbcX

The possible functional significance of the conserved regions of RbcX was investigated by mutational analysis by Sandra Saschenbrecker (Saschenbrecker et al., 2007). Coexpression of RbcL and RbcX (wt or mutants) in *E. coli* demonstrated that the mutations at the conserved residues Y17, Y20, Q29, E32, R70 and R102 were crucial for the function of RbcX (Saschenbrecker et al., 2007).

The above mentioned mutations, which showed noticeable effect, were further characterized using *in vitro* translation experiments, by translating Syn7002-*rbcLS* in the presence of recombinantly purified mutant RbcX proteins (Figure 27).

Synechococcus7002

```
Synechococcus7002   ..............................................................................
Microcystis         ..............................................................................
Synechocystis6803   ..............................................................................
Synechococcus6301   ............................................................................MQ
Thermosynechococcus ..............................................................................
Anabaena7120        ..............................................................................
AnabaenaCA          ..............................................................................
Nodularia           ..............................................................................
Tychonema           ..............................................................................
Arabidopsis_geneI   MVSAWPVVGSPVMDSSSSPCLCLDAHTTGTIRR....KKILGKARNLELGSSFTGSRIVFRLSPKRVSRI
Vitis               MSGALSVVGSSIFDPKTCPCLCLDALPSSNMSRNSVGKKRLSRPGTVEMGSSFVDSWHEWRLEAKSSRR.
Oryza_geneI         ..........MAGVQVMPAVGAVATAEGSCRA...AGQRRGRGVSSLFAGDWRRRPRRAACTARVRGRR
Oryza_Indica        ..........MAGVQVMPAVGAVAAAEGSCRA...AGQRRGRGVSSLFAGDWRRRPRRAACTARVRGRR
Oryza_geneII        ..........MAAVARMAAAVVCLPSSSSSSS...ACRAAPLPWSRGVVVGVRRR.RTVARAARRRGRR
Arabidopsis_geneII  ...................MESSSSLL...HHSYLSYLNPKFGKRPLVSYPLMQSSRK...CKQ
Oryza_geneIII       ..............MASPRCAAVALLHPAGVAA...GGGARRRVLLLDQERPLWGTEVRRRRRRRFSSLE
Chlamydomonas       ...............................MSLSARAAFRAPLSHHRPRPQAYPRA
consensus>50        ..............................................................................
```

Synechococcus7002

```
                                              α1                        α2
                    2000000000000000000000000000000.......00 0000000000
                    1        10        20        30                40
                    :         :         :         :               :
Synechococcus7002   ..............MEFKKVAKETAITLQSYLTYQAVRLISQQLS......ETNPGQAIWLGEFSKR
Microcystis         .......MYDKKVVQDTAKVLQSYLTYQAVRTIIDQLS......ETNPTLAIWLSHYTSS
Synechocystis6803   ......MFMQTKHIAQATVKVLQSYLTYQAVLRIQSELG......ETNPPQAIWLNQYLAS
Synechococcus6301   FMGTASRMASTQRAKPMEMPRISRDTARMLVNYLTYQAVCVIRDQLA......ETNPAGAYRLQVFSAE
Thermosynechococcus .......MDVKHIAKQTTKTLISYLTYQAVRTVIGQLA......ETDPPRSLWLHQPTSQ
Anabaena7120        .......MNLKQIAKDTAKTLQSYLTYQALMTVLAQLG......EMNPPLALWLHTPSV.
AnabaenaCA          .......MNLKQIAKDTAKTLQSYLTYQALRTVLAQLG......ETNPPLALWLHNPSA.
Nodularia           .......MNLKQIAKDTAKTLQSYLTYQALKTVLAQLG......ETNPPLELWLHNPSA.
Tychonema           .......MDLKQVAKDTAKVLASYLTYQSVRIVIAQLS......ETNPPLAIWLNEPSTK
Arabidopsis_geneI   ANRKSKKLLIVNEDVAGNYDDTPGDVQKQIVNYFTYKAVRTVLHQLY......EMNPPQYTWFYNHIIT
Vitis               .QHKARRLLVVSE.VAGQYEDSFEDVKTQILNYFTYKAVRTVMHQLY......EMNPTQYRWLYDPVAA
Oryza_geneI         QQQQQQGLAVVCN.LGGTYDEGFEDIHVQLINVFTYKAVKTVLTQLY......EMNPPSYRWLYNPVAV
Oryza_Indica        QQQQQQGLAVVCN.LGGTYDEGFEDIHVQLINVFTYKAVKTVLTQLY......EMNPPSYRWLYNPVAV
Oryza_geneII        PGRR..GLVIVDE.FAGQYEEGFEDVHTEIMNYFTYKATSTVLHQLY......EMNPPAYTWLYNYVVV
Arabidopsis_geneII  TRICSNKMYVPG..PGBASPE..AKAAKHLHDFFTYVAVRIVSAQLB......SYNPEAYMELREPLDT
Oryza_geneIII       TPPRCSKMYVPG..FGEGSPE..KKAARNLQHFFNYIAVRVVLTQLB......SYNREAYGELMDPVNR
Chlamydomonas       VVTPCARMHIPADSFSGASPE..RKAAVALRSLFTFVAARVVLEQLQGPGGPETTYNQQAYLDLMDPLGT
consensus>50        ........1.........mn..ei..dtak.1.sylt%qavrt!1.#L.........e.#pp...wl.nf...
```

Synechococcus7002

```
                        α3           α4N          η1          α4C
                    0.0000000  00...000000000..000000000000000000000000000
                    50        60        70        80        90       100       110
                    :         :         :         :         :         :         :
Synechococcus7002   HPIQES.DLYLEAMMLENKE...LVLRILTVR..ENLAEGVLEFLPEMVLSQIKQSNGNHRRSLLERLTQ
Microcystis         HSIQDG.EAYIAGLMTENKE...LVLRIMTVR..EHLAEQVLEFLPEMVKTAIINDNTEHRRQLLERLTR
Synechocystis6803   HSIQNG.ETFLTELLDBNKE...LVLRILAVR..EDIAESVLDFLPGMTRNSLAESNIAHRRHLLERLTR
Synechococcus6301   FSFQDG.BAYLAALLNHDRE...LGLRVMTVR..EHLABHILDYLPEMTIAQIQBANINHRRALLERLTG
Thermosynechococcus ESIQDG.ERYLEALFREQPD...LGFRILTVR..EHLAEMVADYLPEMLRAGIQQANLQQRCQQLERMTQ
Anabaena7120        GKVQDG.EAYVKELFREQPD...LALRIMTVR..EHIAEEVAEFLPEMVRSGIQQANMQRRQHLERVTH
AnabaenaCA          GKVQDG.EKYIEELFLBKPD...LALRIMTVR..EHIAEEIAEFLPEMVVTGIQQANMEKRRQHLERMTQ
Nodularia           GKVQDG.EAYLEQLFQAKPD...LALRIMTVR..EHIAAEVAEFLPEMVLTGIQQGNMQRRQHLERMTF
Tychonema           GKIQDG.ELYIRELLLENQD...LGFRIMTVR..EYLAHDVTEFLPEMVRTGIQQSNMEHRRQHLERITQ
Arabidopsis_geneI   NRPTDG.KRFLRALGKESQE...LABRVMITR.LHLYGKWIKKCDHGKIYQEISDENLALMRERLMETVI
Vitis               NKPGEG.KRFIRILVKEKQE...LABRVMVTR.LHLYGKWIKKCNHABIYKEISDENLELMRBRLMETVK
Oryza_geneI         NKPTDG.KVFLRALGKEKQE...LABRVMITR.LHLYSKWIKKCDHAMMYERISDENLALMRERLMETVI
Oryza_Indica        NKPTDG.KVFLRALGKEKQE...LABRVMITR.LHLYSKWIKKCDHAMMYERISDENLALMRERLMETVI
Oryza_geneII        NDPKEG.KHFLIALAKRRQD...LABRVMITR.LHLYSKWIKKCDHAKMYEKISNENLEIMRQRLMETVA
Arabidopsis_geneII  NSVSDG.DKFCATLMRRSSRHMNLALRILEVRSAYCKNDFEWDNMKRLAFKNVDDSNTRLMREYVLETSH
Oryza_geneIII       NSLNDA.DTFCKKLIRESPRHKQLAMRILEVRSAYVKHDFEWDNLKRLSFKMVDBANTKLMRDYVLETSH
Chlamydomonas       PMKGDGGDEWMAAVMRKNHA...LALRLMVREAYL.DEFEWGKTMEMASRETREANTRLMRAAAMASLQ
consensus>50        ..iq#g.e.yl..l..e..e...LalRi$.vR..e...ae.v.eflpem....i.#.Nle.rrq.ler.t.
```

Synechococcus7002

```
                    120       130
                    :         :
Synechococcus7002   VDSSST.DQTEPNPGESDTSEDSE................
Microcystis         TAISQP.ETSELNLDVDDLPNP..................
Synechocystis6803   TVAEVD.NFPSETSNGESNNNDSPPS..............
Synechococcus6301   LGAEPSLPETEVSDRPSDSATPDDASNASHAD........
Thermosynechococcus VSEANVENSNLETPE.........................
Anabaena7120        LSLSNPSPESEQQTI.SDTDWDH.................
AnabaenaCA          VSLSHPSPESEQQQF.SDPDWDNLAS..............
Nodularia           IDTSDSSPQSEQQTS.SDPNSEQ.................
Tychonema           LNWSVATSNPETSSIDSEPNLDNLSS..............
Arabidopsis_geneI   WPSDDTNSR...............................
Vitis               WPSDDTNTEKIG............................
Oryza_geneI         WPTDDTNTEKIG............................
Oryza_Indica        WPTDDTNTEKIG............................
Oryza_geneII        WPTDDTNTSDTAK...........................
Arabidopsis_geneII  VETDSDK.................................
Oryza_geneIII       IEDDN...................................
Chlamydomonas       VWRQEAGGRRCRRRLHVYGGPGRPRQGCRVMGGNGAASPRV
consensus>50        ...dd...e...............................
```

Figure 26. Alignment of amino acid sequence of selected RbcX homologs

Amino acid sequences of RbcX homologs from *Synechococcus sp.* PCC7002 PR-6 (Q44177), *Microcystis aeruginosa* (Q9R3Q3), *Synechocystis sp.* PCC 6803 (Q55670), *Synechococcus elongatus* PCC6301 (Q5MZ09), *Thermosynechococcus elongatus* (Q8DIS6), *Anabaena* sp.PCC7120 (Q44307), *Anabaena sp.* CA (Q44212), *Nodularia harveyana* BECID29 (Q5K2U3), *Tychonema bourrellyi* (O86944), *Arabidopsis thaliana*_gene I (NP_568382), *Vitis vinifera* (CAO21907), *Arabidopsis thaliana*_gene II (NP_567263), *Oryza sativa* (japonica cultivar-group)_gene I (Q84M38), *Oryza sativa* (Indica Group) (EAY92274), *Oryza sativa* (japonica cultivar-group)_gene II (Q7XIH6), *Chlamydomonas reinhardtii* (EDP09788), *Oryza sativa* (japonica cultivar-group)_gene III (Q7EZV3) were aligned using the Multialign sever. Residues indicated in red background are invariable and residues indicated in red are conserved in more than 50% of the sequences; a consensus sequences according to these criteria using the standard symbols of the MultAlin analysis program (Corpet, 1988) is given in the bottom row of the alignment. The secondary structure elements and the residue numbering for RbcX of *Synechococcus* sp. PCC7002 are indicated above the sequences. Residues of the central groove of RbcX contacting the phenylalanines and isoleucines in the C-terminal recognition motif of RbcL are indicated by pink and green circles, respectively. Red circles denote residues at the corner surfaces of RbcX implicated by mutagenesis in binding to an uncharacterized surface of RbcL. Black circles denote structural residues involved in dimerization and blue circles indicate residues forming the hydrophobic core of the four-helix bundle.

The double mutation Y17A, Y20L abolished the formation of soluble RbcL and inturn the formation of RbcL$_8$ (Figure 27A) (Saschenbrecker et al., 2007). The secondary elements of this mutant were similar to those of wild type RbcX (Figure 28A and 28B) and also the crystal structure was virtually identical to that of wild-type RbcX (Saschenbrecker et al., 2007), indicating the essential function of central groove of RbcX.

Interestingly, individual mutations Q29A, R70A and R102A allowed the accumulation of reduced soluble RbcL (Saschenbrecker et al., 2007) and increased RbcL in the pellet fraction (Figure 27C), thereby not supporting the assembly to RbcL$_8$ as efficiently as wild type (Figure 27A and 27B). Notably, all of these mutations mapped to the polar ends of the RbcX dimer, forming symmetrical, contiguous surfaces (Figure 25B).

A

Syn7002-*rbcLS*+GroEL/ES
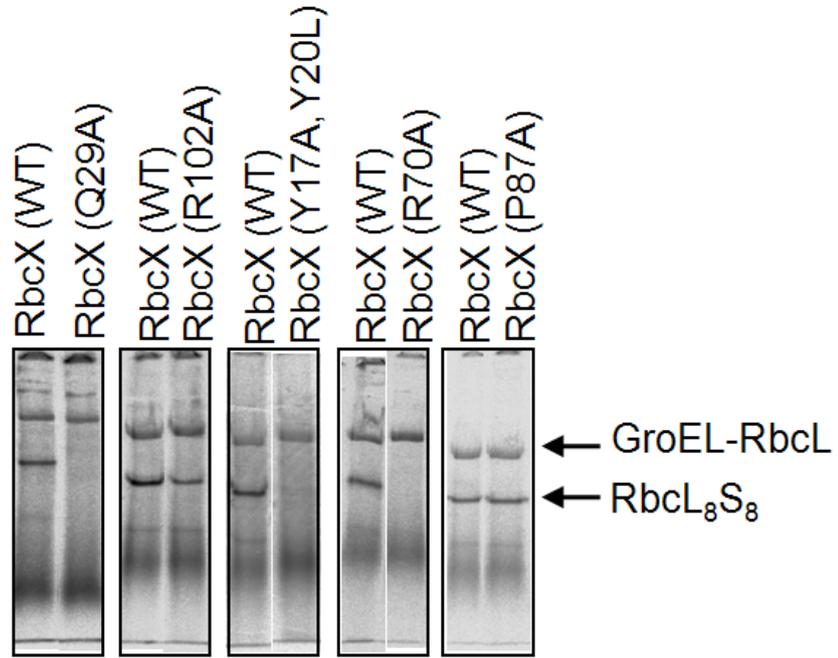
Autoradiogram Native-PAGE

B

C

Autoradiogram
SDS-PAGE
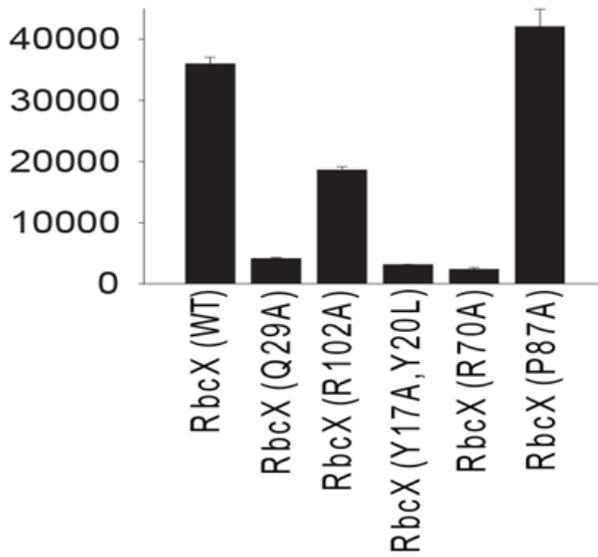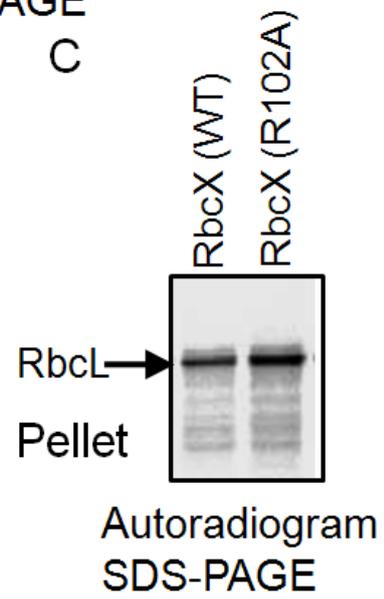
Figure 27. Functional analysis of Syn7002-RbcX mutants in *E.coli*-translation system *in vitro*

(A) *in vitro* cotranslation of Syn7002-*rbcLS* was carried out in the presence of recombinantly purified wild type or mutant RbcX proteins as indicated in the figure. $RbcL_8$ product was monitored on Native-PAGE followed by autoradiography.

(B) Caboxylation activity was measured as explained in the figure 19.

(C) Pellet fraction of the translation product of Syn7002-*rbcLS* carried out in the presence of RbcX wild type (WT) or RbcX (R102A) was analyzed by SDS-PAGE followed by autoradiography.
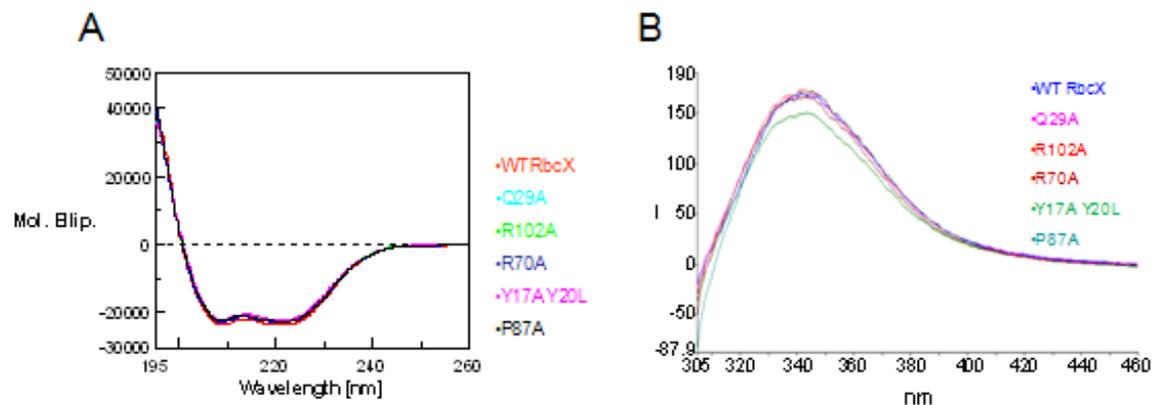
Figure 28. Secondary structure analysis of Syn7002-RbcX mutants

CD analysis (A) and melting curve analysis (B) of recombinantly purified wild type and mutant RbcX proteins was done as explained in materials and methods section.

Coimmunoprecipitation studies upon *E.coli* expressions performed by Sandra Saschenbrecker demonstrated that disruption of the polar peripheral surface of RbcX(Q29A) results in the misassembly of RbcL, though this RbcX(Q29A) can keep RbcL soluble (Saschenbrecker et al., 2007) .

In summary *in vivo* coexpression (Saschenbrecker et al., 2007) and *in vitro* translation experiments identified two independent, functionally critical regions on RbcX, the central hydrophobic crevice and a peripheral polar surface of the molecule. Whereas the central crevice is essential for the production of soluble RbcL subunits, disruption of the polar surface results in the production of soluble but misassembled RbcL, suggesting a role in proper subunit arrangement.

5.3.8 RbcX binds the conserved C-terminal peptide of RbcL

The structural features of the central groove region of RbcX seemed suitable for binding an extended peptide. To identify such a sequence element(s) in RbcL, an array of 12-amino-acid-long, acetylated peptides covering the entire sequence of Syn7002-RbcL with a ten residue overlap (231 peptides in total) was C-terminally attached to a cellulose-PEG membrane. Incubation with Syn7002-RbcX, followed by detection with anti-RbcX antibody, resulted in a strong binding signal with four RbcL peptides sharing the sequence EIKFEFD close to the C terminus of RbcL (Figures 29A and 29B). Other peptides exhibiting weaker binding had no apparent consensus motif. Essentially the same binding pattern was observed with the RbcX polar surface mutant Q29A, while the groove mutant Y17A,Y20L caused the complete loss of RbcX binding to the C-terminal RbcL peptide (Figure 29A).

Similarly, peptides sharing EIKFEFD of cyanobacterial RbcL, EITFNFP of *A.thaliana* RbcL, EIKFEFP of *S. oleracea* RbcL were detected when an array containing C-terminal RbcLs from different organisms including cyanobacteria, red algae and higher plants was incubated with Syn7002-RbcX$_{FLAG}$ protein followed by detection with anti-FLAG antibody (Figure 30A, i). Specific binding to the C-terminal EIKFEFD motif was also detected with RbcX from *Anabaena sp*. CA (Figure 29A). Incubating the array containing C-terminal RbcL from different organisms with either AnaCA-RbcX$_{FLAG}$ or Ana7120-RbcX$_{FLAG}$ and blotting with anti-FLAG antibody also detected EIKFEFD residues of C-terminal Syn7002-RbcL or similar motifs in other organisms (Figure 30A, ii and iii).
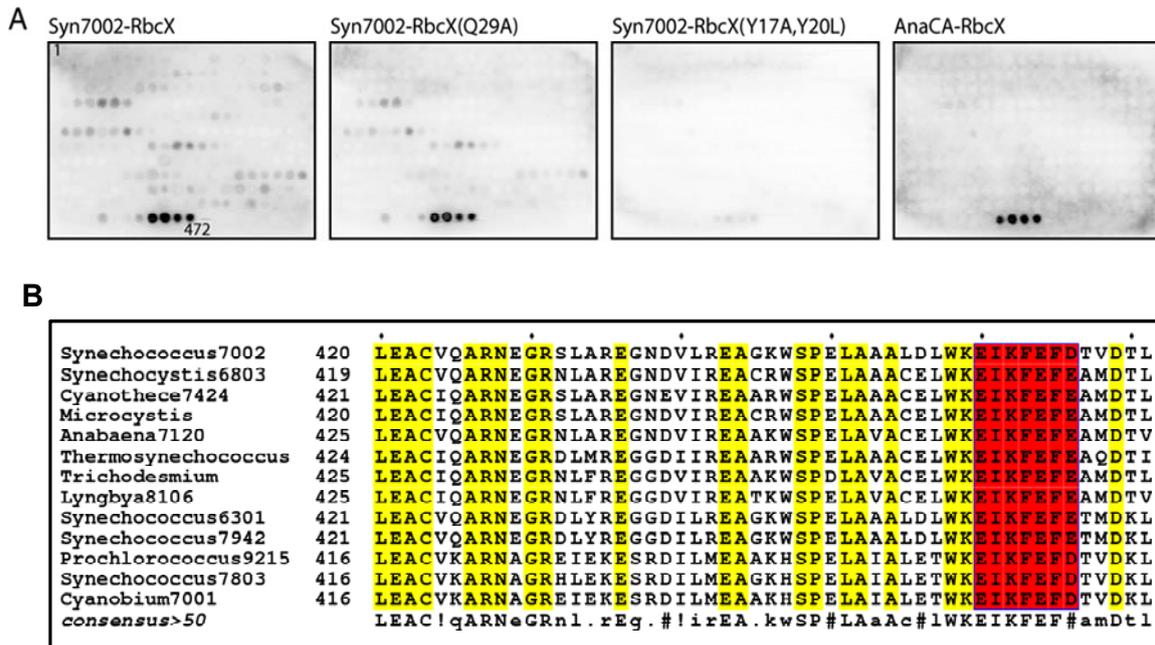
Figure 29. Binding of C-terminal RbcL peptide to RbcX

(A) A cellulose membrane containing an array of overlapping dodecamer peptides covering the sequence of Syn7002-RbcL was probed with the RbcX proteins indicated. Peptide-bound RbcX was visualized by chemiluminescent immunodetection with anti-RbcX antibody (This experiment was performed by Sandra Saschenbrecker) (Saschenbrecker et al., 2007).

(B) Alignment using MultAlin (Corpet, 1988) of C-terminal amino acid sequences of RbcL among cyanobacteria. *Synechococcus sp.* PCC 7002 (BAA03076), *Synechococcus elongatus* PCC 6301 (P00880), *Anabaena sp.* PCC 7120 (Nostoc sp. PCC 7120) (P00879), *Anabaena sp.7120* (P00879), *Synechocystis sp.* 6803 (P54205), *Trichodesmium erythraeum* IMS101 (Q10WH6), *Synechococcus elongatus* PCC 7942 (Q31NB3), *Prochlorococcus marinus* str. MIT 9215 (Q7V6F8), *Synechococcus sp.* WH 7803 (P96486), *Cyanobium sp.* PCC 7001 (CAM91978), *Cyanothece sp.* PCC 7424 (ZP_02975450), *Lyngbya sp.* PCC 8106 (ZP_01622123), *Microcystis aeruginosa* NIES-843 (YP_001659803), *Thermosynechococcus elongatus* BP-1(NP_682296). High consensus level (≥ 90%) is depicted in yellow background. Red background shows the conserved residues on C-terminus of RbcL which were detected in peptide screen by using RbcX protein.

A

(i) Syn7002-RbcX_FLAG anti-FLAG

(ii) Ana CA-RbcX_FLAG anti-FLAG

(iii) Ana 7120-RbcX_FLAG anti-FLAG

1. *Rhodobacter sphaeroides*
2. *Allochromatium vinosum*
3. *Nitrobacter winogradskyi*
4. *Synechococcus sp. PCC7002*
5. *Synechococcus sp. PCC6301*
6. *Anabaena sp. PCC7120*
7. *Chlamydomonas reinhardtii*
8. *Griffithsia pacifica*
9. *Suhria vittata*
10. *Arabidopsis thaliana*
11. *Oryza sativa*
12. *Spinacia oleracea*
13. *Nicotiana tabacum*
14. *Zea mays*
15. *Rhapis excelsa*

B

(i)

(ii)

C

D

Figure 30. Characterization of C-terminal RbcL peptide binding to RbcX

(A) A cellulose membrane containing an array of overlapping peptides covering the sequence of C-terminal Syn7002-RbcL was probed with the RbcX proteins indicated. Peptide-bound RbcX was visualized by chemiluminescent immunodetection with anti-FLAG antibody.

(B) (i) Comparison of C-terminal RbcL sequences. Alignment using MultiAlin of the RbcL sequences of the cyanobacteria, red algae and higher plants mentioned in the array containing RbcL C-terminal residues; (*Rhodobacter sphaeroides* (P27997), *Allochromatium vinosum* (P22849), *Nitrobacter winogradskyi* (AAD41022), *Synechococcus sp.* PCC7002 (Q44176), *Synechococcus sp.* PCC6301 (P00880), *Anabaena sp.* PCC7120 (P00879), *Chlamydomonas reinhardtii* (NP_958405), *Griffithsia pacifica* (AAQ57501), *Suhria vittata* (AAM91864), *Arabidopsis thaliana* (O03042), *Oryza sativa* (BAA00147), *Spinacia oleracea* (CAB88737), *Nicotiana tabacum* (CAA77361), *Zea mays* (CAA60294) and *Rhapis excelsa* (CAC17944). Number of the starting residue in the alignment of corresponding organism is mentioned before the sequence. High consensus level (≥90%) is depicted in brown background. The box highlighted shows the residues detected by RbcX protein binding.

(ii) Comparison of C-terminal RbcL sequences. Alignment using MultiAlin of the RbcL sequences of the cyanobacterial species *Synechococcus sp.* PCC7002 (Q44176), *Synechococcus sp.* PCC6301 (P00880), *Anabaena sp.* PCC 7120 (Nostoc sp. PCC 7120) (P00879); higher plants *Arabidopsis thaliana* (O03042), *Oryza sativa* (BAA00147); the archeal species *Theromococcus kodakaraensis* (O93627) and *Pyrococcus abyssi* (Q9UZD7) and the proteobacterium *Rhodospirillum rubrum* (Q2RRP5). Number of the starting residue in the alignment of corresponding organism is mentioned before the sequence. High consensus level (≥90%) is depicted in yellow background. The box with green background shows the conservation of RbcX binding motif on RbcL C-terminus.

(C) Structure of the complex of peptide EIKFEFD bound to RbcX dimer. The peptide is shown in stick representation; RbcX is represented as a molecular surface with protomers colored white and blue, respectively. N and C termini of the peptide are indicated.

(D) Magnification of boxed area in (C) presenting a view of the refined peptide bound to RbcX. Molecular interactions between peptide and RbcX are highlighted. Dashed lines represent hydrogen bonds. Residues of the RbcX monomers participating in peptide binding are displayed in stick representation below the transparent surface of the molecule and are numbered in white and yellow, respectively. The important hydrophobic residues in the bound peptide are also labeled (Figure C & D were made by Dr. Andreas Bracher) (Saschenbrecker et al., 2007).

These results suggest that this interaction is a general property of cyanobacterial and higher plant RbcX homologs. In the RuBisCO holoenzyme, the C-terminal sequence of RbcL is located at the surface of the complex and has been implicated in regulating catalysis (Zhu, 1998). It is highly conserved among all form I RuBisCO RbcL homologs, not only of cyanobacteria (Figure 29B) but also of higher plants (Figure 30B, i). It is not found in archaeal and form II RuBisCO (Figure 30B, ii), along with the absence of RbcX in the respective organisms.

To obtain structural information on RbcL peptide binding, crystals of Syn7002-RbcX were soaked with the peptide EIKFEFD. One of the three crystallographically independent dimers in the native crystal form of RbcX exhibited difference electron density compatible with a bound peptide at the central recess. This difference density is not centered on the dyad axis of the RbcX dimer, indicating that one orientation is markedly preferred in the asymmetric environment of the crystal lattice. In the refined model, the peptide adopts an extended conformation with the side chains F462 and F464 extending into cavities lined by the side chains of T13, Y17, Y20, I50, and Q51 at the center of the RbcX cleft; the polar side chains of the peptide face outwards and are mostly disordered (Figures 30C and 30D). I460 of the peptide appears to contact Y20, R24, and S45 of one RbcX chain. In addition, the observed geometry is compatible with three hydrogen bonds formed between the peptide backbone and the side chains of Y20, R24, and Q51. The conformations of the two Q51 side chains appear to be mainly governed by van der Waals contacts to the phenyl rings of peptide residues F462 and F464 (Figure 30D).

5.3.8.1   Affinity of RbcX for RbcL peptide

Binding analysis by isothermal titration calorimetry (Figure 31) revealed a relatively low affinity of Syn7002-RbcX for peptide EIKFEFD (K_D ~230 µM), comprising the core binding region. A somewhat higher affinity (K_D ~160 µM) was measured with the longer peptide KEIKFEFDTVD, which includes the conserved residues Lys458 and Asp468. These measurements indicate that the interaction of Syn7002-RbcX with the RbcL C terminus is highly dynamic. Interestingly, substantially higher affinities for these peptides were measured with AnaCA-RbcX (K_D ~60 µM and 5 µM, respectively) (Figure 31). In AnaCA-RbcX, the loop region between helices α2 and α3 is shorter by one residue and contains a lysine at position 49, thus introducing a positive charge in the proximity of the binding groove (Figure 26). The positioning of this lysine may result in an increased affinity for the negatively charged C terminus of RbcL.
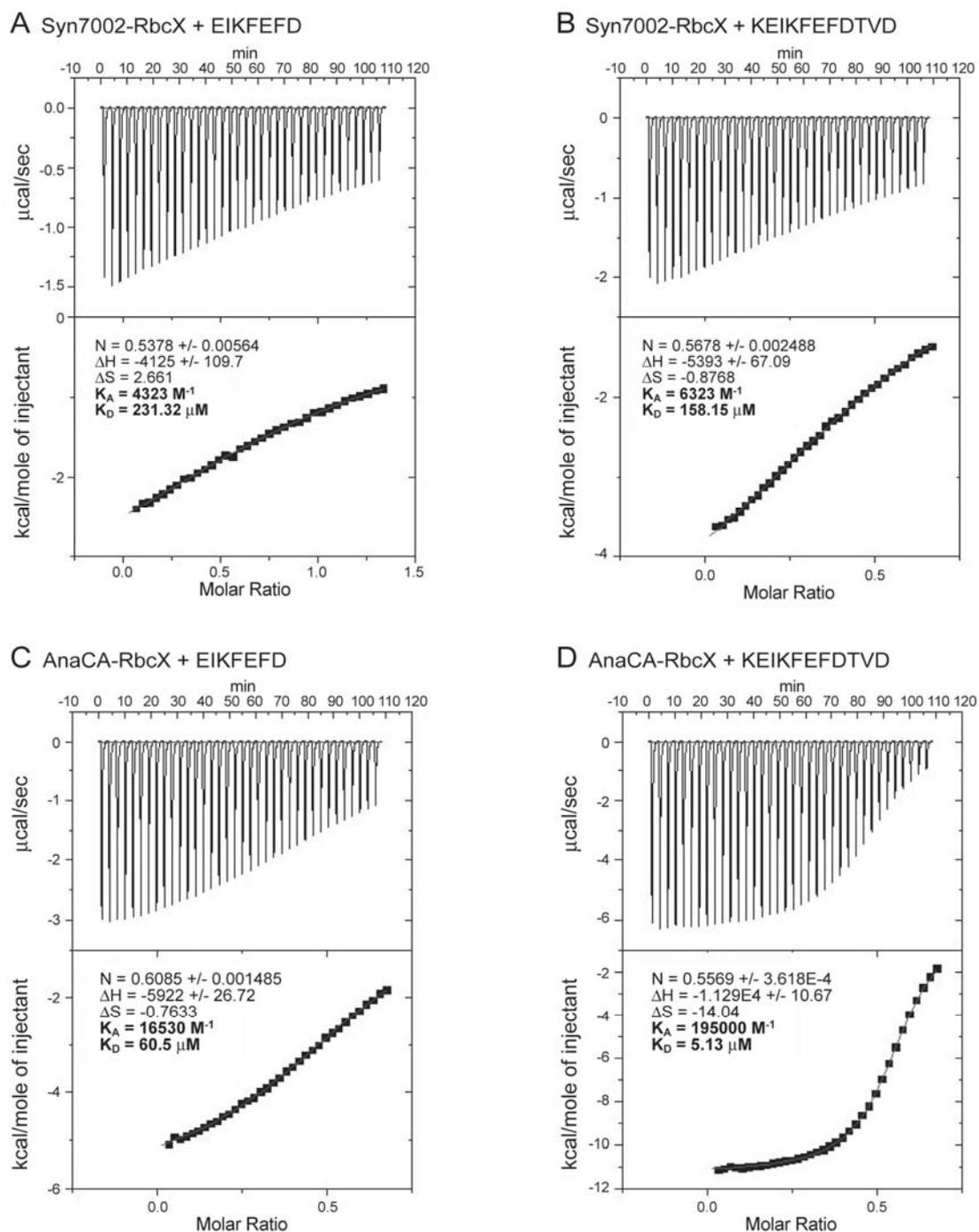
Figure 31. Isothermal titration calorimetry of the RbcX-peptide interaction
Titration of Syn7002-RbcX with Syn7002-RbcL peptide 1. EIKFEFD (A) or peptide 2. KEIKFEFDTVD (B). Titration of AnaCA-RbcX with Syn7002-RbcL peptide 1(C), or peptide 2 (D). The stoichiometry of binding is ~ 1 peptide per RbcX dimer (This experiment was performed by Sandra Saschenbrecker) (Saschenbrecker et al., 2007).

5.3.8.2   Importance of residues F462 and F464 of RbcL C-terminus

As mentioned above the RbcL peptide adopts an extended conformation with the side chains F462 and F464 extending into cavities at the center of the RbcX cleft, which might hint at the importance of these phenylalanines in the interaction of RbcL with RbcX.

So, to know the importance of phenylalanines (F462 and F464) of the EIKFEFD motif in RbcL$_8$ assembly, the oligopeptide in which these phenylalanines were replaced by alanines was used as a competitor for RbcL binding to RbcX. Translating *rbcL in vitro* in the presence of increasing amounts of oligopeptide EIK**F**E**F**D resulted in decreasing amounts of RbcL$_8$ made, proving that EIKFEFD competes with RbcL for RbcX binding (Figure 32A). However, addition of the mutant peptide EIK**A**E**A**D did not reduce the yield of RbcL$_8$ made (Figure 32A), suggesting the importance of the residues F462 and F464 in the RbcX assisted RbcL$_8$ assembly.

5.3.9   Dynamic nature of RbcX-RbcL interaction and its importance in holoenzyme assembly

The difference in affinities for the RbcL C-terminus observed with Syn7002-RbcX and AnaCA-RbcX (Figure 31) was utilized here to determine the functional relevance of the dynamic nature of this interaction. Surprisingly, AnaCA-RbcX, while supporting the expression of soluble RbcL, failed to promote the formation of RbcL$_8$ complexes competent for assembly with RbcS; instead, a high-molecular-weight form of RbcL was observed (Figure 20, Lane 5). Based on static light scattering, the mass of this complex was ~730 kDa, consistent with eight to nine RbcX dimers bound per RbcL$_8$ (Table 1).
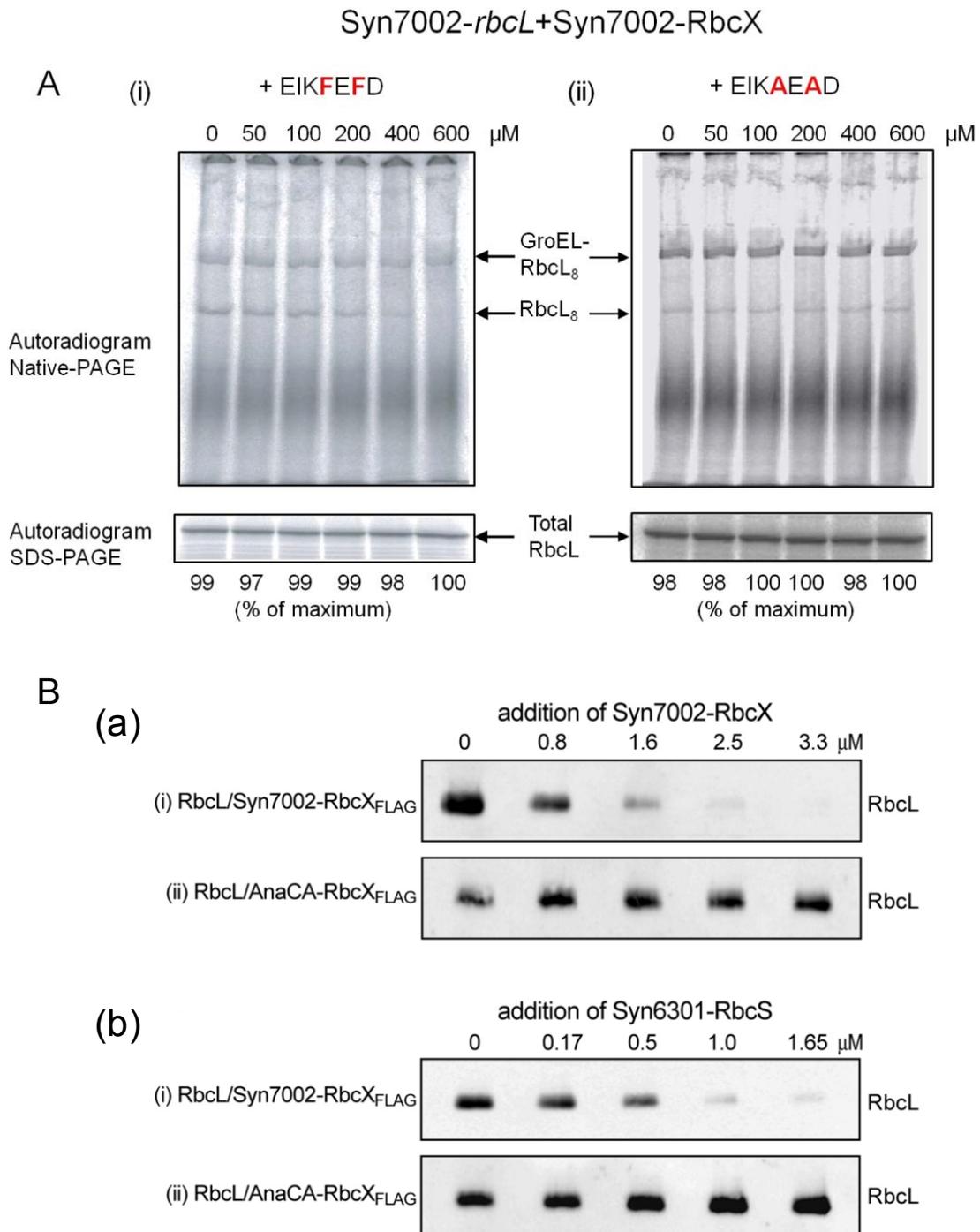
Figure 32. Effects of modulating RbcX binding strength on RbcL assembly and dynamic nature of RbcX-RbcL interaction

(A) Importance of F462 and F464. Syn7002-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine, 40μM Syn7002-RbcX and indicated amounts of oligopeptide, EIK**F**E**F**D

or EIK**A**EA**A**D. The products were analyzed by Native-PAGE and SDS-PAGE, followed by autoradiography.

(B). Dynamic nature of the $RbcL_8$-RbcX interaction

(a) Complexes of (i) Syn7002-RbcL/Syn7002-RbcX$_{FLAG}$ or (ii) Syn7002-RbcL/AnaCA-RbcX$_{FLAG}$ were produced by co-expression in *E. coli* and immunoprecipitated from soluble cell lysate with anti-FLAG affinity beads overnight at 4^0C. Bead-bound complexes were incubated with the indicated concentrations of non-tagged Syn7002-RbcX for 15 min at room temperature. Subsequently, bound protein was eluted from the beads and analyzed for RbcL by SDS-PAGE and immunoblotting.

(b) Complexes of (i) Syn7002-RbcL/Syn7002-RbcX$_{FLAG}$ or (ii) Syn7002-RbcL/AnaCA-RbcX$_{FLAG}$, produced as in (a), were incubated with indicated concentrations of Syn6301-RbcS and subjected to co-immunoprecipitation with anti-FLAG affinity beads for 15 min at room temperature. Bead-bound proteins were analyzed for RbcL by SDS-PAGE and immunoblotting (Experiments in B were performed by Sandra Saschenbrecker) (Saschenbrecker et al., 2007).

Next, the exchangeability of RbcX-bound $RbcL_8$ was analyzed by Sandra Saschenbrecker. RbcL was coexpressed with either FLAG-tagged Syn7002-RbcX or AnaCA-RbcX, and the respective complexes were immunoprecipitated (Figure 32B). Bound $RbcL_8$ was detected by immunoblotting after incubation with increasing concentrations of nontagged Syn7002-RbcX. Half-maximal displacement of RbcL was observed with ~1 µM RbcX, providing an estimate for the apparent affinity of Syn7002-RbcX for $RbcL_8$. Complete displacement of $RbcL_8$ by higher RbcX concentrations occurred within the 15 min incubation time. Thus, the peripheral surface of RbcX contributes to $RbcL_8$ binding over the lower affinity of the central peptide cleft alone (KD ~150 µM). Nevertheless, the $RbcL_8$-RbcX interaction appears to retain a substantial off rate, consistent with the dissociation of the complex upon gel filtration (Figure 29B).

In contrast, Syn7002-RbcX failed to displace $RbcL_8$ from the complex with AnaCA-RbcX$_{FLAG}$ (Figure 32B, a), presumably due to the 30-fold higher affinity of AnaCA-RbcX for the RbcL C-terminus. In an analogous experiment, the displacement reaction was performed with RbcS, which binds to free $RbcL_8$ with an affinity of ~10 nM (Andrews, 1988). $RbcL_8$ was readily displaced from Syn7002-RbcX by submicromolar RbcS concentrations (Figure 32B, b). In contrast, RbcS failed to displace $RbcL_8$ from the complex with AnaCA-RbcX (Figure 32B, b), explaining the nonproductive nature of the latter. These results support the conclusion that

complex formation between RbcL and RbcX must be dynamic in order to support the formation of RuBisCO holoenzyme.

Sample	Theoretical MW (Da)*	Measured MW (Da)	Hydrodynamic Radius (nm)
Syn6301-RbcL$_8$	419581.6	414900	5.0
Syn6301-RbcL$_8$S$_8$	526247.2	491600	5.2
Syn6301-RbcL$_8$S$_8$ /AnaCA-RbcX(His6)$_8$		730000	6.1
AnaCA-RbcX	1550606 (Monomer) 31013.2 (Dimer)	31040	1.8
Syn7002-RbcX	15269.1 (Monomer) 30538.2 (Dimer)	30170	1.8

Table 1. Mass analysis of recombinantly expressed proteins
Molecular mass and hydrodynamic radius of purified proteins analyzed by Multiangle light scattering (FFF-MALS).(*) Theoretical MW was analyzed using Expasy ProtoParam tool (http://www.expasy. org/tools/protparam.html) (FFF-MALS analysis was performed by Dr. Manajit Hayer Hartl).

5.3.10   RbcL$_8$-assembly stages and role of RbcX

RbcL$_8$ assembly is thought to involve the dimerization of folded monomers followed by tetramerization of dimers (Hubbs and Roy, 1993). At which stage does RbcX begin to interact with RbcL subunits in this reaction? To address this question, *rbcL* was translated *in vitro* in *E. coli* lysate supplemented with GroEL/GroES and RbcX$_{FLAG}$. Translation was first performed for 6 min in the absence of labeled amino acid to ensure linear RbcL production, followed by a 6 min pulse with ^{35}S-methionine and a chase with excess unlabeled methionine for up to 160 min. Aliquots of the reaction were withdrawn at different times, translation was stopped by keeping them on ice, and then analyzed by a discontinuous Bis-Tris Native-PAGE system in which protein migration correlates well with molecular mass. In the absence of RbcX, GroEL-bound RbcL (~800 kDa) was visible at the beginning of the chase, followed by the appearance of a fuzzy low-molecular weight band that migrated similarly to a monomeric mutant form of bacterial RuBisCO from *R. rubrum* (50 kDa; Figure 33A) and thus represented RbcL monomers. In addition, a

sharp RbcL band migrating similarly to dimeric RuBisCO from *R. rubrum* was observed. In the presence of excess Syn7002-RbcX, these forms of RbcL decreased over time, and formation of $RbcL_8$ occurred in ~40 minutes (Figure 33A and 33B).

A

Syn7002-*rbcL*+ Syn7002-RbcX_FLAG

B

C

**Syn7002-*rbcL*+
Syn7002-RbcX(Q29A)_FLAG**

D

**Syn7002-*rbcL*+ Syn7002-
RbcX(Y17A,Y20L)_FLAG**

E

Syn7002-*rbcL*

F

Empty plasmid
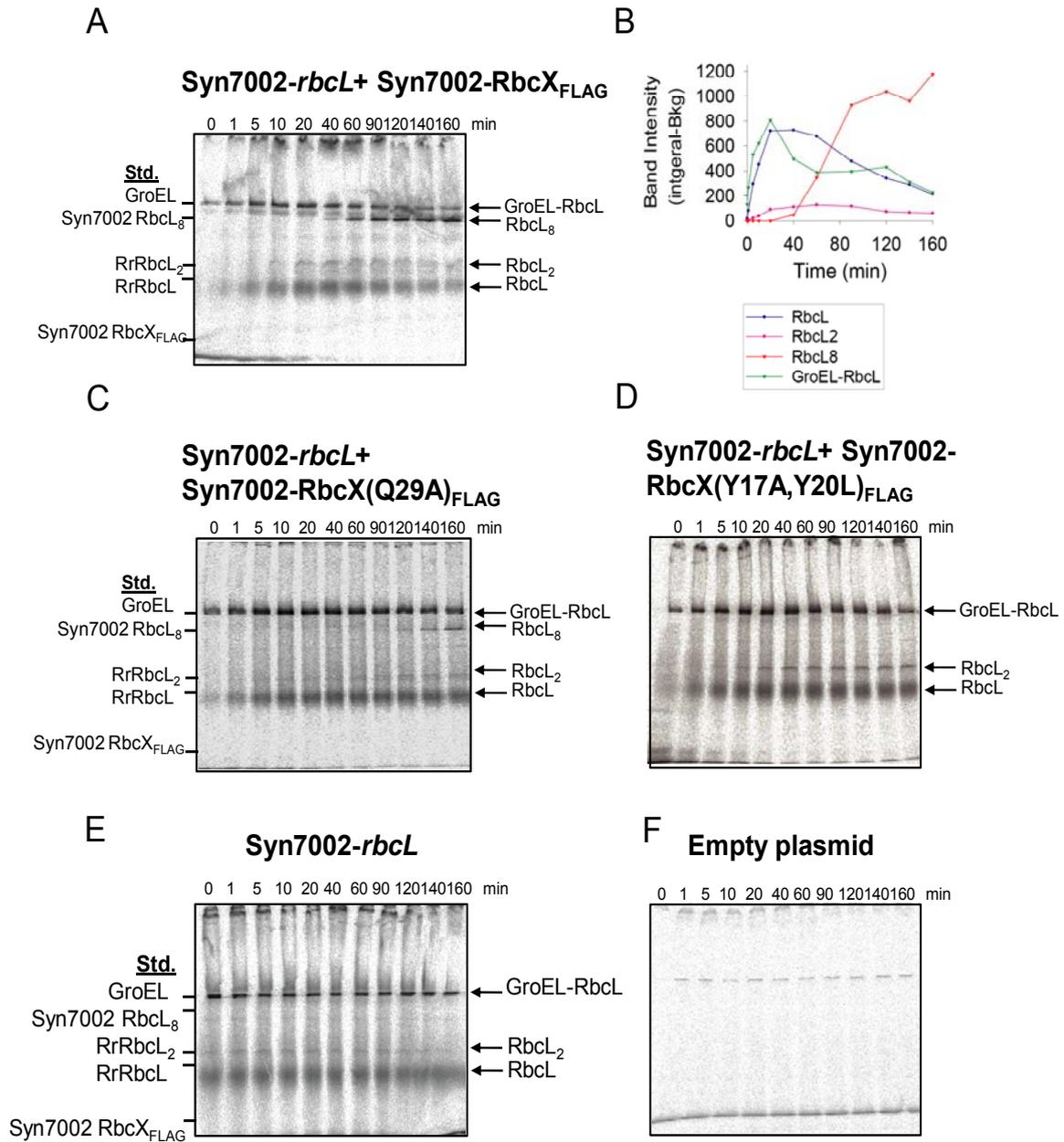
Figure 33. Pulse-chase analysis of RbcL$_8$ assembly

(A) Pulse-chase upon translation with *rbcL* was performed as described in the Materials and Methods section, in presence of FLAG tagged (A & B) Syn7002-RbcX wild type (WT), (C) Syn7002-RbcX (Q29A), (D) Syn7002-RbcX(Y17A,Y20L) or (E) in absence of RbcX or the translation of empty plasmid in absence of RbcX (F), The products were analyzed by Bis-Tris PAGE followed by autoradiography. (B) Quantification of bands corresponding to GroEL-RbcL, RbcL, RbcL$_2$ and RbcL$_8$ of Bis-Tris PAGE in (A).

Similarly, in the presence of Syn7002-RbcX(Q29A), bands corresponding to monomers and dimers appeared; but in contrast to the results with Syn7002-RbcX wild type, appearance of RbcL$_8$ was delayed (~120 minutes) (Figure 33C). In presence of Syn7002-RbcX (Y17A,Y20L), formation of RbcL$_8$ was almost abolished (Figure 33D), consistent with the observations that these mutations disrupt the functionality of RbcX (Figure 27). Wild type RbcX was highly efficient in keeping RbcL soluble and promoted formation of RbcL$_8$ complex, but RbcX (Y17A,Y20L) was highly inefficient in keeping RbcL soluble and so failed to support the formation of RbcL$_8$ complex (Figure 34A), again consistent with the earlier observations (Figure 27).

In absence of RbcX, formation of RbcL$_8$ was not observed upon translation of *rbcL*, though slightly reduced amounts of monomers and dimers could be seen (Figure 33E). None of the intermediates was observed when an empty plasmid was translated, suggesting that these bands are rbcL products (Figure 33F).

When the same pulse-chase experiment was performed with AnaCA-RbcX, two slower-migrating forms of RbcL (RbcL-I1 and RbcL-I2) in the range of 200–350 kDa were detected in equilibrium with RbcL monomer and dimer. RbcL monomer was more efficiently cleared from GroEL and more quantitatively consumed than in the assembly reaction with Syn7002-RbcX. The appearance of RbcL-I1 and RbcL-I2 was followed by the formation of high-molecular-weight RbcL$_8$-RbcX complex (~730 kDa) that migrated close to GroEL (Figure 34B). Notably, formation of RbcL-I1 and RbcL-I2 was not observed upon translation of *rbcL* (F462A,F464A) mutant protein (Figure 34C), indicating that recognition of the C-terminal motif initiates the interaction of RbcX with RbcL.
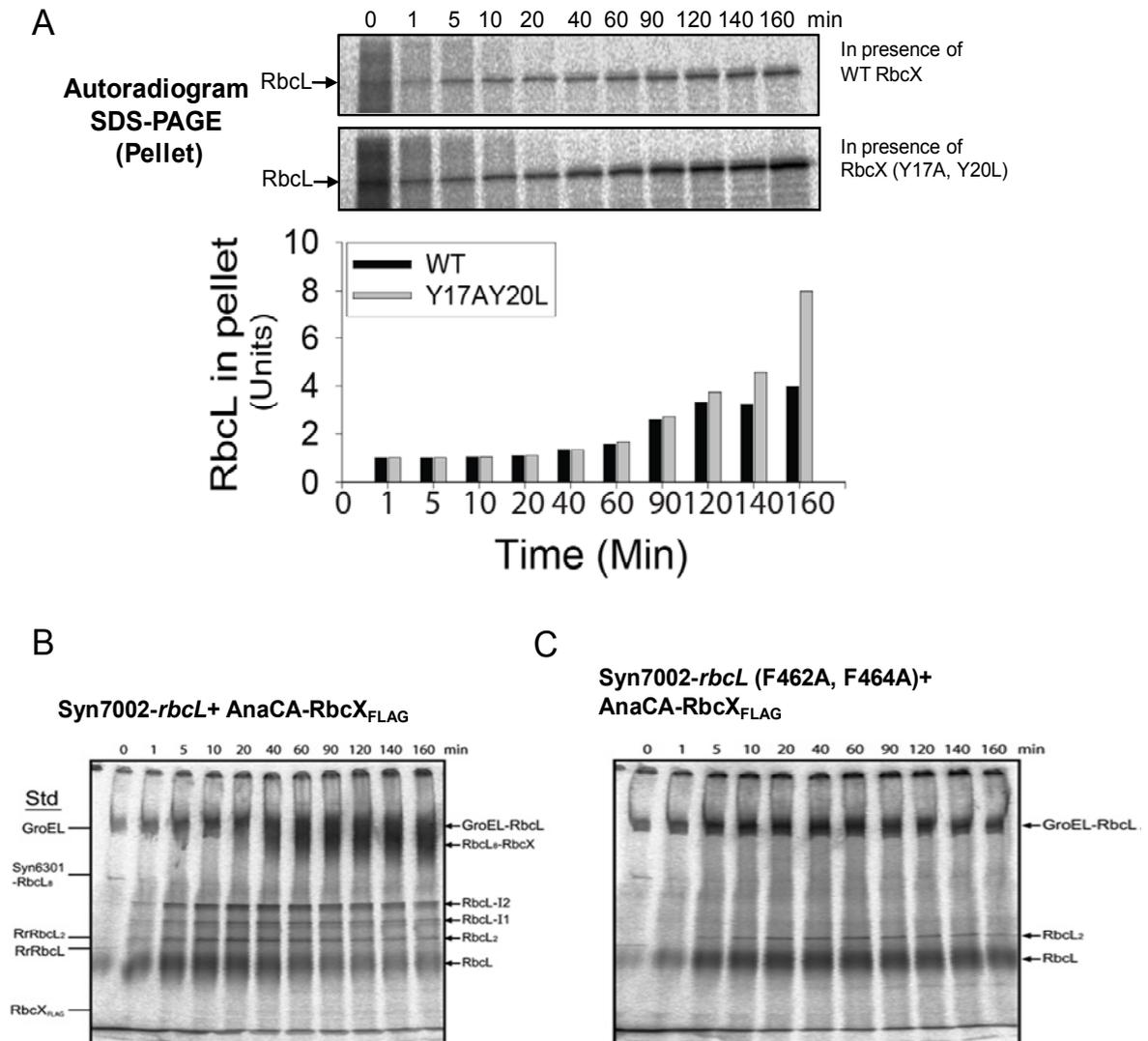
Figure 34. Pulse-chase analysis of RbcL$_8$ assembly

(A) Analysis of pellet fraction in the pulse chase translation of Syn7002-*rbcL* carried out in the presence of RbcX wild type (WT) or RbcX (Y17A,Y20L) by SDS-PAGE. Below is the quantification of bands. Intensity of band in 1 min product is set as 1.

(B) & (C) Pulse chase upon translation of Syn7002-*rbcL* or Syn7002-*rbcL*(F462A, F464A) in the presence of AnaCA-RbcX$_{FLAG}$ respectively.

5.3.10.1 RbcX interacts with early intermediates of RbcL$_8$ assembly

Based on the above experiments, RbcX appears to engage newly-synthesized RbcL subunits immediately after their chaperonin assisted folding at the level of monomers or dimers, shifting the equilibrium toward RbcX-bound intermediates and promoting the formation of RbcL$_8$ core complexes. In order to analyze this in detail, the following *in vitro* translation experiment was done. AnaCA-RbcX was used for the experiment instead of Syn7002-RbcL, as the former has higher affinity to RbcL than the latter.

	1	2	3	4	5	6	7	8	9	10	11
^{35}S –methionine	+						+	+	+	+	
Unlabelled methionine		+									+
Syn7002-*rbcL*		+					+	+	+		+
AnaCA-*rbcX*	+										
Empty vector										+	
GroEL/GroES		+					+	+	+		+
Syn7002-RbcX							+				
AnaCA-RbcX								+			
CAM	+	+					+	+	+	+	+
Add ^{35}S- methionine											+
Mix				1 + 2							
Incubate at 30^{0}C, for Min			10	20	30	40					30

Autoradiogram of Native-PAGE

Std.
GroEL —
Syn7002-RbcL$_8$ —
Rr-RbcL$_2$ —
Rr-RbcL (K168E) —
AnaCA-RbcX —

→ RbcL+ RbcX
→ GroEL-RbcL
← RbcL$_8$
← RbcL monomer + RbcX
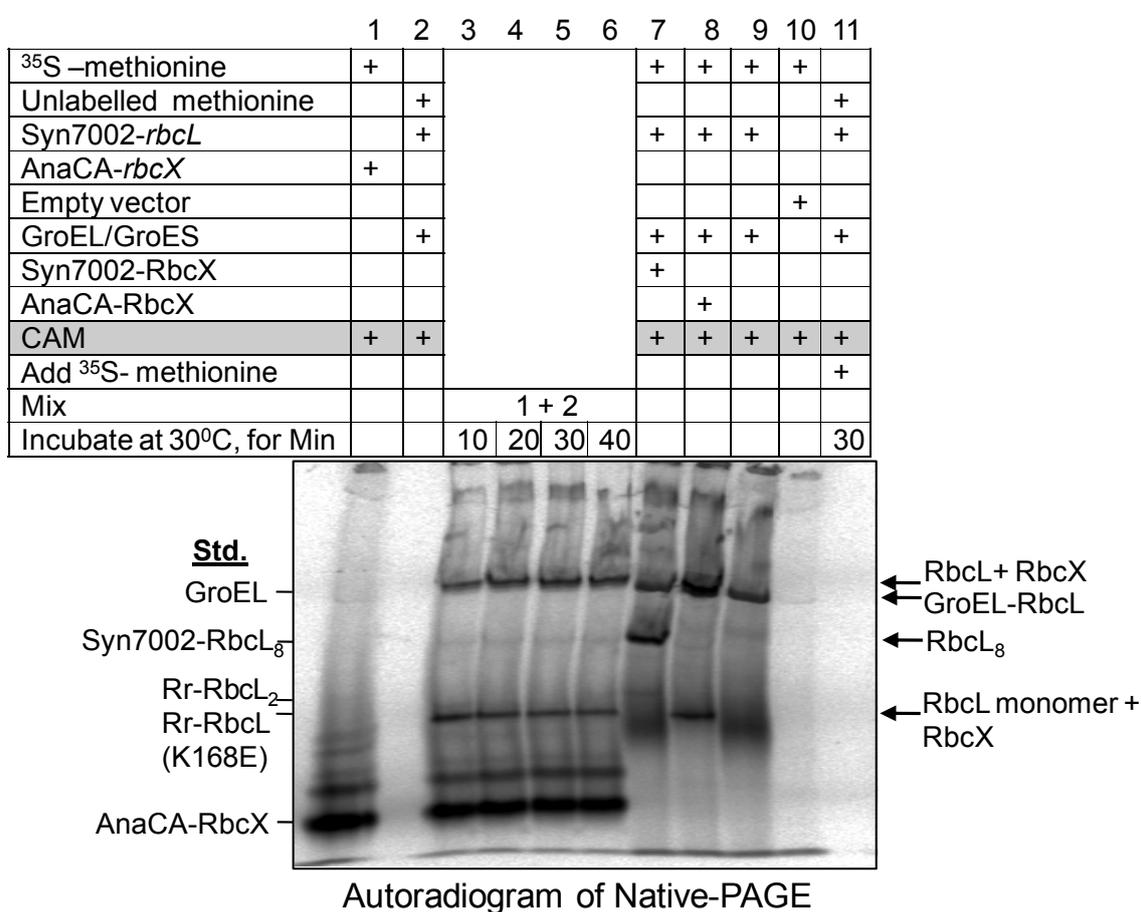
Figure 35. RbcX interacts with early stages of RbcL$_8$ assembly
(Lane 1) AnaCA-*rbcX* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine, (Lane 2) Syn7002-*rbcL* was translated in the presence of unlabelled methionine and GroEL/GroES (0.5μM/1.0μM) at 30^{0}C for 90 minutes, translation was inhibited by the addition of chloramphenicol,

(Lanes 3-6) the products of 1 and 2 were mixed and incubated for time durations mentioned. (Lane 7) Syn7002-*rbcL* was translated in the presence of ^{35}S-methionine, GroEL/GroES (0.5µM/1.0µM) and Syn7002-RbcX or (Lane 8) AnaCA-RbcX or (Lane 9) in absence of RbcX. Translation was inhibited by the addition of CAM. The products were analyzed by Native-PAGE followed by autoradiography. (Lane 10 & 11) *rbcL* was translated in the presence of cold methionine. CAM treatment was followed by ^{35}S-methionine addition. The product was incubated for 30 minutes at 30^0C.

In the first reaction, AnaCA-*rbcX* was translated in the presence of ^{35}S-methionine (Figure 35, Lane 1) and in the second reaction Syn7002-*rbcL* was translated in the presence of cold methionine and GroEL/GroES (Figure 35, Lane 2), followed by inhibiting the translation by the addition of chloramphenicol. The products were mixed and incubated to allow RbcL and RbcX interaction (Figure 35, Lanes 3-6). Since RbcL is non-radiolabelled, radiolabelled products comigrating with *R. rubrum* monomer represented AnaCA-RbcX bound to RbcL monomer (Figure 35, Lanes 3-6), supporting the conclusion that the interaction with RbcX initiate at the level of RbcL monomers. As expected, translation of Syn7002-*rbcL* in the presence of AnaCA-RbcX and ^{35}S-methionine produced larger sized complex and a band corresponding to monomer of RbcL (Figure 35, Lane 8). Translating *rbcL* in presence of cold methionine followed by the addition of ^{35}S-methionine after CAM treatment did not produce any radiolabelled product, suggesting that inhibition of translation by CAM was efficient (Figure 35, Lanes 10 and 11).

5.3.10.2 Mass-spectrometric analysis of RbcL-RbcX interaction

To further confirm the interaction of RbcX with RbcL monomers during RbcL$_8$ assembly, mass-spectrometry based protein identification (LC/MSMS) was performed (Figure 36 and 37). For this purpose, Syn7002-*rbcL* was translated *in vitro* in *E. coli* lysate supplemented with GroEL/GroES (0.5µM/1.0µM). Translation was inhibited by the addition of CAM, an aliquot was collected and RbcX protein was added to the remaining sample, followed by incubation at 30^0C. Aliquots were collected at different time intervals and were analysed by Bis-Tris PAGE (Figure 36). Bands corresponding to GroEL-RbcL (Figure 36, Box 1), RbcL monomers-

0 1 2 4 6 10 15 20 25 30 45 50 min
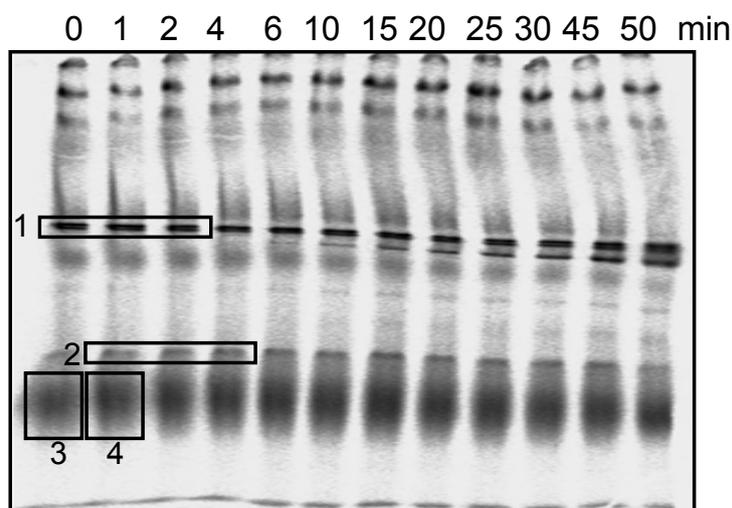
Autoradiogram of Bis-Tris PAGE

LC /MSMS based identification of proteins

1. GroEL +7002 LSU
2. 7002 RbcL
3. Syn7002 RbcL
4. Syn7002 RbcL + Syn7002 RbcX

Figure 36. Mass-spectrometry analysis of RbcX interaction with early stages of RbcL$_8$ assembly

Protocol for *in vitro* translation and preparation of samples for LC/MSMS analysis are mentioned in the materials and methods section. Syn7002-*rbcL* was translated *in vitro* in *E. coli* lysate supplemented with GroEL/GroES (0.5µM/1.0µM). Translation was inhibited by the addition of CAM and aliquot was collected and RbcX protein was added to the remaining product and was incubated further at 30^0C. Aliquots were collected at different time intervals and were analyzed by Bis-Tris PAGE. The above indicated gel slices from Bis-Tris PAGE were further analysed by LC/MSMS. Below is the list of proteins identified and the number corresponds to the gel slice number (Mass-spectrometric analysis was performed in cooperation with Dr. Frank Seidler, Department of Membrane Biochemistry, MPI, Martinsried).

before (Figure 36, Box 3) and after (Figure 36, Box 4) the addition of RbcX and RbcL$_2$ (Figure 36, Box 2) were analysed by LC/MSMS (Figure 37A and 37B). This analysis revealed the presence of RbcX along with RbcL monomers in the aliquot drawn after the addition of RbcX protein showing a peptide coverage of ~90% for RbcX (Figure 36, Box 4 and Figure 37B), but not in the aliquot collected before the

addition of RbcX protein (Figure 36, Box 3 and Figure 37A), thereby confirming the interaction of RbcX with RbcL monomer.

Based on these experiments, RbcX appears to interact with newly-synthesized RbcL monomers immediately after their folding by chaperonin, promoting them to form RbcL$_8$ complex.

A

Significant hits:
E.coli GroES (AAL56005)
Score: 57

Synechococcus *sp.* PCC 7002 RbcL (Q44176)
Score: 35

B (i)

Significant hits:
Synechococcus sp. PCC 7002 RbcX (Q44177)
Score: 176

Synechococcus *sp.* PCC 7002 RbcL (Q44176)
Score: 22

(ii)
```
  1  MEFKKVAKET AITLQSYLTY QAVRLISQQL SETNPGQAIW LGEFSKRHPI
 51  QESDLYLEAM MLENKELVLR ILTVRENLAE GVLEFLPEMV LSQIKQSNGN
101  HRRSLLERLT QVDSSSTDQT EPNPGESDTS EDSE
```

(iii)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
1	301.233000	600.451448	600.395859	0.055589	0	24	0.00086	1	R.ILTVR.E
3	309.216000	616.417448	616.354385	0.063063	0	5	0.035	1	R.SLLER.L
4	315.183000	628.351448	628.390762	-0.039314	0	25	0.00034	1	K.ELVLR.I
228	929.048000	1856.081448	1855.967926	0.113522	0	46	1.2e-006	1	K.ETAITLQSYLTYQAVR.L
63	619.920000	1856.738172	1855.967926	0.770246	0	(9)	0.028	1	K.ETAITLQSYLTYQAVR.L
229	929.611000	1857.207448	1855.967926	1.239522	0	(44)	6.1e-006	1	K.ETAITLQSYLTYQAVR.L
230	929.672000	1857.329448	1855.967926	1.361522	0	(36)	3.7e-005	1	K.ETAITLQSYLTYQAVR.L
64	620.158000	1857.452172	1855.967926	1.484246	0	(11)	0.017	1	K.ETAITLQSYLTYQAVR.L
144	726.518000	2176.532172	2176.017990	0.514182	0	10	0.0054	1	R.HPIQESDLYLEAMMLENK.E + Oxidation (M)
146	731.222000	2190.644172	2192.012894	-1.368722	0	(8)	0.031	1	R.HPIQESDLYLEAMMLENK.E + 2 Oxidation (M)
246	1138.730000	2275.445448	2274.181625	1.263823	0	28	0.00016	1	R.ENLAEGVLEFLPEMVLSQIK.Q + Oxidation (M)
195	816.259000	2445.755172	2445.253906	0.501266	0	25	0.00016	1	R.LISQQLSETNPGQAIWLGEFSK.R
196	816.442000	2446.304172	2445.253906	1.050266	0	(18)	0.00088	1	R.LISQQLSETNPGQAIWLGEFSK.R
197	816.524000	2446.550172	2445.253906	1.296266	0	(20)	0.00046	1	R.LISQQLSETNPGQAIWLGEFSK.R
220	918.818000	2753.432172	2754.111115	-0.678943	0	(1)	0.079	1	R.LTQVDSSSTDQTEPNPGESDTSEDSE.-
221	919.209000	2754.605172	2754.111115	0.494057	0	14	0.006	1	R.LTQVDSSSTDQTEPNPGESDTSEDSE.-

Figure 37. LC/MSMS based protein identification
(A) Probability Based Mowse Score of matching proteins as significant hits for the proteins from the gel slice 3.
(B) (i) Probability Based Mowse Score of matching proteins as significant hits for the proteins from the gel slice 4. (ii) The peptide coverage identified by LC/MSMS for Syn7002-RbcX. Matched peptides shown in red. (iii) Table summarizing the details including peptides identified for Syn7002-RbcX.Ions score is -10*Log (P), where P is the probability that the observed match is a random event. Individual ions scores > 2 indicate identity or extensive homology (p<0.05). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

5.4   Characterization of RbcX from *Arabidopsis thaliana*

The purification of At-RbcX and At-RbcS1A and structural analysis of At-RbcX were done in close collaboration with Bharathi Vasudeva Rao.

In *Arabidopsis thaliana,* RuBisCO large subunits are encoded by *rbcL* gene in plastome (Figure 38). *A. thaliana* possess four different small subunits (RbcS1A, RbcS1B, RbcS2B and RbcS3B) (Figure 39A). Like other higher plants, small subunits of *A. thaliana* are encoded by the *rbcS* genes in the nuclear genome. They contain N-terminal transit peptides (Figure 39A) and after translation on cytosolic ribosomes, the small subunit precursor proteins have to be imported into the chloroplasts, where they are processed and folded to the native state (Gatenby and Ellis, 1990). RuBisCO assembly takes place in the chloroplast. In cyanobacteria, *rbcL* and *rbcS* genes exist on the same operon.

The large subunits of cyanobacteria and higher plants share about 80% sequence identity whereas small subunits share only 40% (Newman and Gutteridge, 1993) (Figure 38 & 39). The most striking difference in the small subunits is found in the βA-βB loop; cyanobacteria have only 10 residues in the βA-βB loop, whereas plants have 22 residues (Spreitzer, 2003) (Figure 39).

The RbcX recognition sequence on the RbcL C-terminus, detected in peptide screens, is conserved in higher plant RbcL (Figure 38) and homologs to RbcX genes are present in plant genomes, exhibiting around 30% similarity to the

cyanobacterial RbcX sequences (Figure 40). It is thus likely that RbcX in plants has an important role in RuBisCO assembly similar to cyanobacterial RbcX.
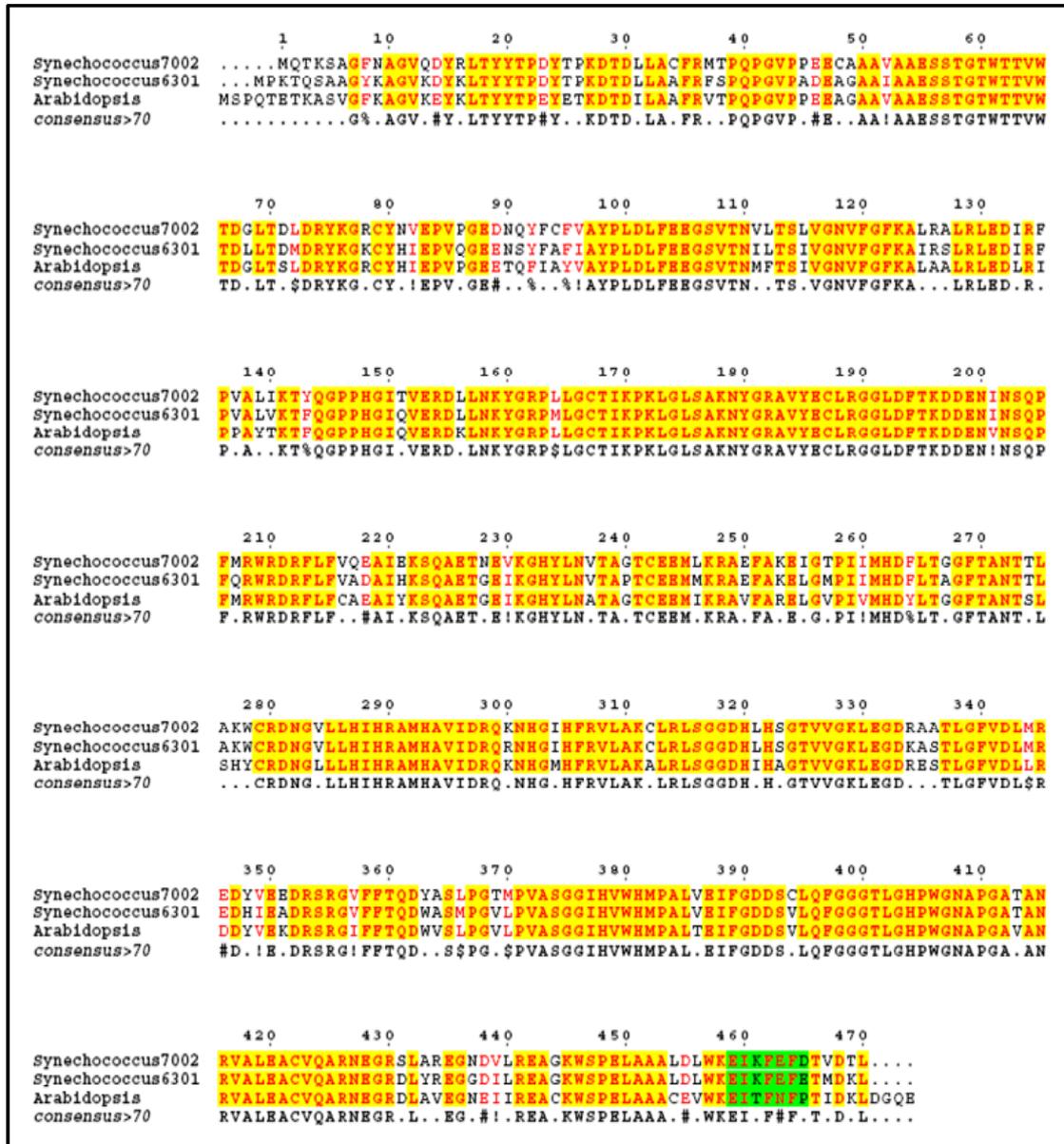
Figure 38. Conservation of RbcX binding motif on the C-terminus of RbcL in *A. thaliana*.

Alignment using MultAlin analysis program (Corpet, 1988) of representative large subunits of RuBisCO from *Synechococcus* sp. PCC7002 (Q44176), *Synechococcus* sp. PCC6301 (P00880) and *Arabidopsis thaliana* (BAA84393). Residues indicated in yellow background are invariable and residues indicated in red are conserved in more than 50% of the sequences. Green box shows the motif to which RbcX binds.

A
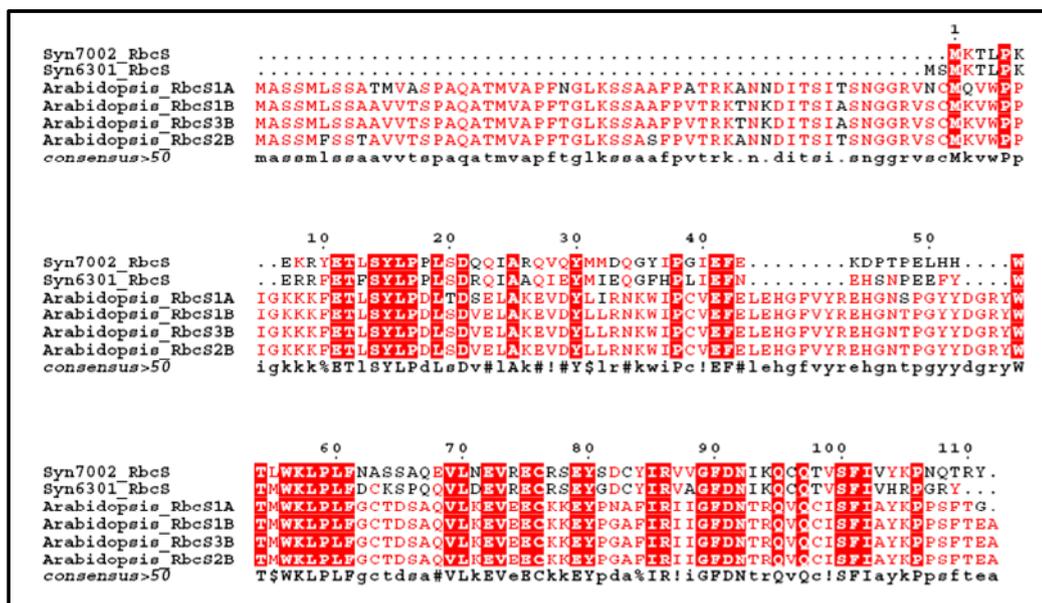
B
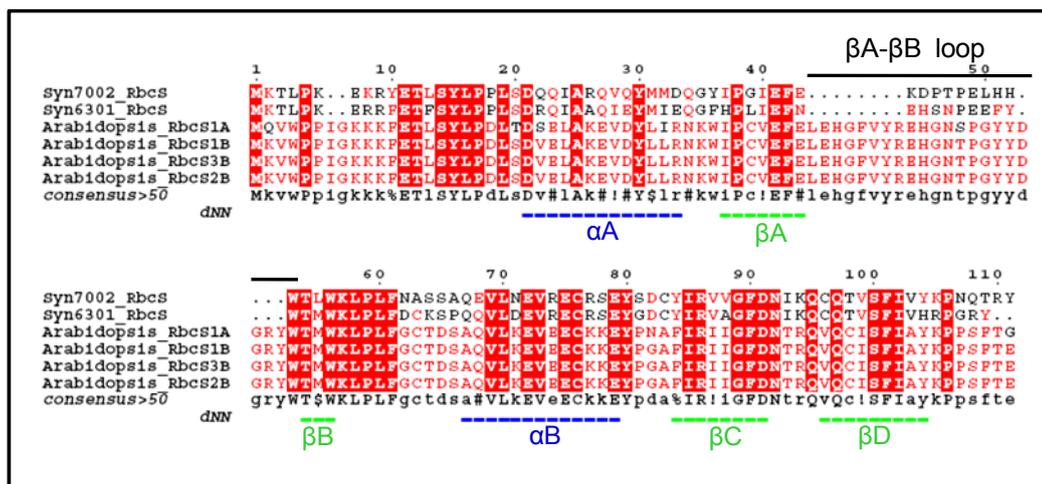
Figure 39. Divergence in the βA-βB loop of RbcS

Alignment using MultAlin analysis program (Corpet, 1988) of representative small subunit primary structures of RuBisCO from *Synechococcus* sp. PCC7002 (Q44178), *Synechococcus* sp. PCC 6301 (P04716), RbcS1A (NP_176880), RbcS1B (NP_198657), RbcS2B (NP_198658), RbcS3B (NP_198657) of *Arabidopsis thaliana* with (A) and without (B) transit peptides. Residues indicated in red background are invariable and residues indicated in red are conserved in more than 50% of the sequences.

A

B
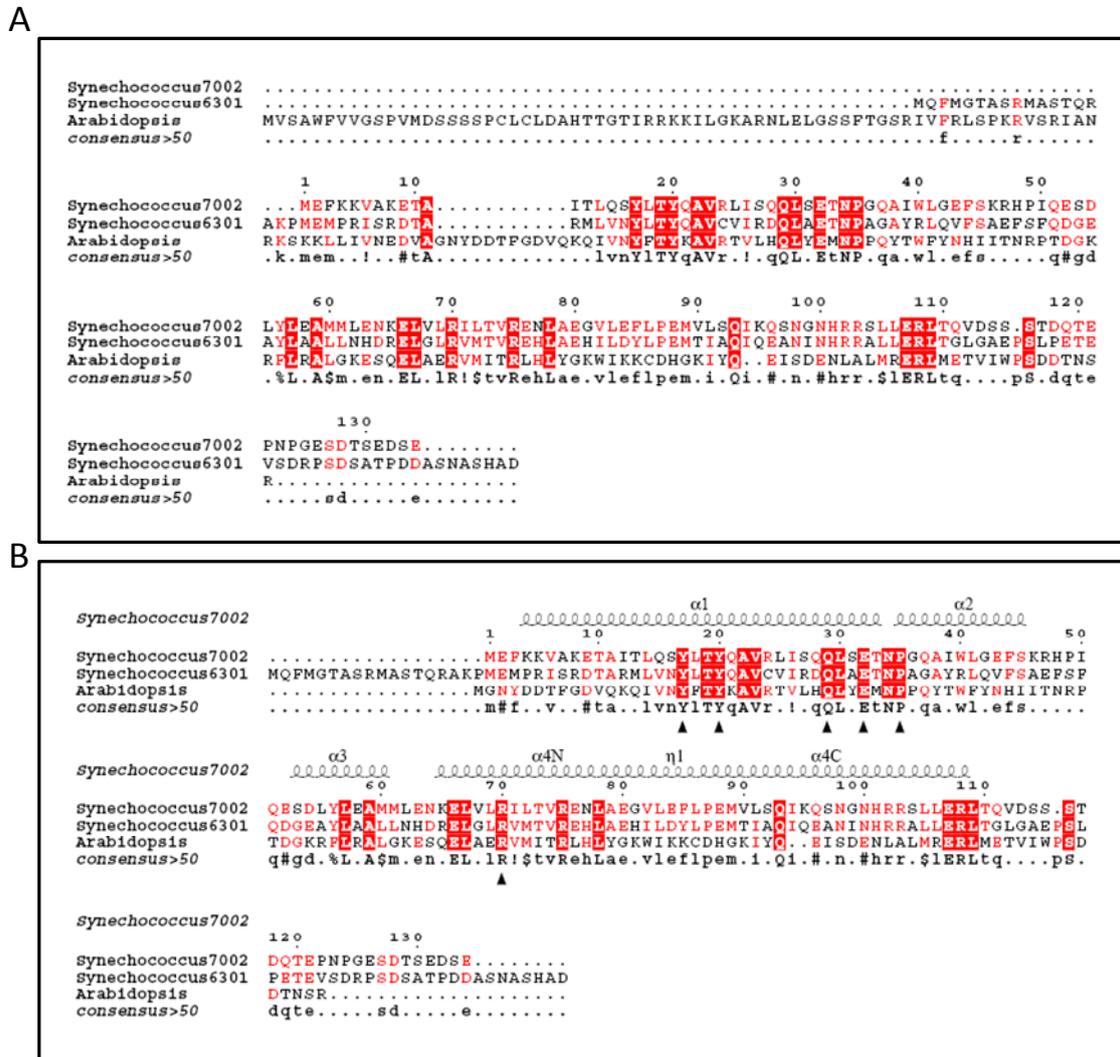
Figure 40. Conservation of crucial residues in *A. thaliana* RbcX

Alignment of amino acid sequences of RbcX homologs from *Synechococcus* sp. PCC7002 PR-6 (Q44177), *Synechococcus elongatus* PCC6301 (Q5MZ09) and *Arabidopsis thaliana*_gene I (NP_568382) (with (A) and (B) without transit peptide) using MultAlin analysis program (Corpet, 1988). Residues indicated in red background are invariable and residues indicated in red are conserved in more than 50% of the sequences; a consensus sequences according to these criteria using the standard symbols of the MultAlin analysis program (Corpet, 1988) is given in the bottom row of the alignment. The secondary structure elements and the residue numbering for RbcX of *Synechococcus* sp. PCC7002 are indicated above the sequences. Residues which are necessary for the proper function of Syn7002-RbcX are indicated by black triangles.

5.4.1 Size determination of *A. thaliana* RbcX

A. thaliana rbcX (At-*rbcX*) was cloned, without the transit peptide sequence, into *pHUE* vector (Catanzariti et al., 2004) which provides a cleavable ubiquitin tag to increase the solubility of RbcX (Figure 41A). The sequence for the transit peptide was predicted using ChloroP program (Emanuelsson et al., 1999) and further confirmed by aligning the amino acid sequence to that of Syn7002-RbcX (Figure 40). RbcX was expressed in *E. coli* and purified as soluble, homogeneous protein, running at an apparent molecular weight of approximately 15 kDa on SDS-PAGE (Figure 41B). To determine its oligomeric state, the protein was subjected to field flow fractionation in conjunction with multiangle light scattering (FFF-MALS) (Figure 42A) and size exclusion chromatography using the gel filtration column Superdex 200 (Figure 42B). Both these methods revealed that At-RbcX forms dimers of ~30 kDa similar to Syn7002-RbcX.
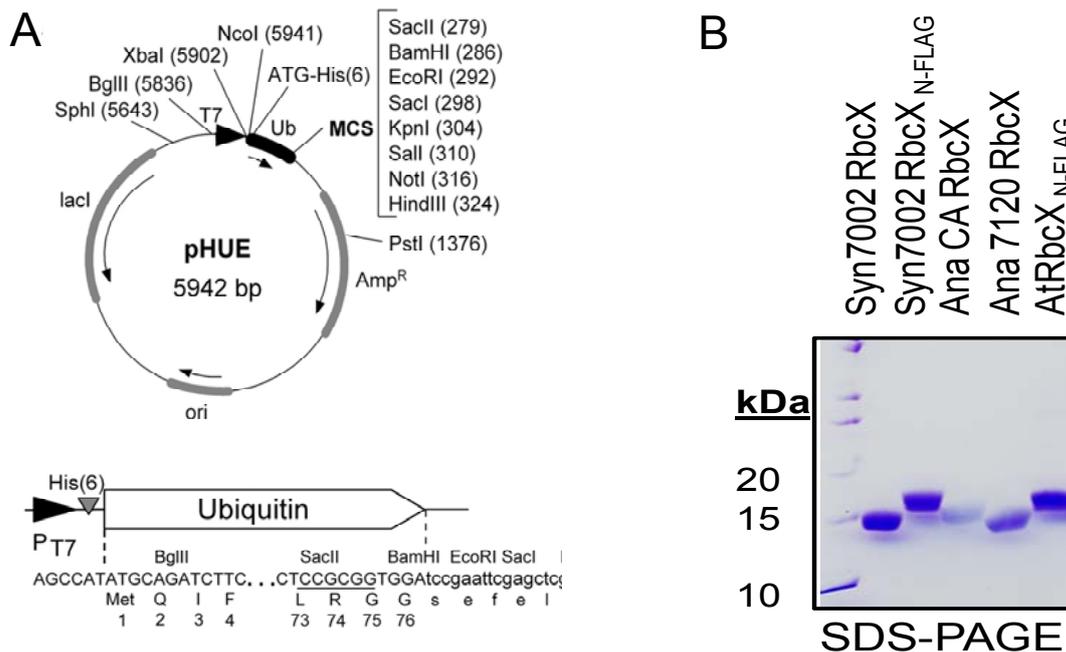
Figure 41. Cloning and purification of RbcX from *A. thaliana*
(A) At-*rbcX* was cloned into *pHUE* vector (Catanzariti et al., 2004) which provides a cleavable Ubiquitin tag and was expressed in *E.coli* and purified as mentioned in Materials and Methods section (Cloning and purification was done in collaboration with Bharathi Vasudeva Rao)
(B) SDS-PAGE followed by comassie staining of purified RbcX proteins.
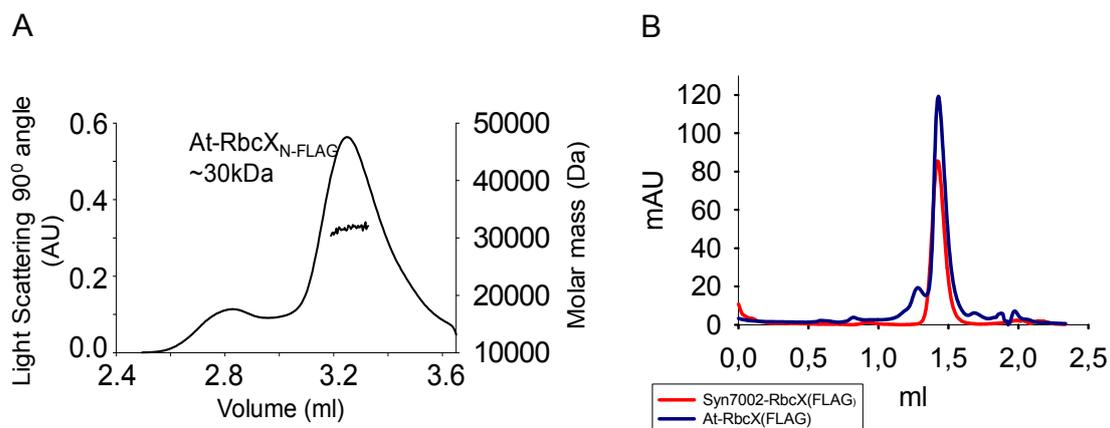
Figure 42 Size determination of At-RbcX using light scattering and gel filtration analysis

(A) FFF-MALS data showing the analysis of recombinantly purified At-RbcX$_{FLAG}$. Horizontal line across the peak indicates molar mass and homogeneity of the sample (FFF-MALS was performed in cooperation with Dr. Manajit Hayer-Hartl).
(B) Gel filtration analysis of Syn7002-RbcX$_{FLAG}$ and At-RbcX$_{FLAG.}$

5.4.2 RbcX from *Synechococcus* sp. PCC7002 and *A. thaliana* possess similar secondary structure

Syn7002-RbcX and At-RbcX share sequence identity of only around 32%. However the crucial residues mentioned earlier in the study of Syn7002-RbcX are conserved in At-RbcX as well (Figure 40).

The secondary structure analysis using circular dichroism demonstrated that At-RbcX is mainly alpha helical, similar to Syn7002-RbcX (Figure 43). The secondary structure predictions of both RbcX proteins, with the server, http://bioinformatik.biochemtech.uni-halle.de/cdnn, indicated similar results and predicted about 73% alpha helix for Syn7002-RbcX (Figure 43B) (which is already known from our studies) and about 65% for At-RbcX (Figure 43C). Moreover, the secondary structure elements that are predicted to exist for Syn7002-RbcX and At-RbcX are very similar in both their extent and position within their respective primary amino acid sequences.
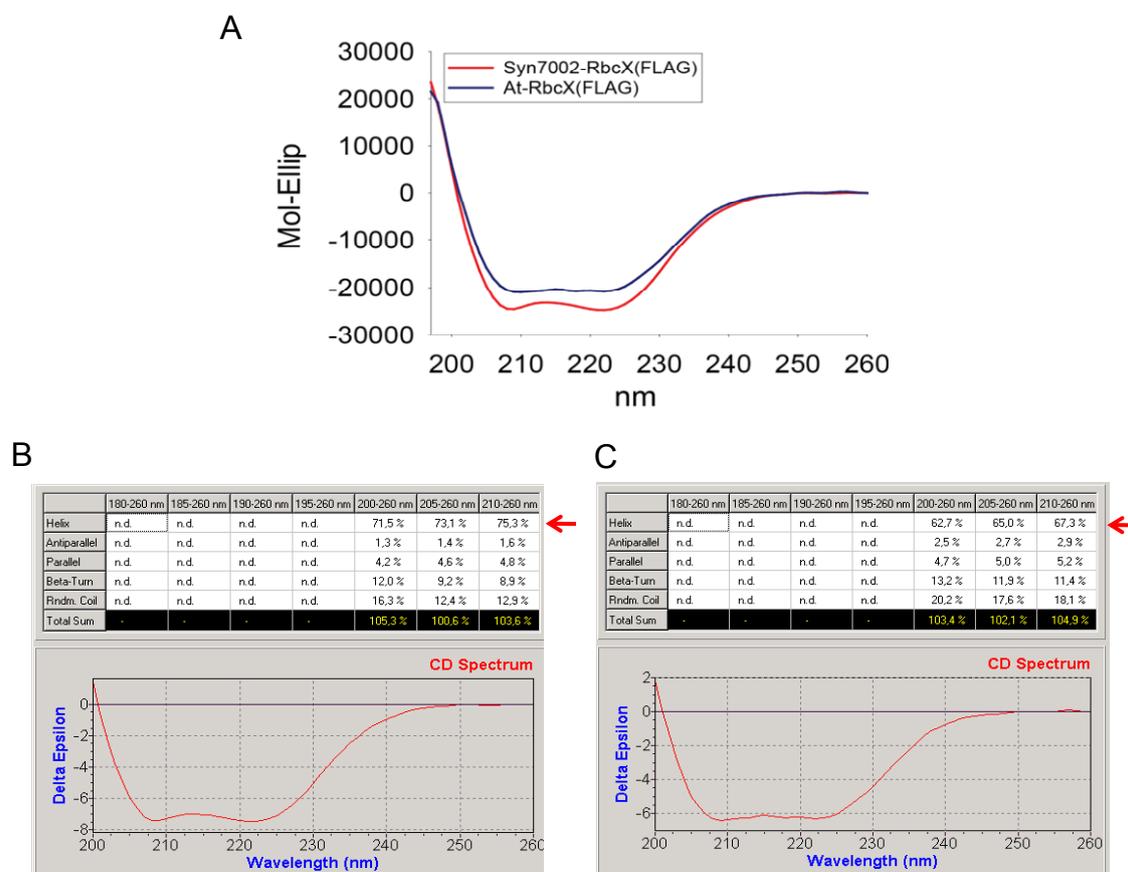
Figure 43. Secondary structure analysis of RbcX from *A. thaliana*
(A) CD analysis of recombinantly purified At-RbcX$_{FLAG}$ and Syn7002-RbcX$_{FLAG}$ proteins
(B) Secondary structure prediction of Syn7002-RbcX and At-RbcX proteins using cdnn server http://bioinformatik.biochemtech.uni-halle.de/cdnn.

5.4.3 Tertiary structure of At-RbcX

A tertiary structure prediction for At-RbcX was performed by sequence alignment based superimposition, using the X-ray structure of *Synechococcus* sp. PCC7002 (Figure 44). Polar residues such as Q29, E32, R70 located at the corner of the RbcX dimer, which are known to be important for the function of RbcX, are conserved in *A. thaliana* RbcX (Figure 44, 40 & 27A). Y17, Y20, and I50 of each monomer, lining the groove in the center of the molecule, are also conserved (Figure 44 & 25B).
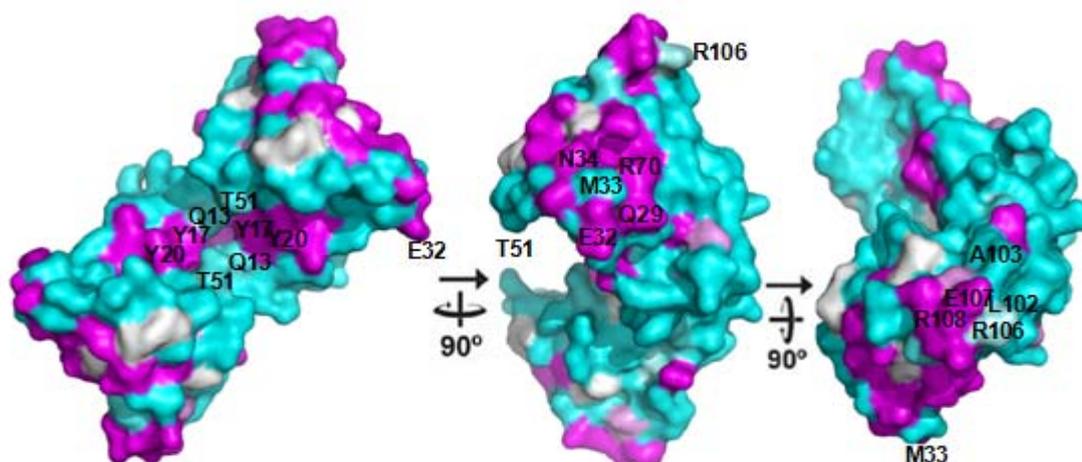
Figure 44. Superimposed model for *A. thaliana* RbcX

Surface conservation in At-RbcX. The similarity score from an alignment of amino acid sequences of Syn7002-RbcX and At-RbcX was plotted onto the accessible surface of the RbcX dimer. Sequence conservation is indicated by a color gradient, indicating highly conserved residues in magenta and variable regions in cyan. The positions of conserved surface residues are indicated (This work was performed in cooperation with Dr. Andreas Bracher).

5.4.4 *A. thaliana* RbcS1A (At-RbcS1A) can complement for Syn7002-RbcS

As mentioned above, the RuBisCO small subunits of higher plants share about 40% sequence identity (Newman and Gutteridge, 1993) (Figure 39) and they mainly differ in their βA-βB loop. In order to analyze the function of *A. thaliana* RuBisCO small subunit At-RbcS1A, the *rbcS1A* gene, without the transit peptide sequence, was cloned into pET vector, expressed in *E. coli* and purified from the inclusion bodies. The sequence for the transit peptide was predicted using ChloroP program (Emanuelsson et al., 1999) and further confirmed by aligning the amino acid sequence to that of Syn7002-RbcS (Figure 39). The purified At-RbcS1A was able to replace Syn7002-RbcS and yielded active hybrid holoenzyme in the *E.coli* based translation system *in vitro* (Figure 45A). Translating Syn7002-*rbcL* in the presence of purified Syn7002-RbcX, GroEL/ES and At-RbcS1A, replacing Syn7002-RbcS, yielded active RuBisCO in a concentration dependent manner. Similar results were also obtained when Syn7002-*rbcLX* was coexpressed with GroEL/ES in *E.coli in vivo* and purified At-RbcS1A was supplemented to the *E.coli* lysate (Figure 45B).
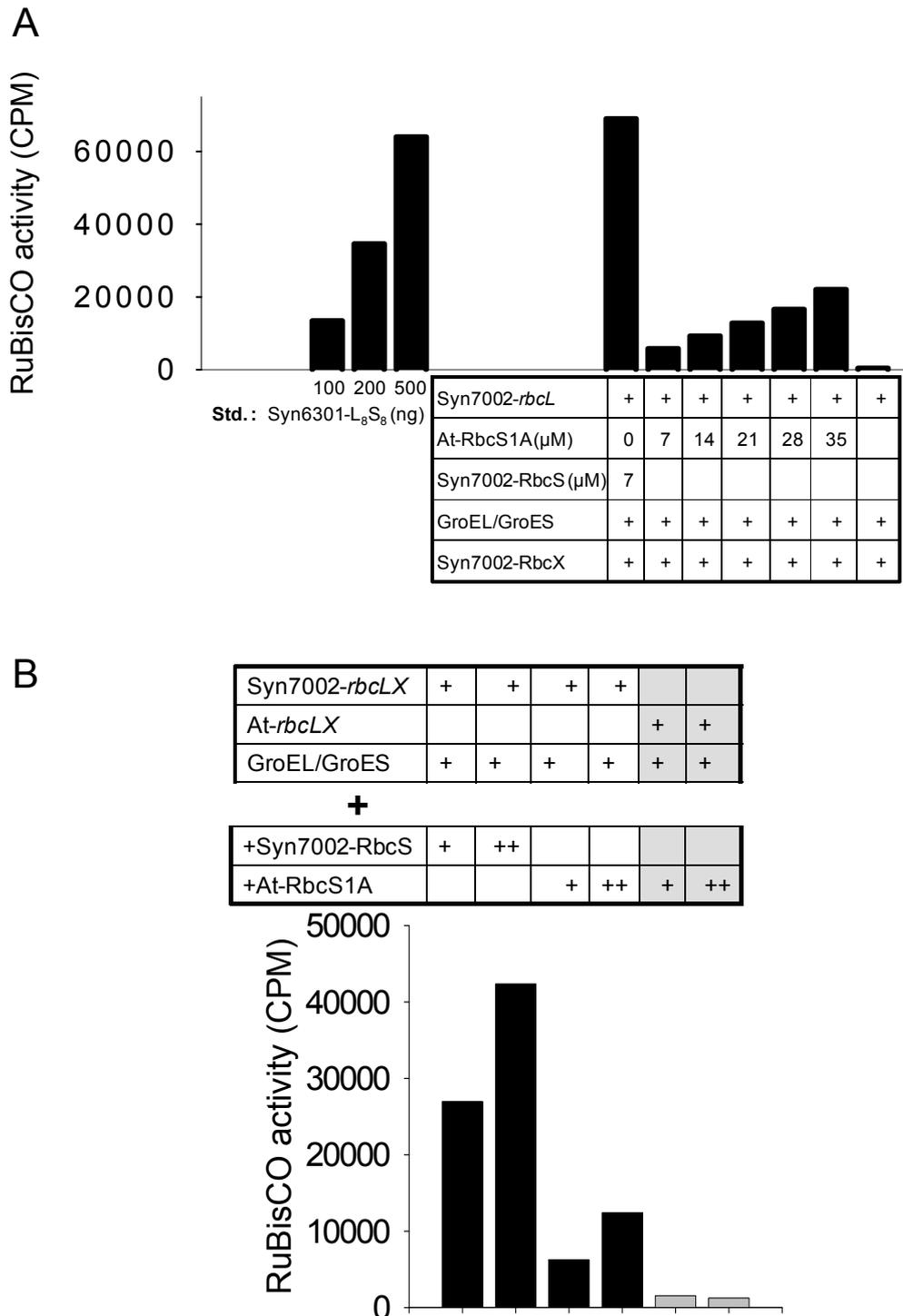
Figure 45. Functional analysis of At-RbcS1A and At-RbcX

(A) Syn7002-*rbcL* was translated in *E.coli* lysate *in vitro* in presence of unlabelled methionine. Lysates were supplemented with purified GroEL/GroES (0.5μM/1.0μM), Syn7002-RbcX (40μM) and Syn7002-RbcS (7 μM) (Lane 1) or At-RbcS1A (7-35 μM) (Lanes 2-6). The translation products were analyzed for the carboxylation activity.

(B) Syn7002-*rbcLX* (Lanes 1 and 2) or At-*rbcLX* (Lanes 3 and 4) were cooverexpressed along with GroEL/GroES in *E. coli* by using IPTG inducible coexpression plasmids. The soluble *E. coli* lysates were supplemented with purified Syn7002-RbcS or At-RbcS1A (7 µM (+) or 14 µM (++)) and carboxylation activity was measured.
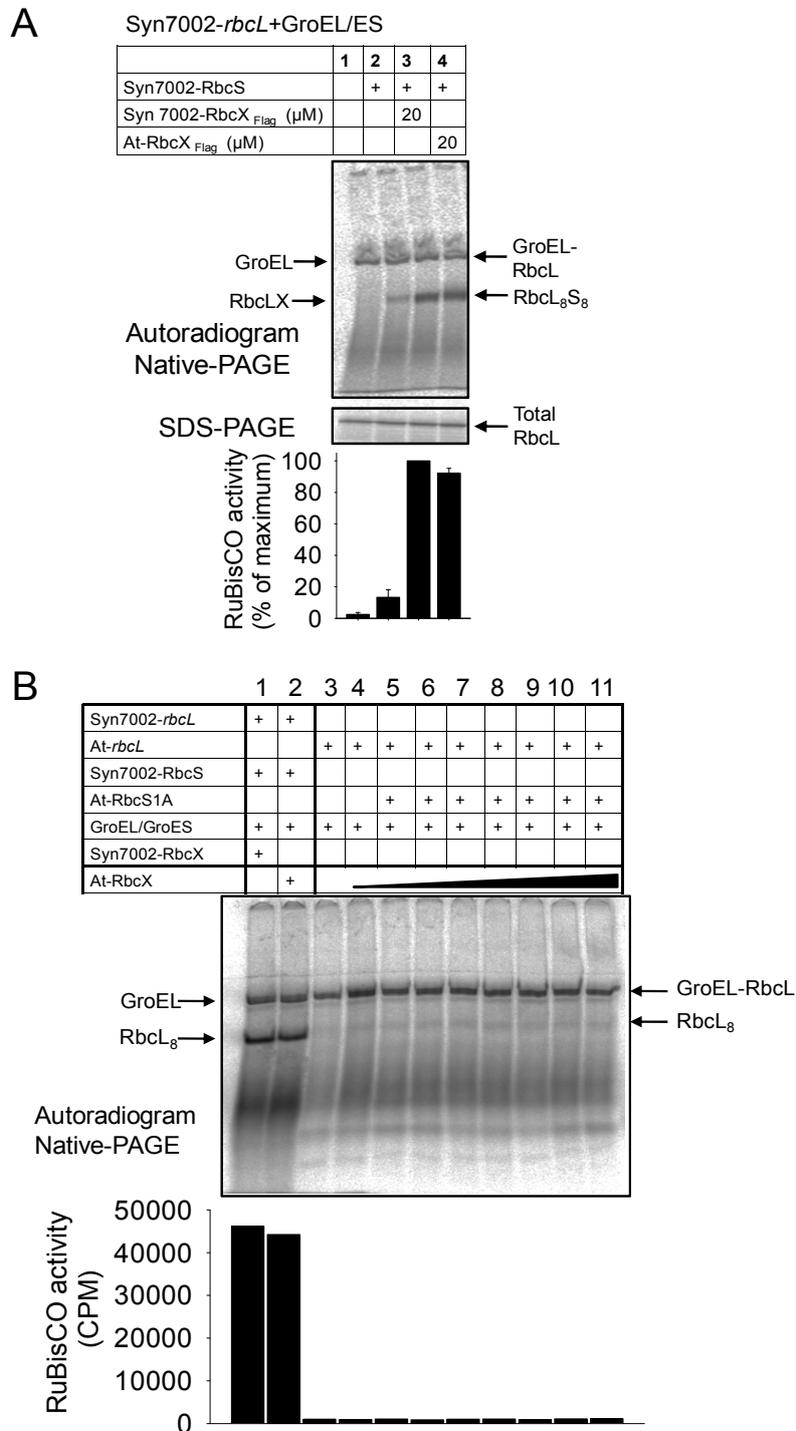
Figure 46. Functional analysis of At-RbcX

(A) Syn7002-*rbcL* was translated in *E.coli* lysate *in vitro* supplemented with ^{35}S-methionine, GroEL/GroES (0.5μM/1.0μM), Syn7002-RbcS (7 μM) (Lanes 2-4) and 20μM Syn7002-RbcX (Lane 3) or At-RbcX (Lanes 4). Translation was inhibited by the addition of CAM. The soluble and total protein in the products of the translation reactions were analyzed by Native-PAGE and SDS-PAGE respectively, followed by autoradiography. For the carboxylation activity, translation reactions were carried out in presence of unlabelled methionine and after the CAM addition products were analyzed for the RuBisCO activity.

(B) Syn7002-*rbcL* (Lanes 1 and 2) or At-*rbcL* (Lanes 3-11) was translated in *E.coli* lysate *in vitro* supplemented with ^{35}S-methionine, GroEL/GroES (0.5μM/1.0μM), Syn7002-RbcS (14 μM) (Lanes 1 and 2) or At-RbcS1A (14 μM) and 20μM Syn7002-RbcX (Lane 1) or up to 20μM of At-RbcX (Lanes 2 and 4-11). Translation was inhibited by the addition of CAM. The soluble products of the translation reactions were analyzed by Native-PAGE followed by autoradiography. For the carboxylation activity, translation reactions were carried out in presence of unlabelled methionine and after the CAM addition products were analyzed for the RuBisCO activity.

5.4.5   At-RbcX can functionally replace Syn7002-RbcX

In order to analyze the role of At-RbcX in the assembly of RuBisCO, the following *in vitro* translation experiment was performed. Translation of Syn7002-*rbcL* in the presence of At-RbcX and Syn7002-RbcS produced $RbcL_8S_8$ active holoenzyme, demonstrating that At-RbcX could complement for Syn7002-RbcX (Figure 46A, Lanes 3-4, and Figure 46B, Lanes 1-2). However, translating At-*rbcL* in the presence of GroEL/ES or Ch-Cpn60/20 and varying amounts of At-RbcX and At-RbcS1A did not yield active RuBisCO (Figure 46B, Lanes 3-11 and Figure 45B, Lanes 5-6), suggesting the need of an additional factor/s for the assembly of the higher plant enzyme.

5.4.6   Characterization of At-RbcX (Q29A)

In case of Syn7002-RbcX, the mutation Q29A, disrupting the peripheral polar surface of RbcX, failed to promote the formation of $Syn7002-RbcL_8$ complexes (Figure 27). To analyze the effect of a similar point mutation in At-RbcX, Syn7002-*rbcL* was translated in the presence of At-RbcX(Q29A). This RbcX mutant failed to

produce RbcL$_8$ complexes, and in turn active RuBisCO upon addition of RbcS (Figure 47, Lane 4).

Syn7002-*rbcL*+GroEL/ES
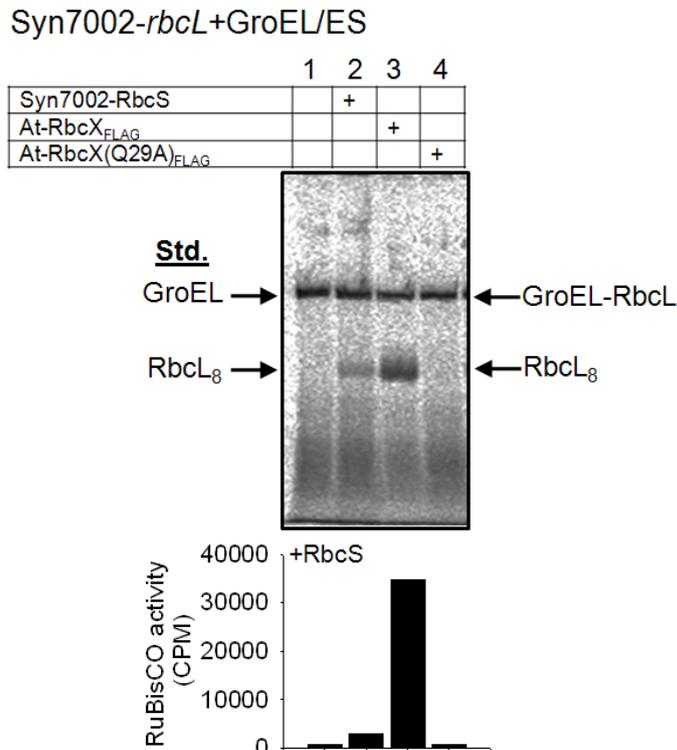
Figure 47. Q29A mutation disrupts At-RbcX function
Syn7002-*rbcL* was translated in *E. coli* lysate *in vitro* in the presence of ^{35}S-methionine. GroEL/GroES (0.5µM/1.0µM), Syn7002-RbcS (7 µM), At-RbcX (10 µM) and At-RbcX (Q29A) (10 µM) were supplemented to the reactions wherever indicated. RbcL$_8$ formation in products was monitored by Native-PAGE followed by autoradiography. For the carboxylation activity translation reactions were carried out in presence of unlabelled methionine and after the CAM addition, products were analyzed for the RuBisCO activity upon addition of RbcS.

5.4.7   At-RbcX binds the conserved C-terminal RbcL peptide similar to Syn7002-RbcX

To test whether *A. thaliana* RbcX can bind to the C-terminal EIKFEFD or a similar motif on RbcL, an array of acetylated peptides covering the C-terminal sequence of RbcL from different organisms, including cyanobacteria, red algae and higher plants (with a ten residue overlap) was incubated with At-RbcX$_{FLAG}$. The detection using anti-FLAG antibody resulted in a strong binding signal with four RbcL peptides sharing the sequence EIKFEFD or a similar motif close to the C-terminus of RbcL (Figure 48B) from most organisms including *Synechococcus* sp. PCC7002 and *A. thaliana*, similar to Syn7002-RbcX (Figures 48A and 29A). Essentially the same binding pattern was observed with the RbcX polar surface mutant Q29A,

albeit with reduced efficiency (Figure 48A). While binding of Syn7002-RbcX (Q29A) to Syn7002-RbcL peptide was stronger, binding to At-RbcL peptide was weaker (Figure 48A).

These results again confirm that the interaction with the RbcL C-terminus is a general property of cyanobacterial and higher plant RbcX homologues. Surprisingly, At-RbcX was able to detect some residues in the C-terminus of RbcL in red algae as well (in which there are no reports yet about RbcX homologues) (Figure 48A), hinting at the possible presence of an RbcX homologue.
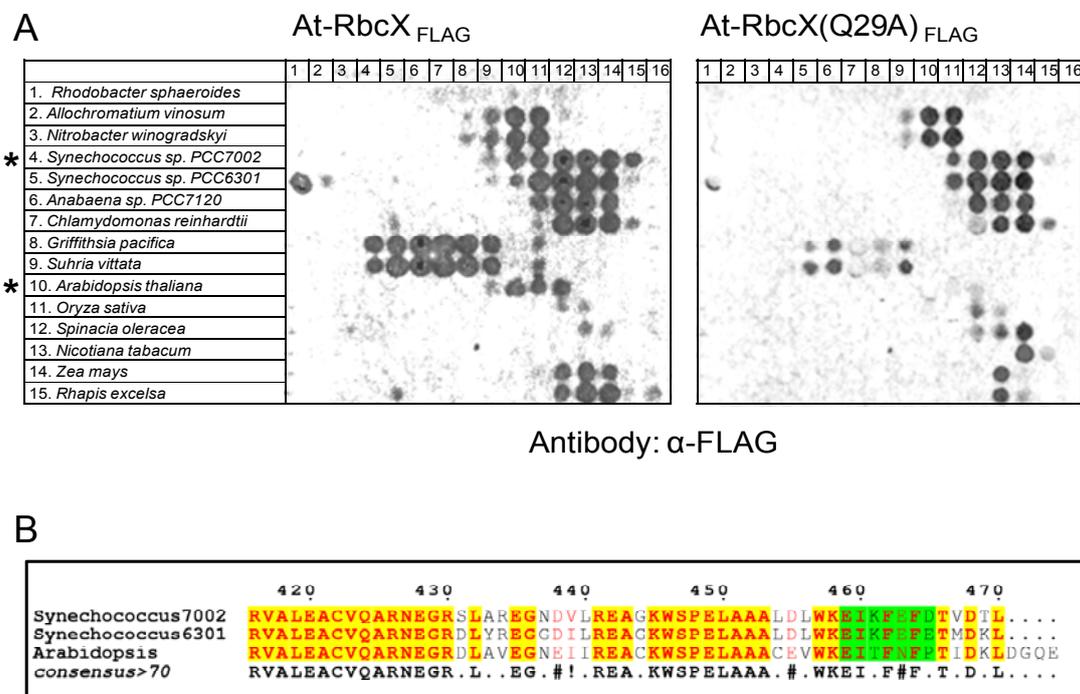
Figure 48. Binding of C-terminal RbcL peptide to At-RbcX
(A) A cellulose membrane containing an array of overlapping peptides covering the sequence of C-terminal Syn7002-RbcL was probed with the RbcX proteins indicated. Peptide-bound RbcX was visualized by chemiluminescent immunodetection with anti-FLAG antibody. RbcL C-terminal sequences of the corresponding organisms are as mentioned in the Figure 30B(i).
(B) Comparison of C-terminal RbcL sequences. Alignment using MultiAlin of the RbcL sequences of *Synechococcus sp.* PCC7002 (Q44176), *Synechococcus sp*. PCC6301 (P00880) and *Arabidopsis thaliana* (O03042). High consensus level (≥90%) is depicted in yellow background. The box highlighted shows the residues detected by RbcX protein binding.

6 Discussion

The yield of plants depends mainly on the efficiency of their photosynthesis and in turn on the functions of key enzymes like RuBisCO. But biochemically and physiologically RuBisCO is inefficient, mainly because it is substrate unspecific and participates in two competing reactions, photosynthesis and photorespiration which occur at the same active site (Figure 10 and 13B). Photosynthesis is the process through which plants fix carbon from the atmosphere, whereas photorespiration is an apparently wasteful reaction that leads to a loss of a considerable amount of fixed carbon (Hartman and Harpel, 1994; Mann, 1999) and consumes energy. Moreover, RuBisCO is a poor catalyst with a small turn-over number that is among the lowest for any known biological catalyst (Cleland et al., 1998; Tabita, 1999). This presumably accounts for the amazing abundance of RuBisCO in the biosphere.

in vitro mutagenesis of RuBisCOs may be one of the possible solutions to overcome these limitations. But *in vivo* expression studies in *E. coli* and *in vitro* reconstitutions using Form I RuBisCO have not been successful (Andrews and Lorimer, 1985; Chaudhari and Roy, 1989; Cloney et al., 1993; van der Vies et al., 1986). GroEL (Ch-Cpn60) and GroES (Ch-Cpn20/10) chaperones have been shown to assist in the folding of newly translated RbcL polypeptides (Ferreyra et al., 1993; Gatenby and Ellis, 1990; Gatenby et al., 1990; Goloubinoff et al., 1989; Gutteridge, 1995; Lubben et al., 1989). In addition, recently, RbcX has been implicated in the enhanced production of active RuBisCO (Larimer and Soper, 1993; Li and Tabita, 1997; Onizuka et al., 2004). In some cyanobacteria, the *rbcX* gene exists in the intergenic space between *rbcL* and *rbcS,* while in others it exists in a different operon. However, the exact role of RbcX is not clear. In the present study, cyanobacterial RuBisCO, which is very similar to the RuBisCO from higher plants, was used to analyze the folding and assembly of Form I RuBisCO, and the role of RbcX in this process.

6.1 GroEL/ES assisted folding of RuBisCO

RuBisCO is one of the stringent substrates (class 3 substrates) of GroEL; it folds efficiently only when it is enclosed in the GroEL/GroES cage (Brinker et al., 2001). Consistently, the *in vitro* translation of *rbcL* in GroEL depleted lysate did not result in RbcL$_8$ complex, even in the presence of other necessary factors, such as RbcX and/or minor amounts of other endogenous chaperones present in *E.coli* lysate (Figure 21).

6.2 RbcX mediated assembly of RuBisCO

The present study defines RbcX as an assembly chaperone of form I RuBisCO from cyanobacteria and resulted in the following model of RuBisCO folding and assembly: Newly-synthesized RbcL subunits bind to GroEL (Figure 49, step 1) and are folded by GroEL/GroES system (Figure 49, steps 2 and 3). Additional chaperones such as DnaK/DnaJ/GrpE might be involved upstream of GroEL/GroES in stabilizing and preventing aggregation of RbcL chains upon synthesis on ribosomes (Figure 49, step 1) (Checa and Viale, 1997). RbcL recycles to GroEL until it reaches its native state (Figure 23). The functional unit of RbcX is a dimer. RbcX interacts with RbcL subunits (Figure 49, step 4), most probably with RbcL monomers immediately upon release from GroEL (Figure 49, step 4) and so stabilizes assembly intermediates in the formation of RbcL$_8$ core complexes (Figure 49, step 5 and 6). Alternatively, RbcL monomers may spontaneously associate to form dimers without any assistance from other factors, as in case of Form II RuBisCO from *R. rubrum.* RbcX then binds these dimers and stabilize them for further assembly (Figure 49, step 5, brackets). Consistent with the first hypothesis, RbcX proteins synthesized in the *E.coli* translation system *in vitro* (in presence of labeled ^{35}S-methionine) were able to interact with newly-synthesized RbcL monomers generated in the presence of unlabelled methionine (Figure 35). In addition, LC/MSMS based mass spectrometry analysis detected the presence of RbcX protein with RbcL monomers (Figures 36 and 37).

The central groove of RbcX recognizes the C-terminal peptide of RbcL (Figures 29, 30 and 49). The disruption of the RbcX groove results in the loss of RbcX function in maintaining RbcL soluble and hence leads to the aggregation of RbcL (Figures 29A, 33D and 34A). The central groove shields the hydrophobic side chains of the RbcL C-terminal sequence and serves to position peripheral binding surfaces of RbcX in proximity to RbcL.
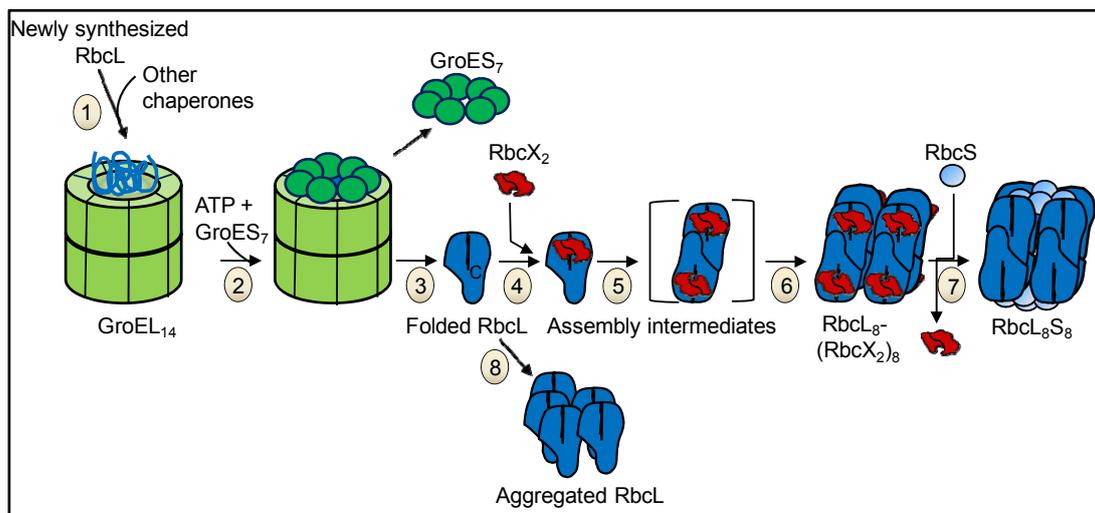
Figure 49. Working model of GroEL/GroES and RbcX function in cyanobacterial RuBisCO assembly

Newly synthesized RbcL peptides interact with chaperones including GroEL (step 1) and are folded by the GroEL/ES system (step 2); upon ATP hydrolysis the folded RbcL subunits are released from the GroEL chamber (step 3). The assembly chaperone RbcX (dimer) binds the C-terminal RbcL peptide (step 4 or 5) and supports the subsequent assembly to $RbcL_8$ core complex (steps 4-6). Individual RbcS (small) subunits are spontaneously added on top of the $RbcL_8$ core, displacing the RbcX (step 7) to yield $RbcL_8S_8$ RuBisCO holoenzyme. In the absence of RbcX, RbcL subunits tend to misassemble and form assembly incompetent aggregates (step 8).

The dynamic nature of the RbcX-RbcL interaction (Figure 32B) facilitates the eventual displacement of RbcX by RbcS subunits which is essential for the formation of functional holoenzyme (Figure 49, step 7). Unlike more general chaperones Hsp70 and GroEL, RbcX recognizes structural features specific to RbcL (Figures 29 and 30). Though the dependence on RbcX varies among RuBisCO homologs, the conservation of the RbcX recognition motif at the C-terminus of RbcL across many organisms possessing Form I RuBisCO, suggests a

general role of RbcX in ensuring efficient RuBisCO assembly. In the absence of RbcX, RbcL subunits misassemble resulting in aggregation (Figure 49, Step 8 and Figure 22).

6.3 Structural aspects of RbcX and RbcL interaction

The arc-shaped RbcX dimer has a central binding cleft for the binding of C-terminal peptide, EIKFEFD, of RbcL (residues 459 to 465 in Syn7002-RbcL) (Figures 30C, 30D and 48). The peptide is bound in an extended conformation with the side chains F462 and F464 occupying hydrophobic pockets of the RbcX cleft (Figure 30D). In addition, hydrogen bonds between the peptide backbone and RbcX residues lining the groove contribute to specific binding. Although RbcX is all α-helical, this binding mode is generally reminiscent of the interaction of extended hydrophobic peptides with the β sandwich-binding cleft of the general chaperone Hsp70 (Zhu et al., 1996). As the ends of polypeptide chains are often surface exposed, specific recognition of N- or C-terminal segments may be employed more generally by assembly chaperones.

Since the C-terminal RbcL peptide is conserved among many photosynthetic organisms having hexadecameric RuBisCO (Figure 30B), it might have an important structural and/or functional role. Because disruption of peptide binding abolished the formation of soluble, assembly competent RbcL protein (Figure 32A), RbcX appears to protect this sequence from undergoing aberrant interactions. The C-terminal peptide of RbcL is located on the outer surface of assembled RuBisCO and is assumed to have a regulatory function in catalysis (Figure 50A). It is thought to cycle during the enzymatic reaction between a more open conformation and a tightly bound state, stabilizing the lid segment of the active site to enclose the substrate ribulose-bisphosphate (or the transition state analog, carboxyarabinitol bisphosphate [CABP] (Zhu, 1998). Formation of the closed state involves Asp468, immediately adjacent to the RbcX-recognition motif. In the crystal structure of RuBisCO from *Synechococcus* sp. PCC6301 (in complex with CABP; Figure 50A), the C-terminal RbcL peptide is not covered by RbcS but would nevertheless be

inaccessible to RbcX as it is not sufficiently detached from the complex, and its hydrophobic side chains are only partially solvent exposed. Thus, RbcX may interact with the C-terminal sequence until its attachment to the main body of RuBisCO occurs.

The central RbcX-binding cleft functions in cooperation with polar binding surfaces around the corners of the RbcX dimer. Mutation of the peripheral region (Q29A) in RbcX impaired RbcL assembly in a manner distinct from that observed upon disabling the central peptide cleft (Y17A, Y20L) (Figures 27, 47 and 48A). The formation of soluble but misassembled RbcL of >800 kDa under these conditions (Saschenbrecker et al., 2007), indicates that the peripheral RbcX surface promotes the proper alignment of RbcL dimers for efficient assembly of $RbcL_8$ complexes, ultimately to be stabilized by RbcS (Curmi et al., 1992).

Binding to the RbcL C-terminus would serve as an anchor to position the peripheral binding surfaces, perhaps allowing RbcX to bridge adjacent RbcL subunits. Assuming the α-helix preceding the C-terminal binding motif in RbcL remains in place, one edge of RbcX would be able to contact the gap between adjacent RbcL dimers, which is occupied by RbcS in the holoenzyme structure (Figure 50).

How exactly RbcL interacts with the corner regions of RbcX remains to be defined, but an overlap of RbcX- and RbcS-binding regions could contribute to efficient RbcX displacement.

6.4 Dynamic nature of RbcX function

The binding and release of RbcX from its substrate is independent of ATP-regulated conformational changes. The fast off rate between $RbcL_8$ and RbcX favors the RbcX displacement by RbcS, leading to the stable interaction of RbcS and $RbcL_8$ (Andrews, 1988) which is also necessary to stabilize the $RbcL_8$ core and to form the $RbcL_8S_8$ holoenzyme (Figure 32B). It seems that the assembly

chaperones might follow a strategy involving a relatively weak interaction coupled with high dynamics, to achieve efficient product release. The binding parameters imply that the affinities between RbcX and $RbcL_8$ or RbcS and $RbcL_8$ have been optimized for each organism, which allows recycling of RbcX.
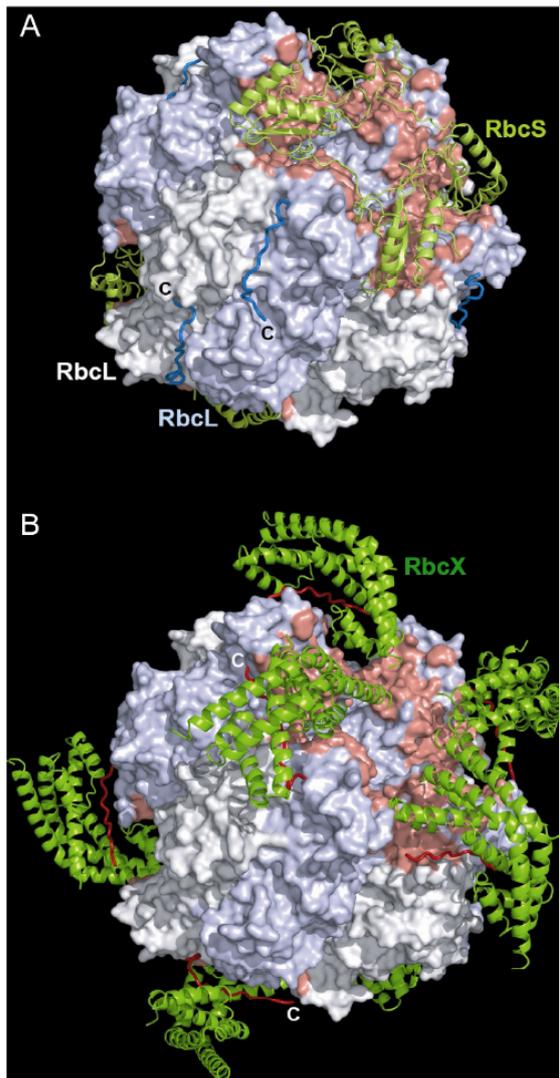
Figure 50. Crystal structure of Syn6301-$RbcL_8S_8$ and model of the $RbcL_8$-RbcX complex.

(A) Structure of RuBisCO of *Synechococcus* sp. PCC6301 bound to CABP (Newman et al., 1993). RbcL subunits are shown in surface presentation, with the exception of the C-terminal peptides (amino acids 460-475), which are presented as blue coils. This conformation is found only in complex with the enzymatic substrate. Subunits of RbcL dimers are colored white and pale blue and residues interacting with RbcS (green ribbons) are depicted in red.

(B) Hypothetical model of eight RbcX (dimers) binding to the assembled $RbcL_8$ core complex using the same representation as in (A). The model is based on the structures of the Syn6301-$RbcL_8S_8$ complex and the RbcX dimer. One edge of RbcX fills the gap between adjacent RbcL dimers, masking most of the surface covered by the globular domain of RbcS in the holoenzyme complex. In this arrangement, the C-termini of RbcL (in red) kink horizontally immediately N-terminal to the recognition motif, leaving the secondary structure composition of RbcL intact (The figure was prepared by Dr. Andreas Bracher).

It seems that RbcX works efficiently for its own RuBisCO. The affinities between RbcX and RbcL of different species vary, even though the proteins are evolutionarily highly related (Figure 31). One such example observed was the

interaction between *Synechococcus* sp. PCC7002-RbcL (Syn7002-RbcL) and *Anabaena* CA-RbcX (Ana-RbcX). The Ana-RbcX binds the C-terminal Syn7002-RbcL peptide with ~30-fold higher affinity than that of Syn7002-RbcX for the C-terminal of Syn7002-RbcL (Figure 31), demonstrating the significance of this adaptation. Formation of a stable complex, most likely Syn7002-RbcL$_8$-AnaCA-RbcX$_8$, was observed with Ana-RbcX (Figure 20, Lane 5 and Figure 32B, a), from which RbcX could not be displaced by RbcS (Figure 32B, b). This dead-end complex presumably defines the otherwise transitory end-state of RbcX activity in the assembly pathway prior to entry of RbcS.

6.5 Implications for the assembly of RuBisCO and other oligomeric proteins

The small turn over number of RuBisCO presumably, explains the amazing abundance of this enzyme (~500 active sites of RuBisCO for each CO_2 molecule in the chloroplast) (Palmer, 1996). So then how do photosynthetic organisms, including plants and green algae, coordinate the synthesis and assembly of such massive amounts of RuBisCO? This process necessitates an assembly chaperone, such as RbcX, in addition to the folding chaperones.

It has been proposed that the complex form I RuBisCO is derived from the simpler dimeric form II RuBisCO as an adaptation to the increased oxygen and lowered carbon dioxide in the atmosphere. The structural complexity might be the other reason for the necessity of assembly chaperone/s, specially when considering the complex role of the C-terminal RbcL segment in assembly and catalysis of form I RuBisCO (Zhu, 1998). *Synechococcus* sp. PCC7002-RuBisCO needs RbcX for its assembly. Similarly, the form I RuBisCO of other cyanobacteria and higher plants might depend on RbcX for assembly. However, translation of *A. thaliana-rbcL* in *E. coli* lysate *in vitro* in the presence of At-RbcS1A, At-RbcX and GroEL/ES or At-Ch-Cpn60/20 failed to form active enzyme (Figure 46B). In these preliminary experiments, only one type of small subunit (At-RbcS1A) (among four types) was used with combinations of chaperones (GroEL/ES, Ch-Cpn60/20 or 10).

In plants, the RuBisCO assembly might be more complex phenomenon as it takes place in the separate compartment, the chloroplast (Roy, 1989) to where small subunits have to be imported and processed before they interact with the $RbcL_8$ complex. The failure of heterologous form I RuBisCOs to assemble upon transgenic expression in plants (John Andrews and Whitney, 2003) may suggest an incompatibility of these enzymes with plant RbcX. In plants, RbcX is also encoded by a nuclear gene and has to be imported into the chloroplasts to perform its function. These complexity needs to be taken into account when performing *in vivo* expression studies. However, the present study strongly suggests that the plant chaperonin Ch-Cpn60/20 or 10 and RbcX have a defined role in the assembly of RuBisCO. The strong binding of higher plant RbcX to the C-terminal RbcL peptide from red algae (Figure 48A) suggests the possible existence of an RbcX or a similar assembly factor in red algae, which remains to be identified.

Overall, these findings suggest that in many cases the formation of complex oligomeric structures in cells might need machinery similar to RbcX, in addition to folding chaperones. The present study advances our understanding of the assembly of complex oligomeric proteins and should facilitate the search for similar factors assisting in the assembly of complex proteins. Future attempts to improve the catalytic performance of RuBisCO from crop plants through protein engineering may have to take the function of RbcX into consideration.

7   References

Agashe, V.R., Guha, S., Chang, H.C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: Increase in yield at the expense of folding speed. Cell *117*, 199-209.

Agashe, V.R., and Hartl, F.U. (2000). Roles of molecular chaperones in cytoplasmic protein folding. Seminars in Cell & Developmental Biology *11*, 15-25.

Allen, J.F., and Martin, W. (2007). Evolutionary biology: out of thin air. Nature *445*, 610-612.

Andersson, I. (1996). Large structures at high resolution: the 1.6 A crystal structure of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol bisphosphate. J Mol Biol *259*, 160-174.

Andersson, I. (2008). Catalysis and regulation in Rubisco. Journal of experimental botany *59*, 1555-1568.

Andersson, I., and Taylor, T.C. (2003). Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. Archives of biochemistry and biophysics *414*, 130-140.

Andrews, T.J. (1988). Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits. J Biol Chem *263*, 12213-12219.

Andrews, T.J., and Ballment, B. (1983). The function of the small subunits of ribulose bisphosphate carboxylase-oxygenase. J Biol Chem *258*, 7514-7518.

Andrews, T.J., and Lorimer, G.H. (1985). Catalytic properties of a hybrid between cyanobacterial large subunits and higher plant small subunits of ribulose bisphosphate carboxylase-oxygenase. J Biol Chem *260*, 4632-4636.

Anfinsen, C.B. (1972). The formation and stabilization of protein structure. Biochem J *128*, 737-749.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A.,Struhl, K. (1997). Current Protocols in Molecular Biology (NY, John Wiley & Sons).

Baneyx, F., Bertsch, U., Kalbach, C.E., van der Vies, S.M., Soll, J., and Gatenby, A.A. (1995). Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. J Biol Chem *270*, 10695-10702.

Beatrix, B., Sakai, H., and Wiedmann, M. (2000). The alpha and beta subunit of the nascent polypeptide-associated complex have distinct functions. J Biol Chem *275*, 37838-37845.

Bertsch, U., Soll, J., Seetharam, R., and Viitanen, P.V. (1992). Identification, characterization, and DNA sequence of a functional "double" groES-like chaperonin from chloroplasts of higher plants. Proc Natl Acad Sci U S A *89*, 8696-8700.

Booth, C.R., Meyer, A.S., Cong, Y., Topf, M., Sali, A., Ludtke, S.J., Chiu, W., and Frydman, J. (2008). Mechanism of lid closure in the eukaryotic chaperonin TRiC/CCT. Nature structural & molecular biology *15*, 746-753.

Boston, R.S., Viitanen, P.V., and Vierling, E. (1996). Molecular chaperones and protein folding in plants. Plant Mol Biol *32*, 191-222.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry *72*, 248-254.

Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., and Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. Cell *107*, 223-233.

Calvin, M. (1989). Forty years of photosynthesis and related activities. Photosynth Res *21*, 3-16.

Catanzariti, A.M., Soboleva, T.A., Jans, D.A., Board, P.G., and Baker, R.T. (2004). An efficient system for high-level expression and easy purification of authentic recombinant proteins. Protein Sci *13*, 1331-1339.

Chaudhari, P., and Roy, H. (1989). Delayed Osmotic Effect on in Vitro Assembly of RuBisCO : Relationship to Large Subunit-Binding Protein Complex Dissociation. Plant Physiol *89*, 1366-1371.

Checa, S.K., and Viale, A.M. (1997). The 70-kda heat-shock protein dnak chaperone system is required for the productive folding of ribulose-bisphosphate carboxylase subunits in escherichia coli. European Journal of Biochemistry *248*, 848-855.

Chitnis, P.R., and Nelson, N. (1991). Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium Synechocystis sp. PCC 6803. J Biol Chem *266*, 58-65.

Cleland, W.W., Andrews, T.J., Gutteridge, S., Hartman, F.C., and Lorimer, G.H. (1998). Mechanism of Rubisco: The Carbamate as General Base. Chemical reviews *98*, 549-562.

Cliff, M.J., Kad, N.M., Hay, N., Lund, P.A., Webb, M.R., Burston, S.G., and Clarke, A.R. (1999). A kinetic analysis of the nucleotide-induced allosteric transitions of GroEL. J Mol Biol *293*, 667-684.

Cloney, L.P., Bekkaoui, D.R., and Hemmingsen, S.M. (1993). Co-expression of plastid chaperonin genes and a synthetic plant Rubisco operon in Escherichia coli. Plant Mol Biol *23*, 1285-1290.

Cloney, L.P., Bekkaoui, D.R., Wood, M.G., and Hemmingsen, S.M. (1992a). Assessment of plant chaperonin-60 gene function in Escherichia coli. J Biol Chem *267*, 23333-23336.

Cloney, L.P., Wu, H.B., and Hemmingsen, S.M. (1992b). Expression of plant chaperonin-60 genes in Escherichia coli. J Biol Chem *267*, 23327-23332.

Coligan, J.E., Dunn, B.M., Speicher, D.W., Wingfield, P.T., Ploegh, H.L. (2000). Current Protocols in Protein Science (John Wiley & Sons, NY.).

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res *16*, 10881-10890.

Curmi, P.M., Cascio, D., Sweet, R.M., Eisenberg, D., and Schreuder, H. (1992). Crystal structure of the unactivated form of ribulose-1,5-bisphosphate carboxylase/oxygenase from tobacco refined at 2.0-A resolution. J Biol Chem *267*, 16980-16989.

Delwiche, C.F., and Palmer, J.D. (1996). Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. Mol Biol Evol *13*, 873-882.

Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature *400*, 693-696.

Dickson, R., Weiss, C., Howard, R.J., Alldrick, S.P., Ellis, R.J., Lorimer, G., Azem, A., and Viitanen, P.V. (2000). Reconstitution of higher plant chloroplast chaperonin 60 tetradecamers active in protein folding. J Biol Chem *275*, 11829-11835.

Dobson, C.M. (1999). Protein misfolding, evolution and disease. Trends Biochem Sci *24*, 329-332.

Drzymalla, C., Schroda, M., and Beck, C.F. (1996). Light-inducible gene HSP70B encodes a chloroplast-localized heat shock protein in Chlamydomonas reinhardtii. Plant Mol Biol *31*, 1185-1194.

Dunn, A.Y., Melville, M.W., and Frydman, J. (2001). Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT. Journal of structural biology *135*, 176-184.

Echave, P., Esparza-Ceron, M.A., Cabiscol, E., Tamarit, J., Ros, J., Membrillo-Hernandez, J., and Lin, E.C. (2002). DnaK dependence of mutant ethanol oxidoreductases evolved for aerobic function and protective role of the chaperone against protein oxidative damage in Escherichia coli. Proc Natl Acad Sci U S A *99*, 4626-4631.

Ellis, J. (1987). Proteins as molecular chaperones. Nature *328*, 378-379.

Ellis, R.J. (1979). The most abundant protein in the world. Trends in Biochemical Sciences *4*, 241-244.

Ellis, R.J. (2001a). Macromolecular crowding: an important but neglected aspect of the intracellular environment. Curr Opin Struct Biol *11*, 114-119.

Ellis, R.J. (2001b). Molecular chaperones: inside and outside the Anfinsen cage. Curr Biol *11*, R1038-1040.

Ellis, R.J. (2006). Molecular chaperones: assisting assembly in addition to folding. Trends Biochem Sci *31*, 395-401.

Ellis, R.J., and Hartl, F.U. (1999). Principles of protein folding in the cellular environment. Curr Opin Struct Biol *9*, 102-110.

Ellis, R.J., Robinson, C., van der Vies, S.M., and Kirwin, P.M. (1988). Processing and assembly of chloroplast proteins. Biochemical Society transactions *16*, 703-704.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci *8*, 978-984.

Emlyn-Jones, D., Woodger, F.J., Price, G.D., and Whitney, S.M. (2006). RbcX can function as a Rubisco chaperonin, but is non-essential in Synechococcus PCC7942 Plant Cell Physiol *47* 1630-1640.

Erbse, A., Mayer, M.P., and Bukau, B. (2004). Mechanism of substrate recognition by Hsp70 chaperones. Biochemical Society transactions *32*, 617-621.

Ferreyra, R.G., Soncini, F.C., and Viale, A.M. (1993). Cloning, characterization, and functional expression in Escherichia coli of chaperonin (groESL) genes from the phototrophic sulfur bacterium Chromatium vinosum. Journal of bacteriology *175*, 1514-1523.

Fischer, H.M., Schneider, K., Babst, M., and Hennecke, H. (1999). GroEL chaperonins are required for the formation of a functional nitrogenase in Bradyrhizobium japonium. Arch Microbiol *171*, 279-289.

Fitchen, J.H., Knight, S., Andersson, I., Branden, C.I., and McIntosh, L. (1990). Residues in three conserved regions of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase are required for quaternary structure. Proc Natl Acad Sci U S A *87*, 5768-5772.

Fredriksson, A., Ballesteros, M., Dukan, S., and Nystrom, T. (2005). Defense against protein carbonylation by DnaK/DnaJ and proteases of the heat shock regulon. Journal of bacteriology *187*, 4207-4213.

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J.S., Tempst, P., and Hartl, F.U. (1992). Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. Embo J *11*, 4767-4778.

Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F.U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. Nature *370*, 111-117.

Funfschilling, U., and Rospert, S. (1999). Nascent polypeptide-associated complex stimulates protein import into yeast mitochondria. Molecular biology of the cell *10*, 3289-3299.

Gatenby, A.A. (1984). European Journal of Biochemistry *144*, 361-366.

Gatenby, A.A., and Ellis, R.J. (1990). Chaperone function: the assembly of ribulose bisphosphate carboxylase-oxygenase. Annual review of cell biology *6*, 125-149.

Gatenby, A.A., Lubben, T.H., Ahlquist, P., and Keegstra, K. (1988). Imported large subunits of ribulose bisphosphate carboxylase/oxygenase, but not imported beta-ATP synthase subunits, are assembled into holoenzyme in isolated chloroplasts. Embo J *7*, 1307-1314.

Gatenby, A.A., van der Vies, S.M., and Rothstein, S.J. (1987). Co-expression of both the maize large and wheat small subunit genes of ribulose-bisphosphate carboxylase in Escherichia coli. European journal of biochemistry / FEBS *168*, 227-231.

Gatenby, A.A., Viitanen, P.V., and Lorimer, G.H. (1990). Chaperonin assisted polypeptide folding and assembly: implications for the production of functional proteins in bacteria. Trends in biotechnology *8*, 354-358.

Gautschi, M., Lilie, H., Funfschilling, U., Mun, A., Ross, S., Lithgow, T., Rucknagel, P., and Rospert, S. (2001). RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. Proc Natl Acad Sci U S A *98*, 3762-3767.

Gautschi, M., Mun, A., Ross, S., and Rospert, S. (2002). A functional chaperone triad on the yeast ribosome. Proc Natl Acad Sci U S A *99*, 4209-4214.

Geissler, S., Siegers, K., and Schiebel, E. (1998). A novel protein complex promoting formation of functional alpha- and gamma-tubulin. Embo J *17*, 952-966.

Genevaux, P., Georgopoulos, C., and Kelley, W.L. (2007). The Hsp70 chaperone machines of Escherichia coli: a paradigm for the repartition of chaperone functions. Molecular microbiology *66*, 840-857.

Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. EMBO Reports *5*, 195-200.

Gest, H. (1993). Photosynthetic and quasi-photosynthetic bacteria. FEMS MIcrobiology Letters *112*, 1-5.

Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S., and Mische, S.M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis *20*, 601-605.

Gibson, J.L., and Tabita, F.R. (1977). Different molecular forms of D-ribulose-1,5-bisphosphate carboxylase from Rhodopseudomonas sphaeroides. J Biol Chem *252*, 943-949.

Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Analytical biochemistry *182*, 319-326.

Glatz, A., Horvath, I., Varvasovszki, V., Kovacs, E., Torok, Z., and Vigh, L. (1997). Chaperonin genes of the Synechocystis PCC 6803 are differentially regulated under light-dark transition during heat stress. Biochem Biophys Res Commun *239*, 291-297.

Goloubinoff, P., Christeller, J. T., Gatenby, A. A., and Lorimer, G. H. (1989). Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfoleded state depends on two chaperonin proteins and Mg-ATP. Nature *342*, 884-889.

Goloubinoff, P., Gatenby, A.A., and Lorimer, G.H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli. Nature *337*, 44-47.

Goloubinoff, P., Gatenby, A.A., and Lorimer, G.H. (1989b). GroE heatshock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli Nature 337, 44-47.

Gurevitz, M., Somerville, C.R., and McIntosh, L. (1985). Pathway of assembly of ribulosebisphosphate carboxylase/oxygenase from Anabaena 7120 expressed in Escherichia coli. Proc Natl Acad Sci U S A *82*, 6546-6550.

Gutsche, I., Essen, L.O., and Baumeister, W. (1999). Group II chaperonins: new TRiC(k)s and turns of a protein folding machine. J Mol Biol *293*, 295-312.

Gutteridge, S.G., A. A. (1995). Rubisco Synthesis, Assembly, Mechanism, and Regulation. Plant Cell *7*, 809-819.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. J Mol Biol *166*, 557-580.

Hansen, S., Hough, E., and Andersen, K. (1999a). Purification, crystallization and preliminary X-ray studies of two isoforms of Rubisco from Alcaligenes eutrophus. Acta Crystallogr D Biol Crystallogr *55*, 310-313.

Hansen, S., Vollan, V.B., Hough, E., and Andersen, K. (1999b). The crystal structure of rubisco from Alcaligenes eutrophus reveals a novel central eight-stranded beta-barrel formed by beta-strands from four subunits. J Mol Biol *288*, 609-621.

Hansen, W.J., Cowan, N.J., and Welch, W.J. (1999c). Prefoldin-nascent chain complexes in the folding of cytoskeletal proteins. The Journal of cell biology *145*, 265-277.

Hanson, T.E., and Tabita, F.R. (2001). A ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from Chlorobium tepidum that is involved with sulfur metabolism and the response to oxidative stress. Proc Natl Acad Sci U S A *98*, 4397-4402.

Harrison, C.J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., and Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science *276*, 431-435.

Hartl, F.U. (1996a). Molecular chaperones in cellular protein folding. Nature *381*, 571-579.

Hartl, F.U. (1996b). Molecular chaperones in cellular protein folding. Nature *381*, 571-580.

Hartl, F.U., and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. Science *295*, 1852-1858.

Hartman, F.C., and Harpel, M.R. (1994). Structure, function, regulation, and assembly of D-ribulose-1,5-bisphosphate carboxylase/oxygenase. Annu Rev Biochem *63*, 197-234.

Haslberger, T., Weibezahn, J., Zahn, R., Lee, S., Tsai, F.T., Bukau, B., and Mogk, A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol Cell *25*, 247-260.

Hayashi, N.R., Arai, H., Kodama, T., and Igarashi, Y. (1997). The novel genes, cbbQ and cbbO, located downstream from the RubisCO genes of Pseudomonas hydrogenothermophila, affect the conformational states and activity of RubisCO. Biochem Biophys Res Commun *241*, 565-569.

Hayer-Hartl, M.K., Weber, F., and Hartl, F.U. (1996). Mechanism of chaperonin action - groes binding and release can drive groel-mediated protein folding in the absence of atp hydrolysis. EMBO Journal *15*, 6111-6121.

Hemmingsen, S.M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988). Nature *333*, 330-334.

Hill, J.E., and Hemmingsen, S.M. (2001). Arabidopsis thaliana type I and II chaperonins. Cell stress & chaperones *6*, 190-200.

Hirohashi, T., Nishio, K., and Nakai, M. (1999). cDNA sequence and overexpression of chloroplast chaperonin 21 from Arabidopsis thaliana. Biochimica et biophysica acta *1429*, 512-515.

Houry, W.A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F.U. (1999). Identification of in vivo substrates of the chaperonin GroEL. Nature *402*, 147-154.

I. Andersson, S.K., G. Schneider, Y. Lindqvist, T. Lundqvist, C.I. Branden and G.H. Lorimer. (1989). Nature *337*, 229-234.

Imker, H.J., Fedorov, A.A., Fedorov, E.V., Almo, S.C., and Gerlt, J.A. (2007). Mechanistic diversity in the RuBisCO superfamily: the "enolase" in the methionine salvage pathway in Geobacillus kaustophilus. Biochemistry *46*, 4077-4089.

John Andrews, T., and Whitney, S.M. (2003). Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants. Archives of biochemistry and biophysics *414*, 159-169.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M.*, et al.* (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209-220.

Kim, S., Willison, K.R., and Horwich, A.L. (1994). Cystosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains. Trends Biochem Sci *19*, 543-548.

Kitano, K., Maeda, N., Fukui, T., Atomi, H., Imanaka, T., and Miki, K. (2001). Crystal structure of a novel-type archaeal rubisco with pentagonal symmetry. Structure *9*, 473-481.

Knight, S., Andersson, I., and Branden, C.I. (1990). Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 A resolution. Subunit interactions and active site. J Mol Biol *215*, 113-160.

Koonin, E.V., and van der Vies, S.M. (1995). Conserved sequence motifs in bacterial and bacteriophage chaperonins. Trends Biochem Sci *20*, 14-15.

Koumoto, Y., Shimada, T., Kondo, M., Takao, T., Shimonishi, Y., Hara-Nishimura, I., and Nishimura, M. (1999). Chloroplast Cpn20 forms a tetrameric structure in Arabidopsis thaliana. Plant J *17*, 467-477.

Kovacs, E., van der Vies, S.M., Glatz, A., Torok, Z., Varvasovszki, V., Horvath, I., and Vigh, L. (2001). The chaperonins of Synechocystis PCC 6803 differ in heat inducibility and chaperone activity. Biochem Biophys Res Commun *289*, 908-915.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., and Hartl, F.U. (1992). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature *356*, 683-689.

Lanzetta, P.A., Alvarez, L.J., Reinach, P.S., and Candia, O.A. (1979). An improved assay for nanomole amounts of inorganic phosphate. Analytical biochemistry *100*, 95-97.

Larimer, F.W., and Soper, T.S. (1993). Overproduction of Anabaena 7120 ribulose-bisphosphate carboxylase/oxygenase in Escherichia coli. Gene *126*, 85-92.

Lauring, B., Kreibich, G., and Weidmann, M. (1995a). The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. Proc Natl Acad Sci U S A *92*, 9435-9439.

Lauring, B., Sakai, H., Kreibich, G., and Wiedmann, M. (1995b). Nascent polypeptide-associated complex protein prevents mistargeting of nascent chains to the endoplasmic reticulum. Proc Natl Acad Sci U S A *92*, 5411-5415.

Lee, B., and Tabita, F.R. (1990). Purification of recombinant ribulose-1,5-bisphosphate carboxylase/oxygenase large subunits suitable for reconstitution and assembly of active L8S8 enzyme. Biochemistry *29*, 9352-9357.

Lee, B.G., Read, B.A., and Tabita, F.R. (1991). Catalytic properties of recombinant octameric, hexadecameric, and heterologous cyanobacterial/bacterial ribulose- 1,5-bisphosphate carboxylase/oxygenase. Archives of biochemistry and biophysics *291*, 263-269.

Lehel, C., Los, D., Wada, H., Gyorgyei, J., Horvath, I., Kovacs, E., Murata, N., and Vigh, L. (1993). A second groEL-like gene, organized in a groESL operon is present in the genome of Synechocystis sp. PCC 6803. J Biol Chem *268*, 1799-1804.

Leroux, M.R., and Hartl, F.U. (2000). Protein folding: versatility of the cytosolic chaperonin TRiC/CCT. Curr Biol *10*, R260-264.

Li, H., Sawaya, M.R., Tabita, F.R., and Eisenberg, D. (2005). Crystal structure of a RuBisCO-like protein from the green sulfur bacterium Chlorobium tepidum. Structure *13*, 779-789.

Li, L.A., and Tabita, F.R. (1997). Maximum activity of recombinant ribulose 1,5-bisphosphate carboxylase/oxygenase of Anabaena sp. strain CA requires the product of the rbcX gene. Journal of bacteriology *179*, 3793-3796.

Lin, Z., and Rye, H.S. (2006). GroEL-mediated protein folding: making the impossible, possible. Critical reviews in biochemistry and molecular biology *41*, 211-239.

Lubben, T.H., Donaldson, G.K., Viitanen, P.V., and Gatenby, A.A. (1989). Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. Plant Cell *1*, 1223-1230.

Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I., and Koonin, E.V. (1999). Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome research *9*, 608-628.

Mann, C.C. (1999). Genetic engineers aim to soup up crop photosynthesis. Science *283*, 314-316.

Markesich, D.C., Gajewski, K.M., Nazimiec, M.E., and Beckingham, K. (2000). bicaudal encodes the Drosophila beta NAC homolog, a component of the ribosomal translational machinery*. Development (Cambridge, England) *127*, 559-572.

Martel, R., Cloney, L.P., Pelcher, L.E., and Hemmingsen, S.M. (1990). Unique composition of plastid chaperonin-60: alpha and beta polypeptide-encoding genes are highly divergent. Gene *94*, 181-187.

Martin-Benito, J., Boskovic, J., Gomez-Puertas, P., Carrascosa, J.L., Simons, C.T., Lewis, S.A., Bartolini, F., Cowan, N.J., and Valpuesta, J.M. (2002). Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. Embo J *21*, 6377-6386.

Martin-Benito, J., Gomez-Reino, J., Stirling, P.C., Lundin, V.F., Gomez-Puertas, P., Boskovic, J., Chacon, P., Fernandez, J.J., Berenguer, J., Leroux, M.R.*, et al.* (2007). Divergent substrate-binding mechanisms reveal an evolutionary specialization of eukaryotic prefoldin compared to its archaeal counterpart. Structure *15*, 101-110.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., and Hartl, F.U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. Nature *352*, 36-42.

Mayer, M.P., Schroder, H., Rudiger, S., Paal, K., Laufen, T., and Bukau, B. (2000). Multistep mechanism of substrate binding determines chaperone activity of Hsp70. Nat Struct Biol *7*, 586-593.

Mayhew, M., da Silva, A.C., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1996). Protein folding in the central cavity of the GroEL-GroES chaperonin complex. Nature *379*, 420-426.

McClellan, A.J., Scott, M.D., and Frydman, J. (2005a). Folding and quality control of the VHL tumor suppressor proceed through distinct chaperone pathways. Cell *121*, 739-748.

McClellan, A.J., Tam, S., Kaganovich, D., and Frydman, J. (2005b). Protein quality control: chaperones culling corrupt conformations. Nature cell biology *7*, 736-741.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003). Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell *113*, 369-381.

Milos, P., and Roy, H. (1984). ATP-released large subunits participate in the assembly of RuBP carboxylase. J Cell Biochem *24*, 153-162.

Mizohata, E., Matsumura, H., Okano, Y., Kumei, M., Takuma, H., Onodera, J., Kato, K., Shibata, N., Inoue, T., Yokota, A.*, et al.* (2002). Crystal structure of activated ribulose-1,5-bisphosphate carboxylase/oxygenase from green alga Chlamydomonas reinhardtii complexed with 2-carboxyarabinitol-1,5-bisphosphate. J Mol Biol *316*, 679-691.

Moller, I., Beatrix, B., Kreibich, G., Sakai, H., Lauring, B., and Wiedmann, M. (1998). Unregulated exposure of the ribosomal M-site caused by NAC depletion results in delivery of non-secretory polypeptides to the Sec61 complex. FEBS Lett *441*, 1-5.

Morse, D., Salois, P., Markovic, P., and Hastings, J.W. (1995). A nuclear-encoded form II RuBisCO in dinoflagellates. Science *268*, 1622-1624.

Mujacic, M., Bader, M.W., and Baneyx, F. (2004). Escherichia coli Hsp31 functions as a holding chaperone that cooperates with the DnaK-DnaJ-GrpE system in the management of protein misfolding under severe stress conditions. Molecular microbiology *51*, 849-859.

Murphy, R.M. (2002). Peptide aggregation in neurodegenerative disease. Annual Review of Biomedical Engineering *4*, 155-174.

Musgrove, J.E., Johnson, R.A., and Ellis, R.J. (1987). Dissociation of the ribulosebisphosphate-carboxylase large-subunit binding protein into dissimilar subunits. European journal of biochemistry / FEBS *163*, 529-534.

Newman, J., Branden, C.I., and Jones, T.A. (1993). Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301. Acta Crystallogr D Biol Crystallogr *49*, 548-560.

Newman, J., and Gutteridge, S. (1993). The X-ray structure of Synechococcus ribulose-bisphosphate carboxylase/oxygenase-activated quaternary complex at 2.2-A resolution. J Biol Chem *268*, 25876-25886.

Nishio, K., Hirohashi, T., and Nakai, M. (1999). Chloroplast chaperonins: evidence for heterogeneous assembly of alpha and beta Cpn60 polypeptides into a chaperonin oligomer. Biochem Biophys Res Commun *266*, 584-587.

Ohtaki, A., Kida, H., Miyata, Y., Ide, N., Yonezawa, A., Arakawa, T., Iizuka, R., Noguchi, K., Kita, A., Odaka, M.*, et al.* (2008). Structure and molecular dynamics simulation of archaeal prefoldin: the molecular mechanism for binding and recognition of nonnative substrate proteins. J Mol Biol *376*, 1130-1141.

Onizuka, T., Endo, S., Akiyama, H., Kanai, S., Hirano, M., Yokota, A., Tanaka, S., and Miyasaka, H. (2004). The rbcX gene product promotes the production and assembly of ribulose-1,5-bisphosphate carboxylase/oxygenase of Synechococcus sp. PCC7002 in Escherichia coli. Plant & cell physiology *45*, 1390-1395.

Ort, D.R. (1994). Photosynthesis. Encyclopedia of Agricultural Sciences *3* 187-195.

Palmer, J.D. (1996). Rubisco surprises in dinoflagellates. Plant Cell *8*, 343-345.

Parikh, M.R., Greene, D.N., Woods, K.K., and Matsumura, I. (2006). Directed evolution of RuBisCO hypermorphs through genetic selection in engineered E.coli. Protein Eng Des Sel *19*, 113-119.

Paul, K., Morell, M.K., and Andrews, T.J. (1991). Mutations in the small subunit of ribulosebisphosphate carboxylase affect subunit binding and catalysis. Biochemistry *30*, 10019-10026.

Pfund, C., Lopez-Hoyo, N., Ziegelhoffer, T., Schilke, B.A., Lopez-Buesa, P., Walter, W.A., Wiedmann, M., and Craig, E.A. (1998). The molecular chaperone Ssb from Saccharomyces cerevisiae is a component of the ribosome-nascent chain complex. Embo J *17*, 3981-3989.

Pushkin, A., Tsuprun, V., Solojeva, N., Shubin, V., Evstigneeva, Z., and Kretovich, W. (1982). High molecular weight pea leaf protein similar to the groE protein of escherichia coli. Biochim Biophys Acta *704*, 379-384.

Ranson, N.A., Burston, S.G., and Clarke, A.R. (1997). Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. J Mol Biol *266*, 656-664.

Ranson, N.A., White, H.E., and Saibil, H.R. (1998). Chaperonins. The Biochemical journal *333 (Pt 2)*, 233-242.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical chemistry *75*, 663-670.

Roessner, D., Kulicke, W.M. (1994). On-line coupling of flow field-flow fractionation and multiangle laser light scattering. . J Chromatogr *A. 687* 249-258.

Roseman, A.M., Chen, S., White, H., Braig, K., and Saibil, H.R. (1996). The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell *87*, 241-251.

Rowan, R., Whitney, S.M., Fowler, A., and Yellowlees, D. (1996). Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. Plant Cell *8*, 539-553.

Roy, H., Cannon, S., and Gilson, M. (1988). Assembly of Rubisco from native subunits. Biochimica et biophysica acta *957*, 323-334.

Roy, H.P.C. (1989). Rubisco Assembly: A Model System for Studying the Mechanism of Chaperonin Action. Plant Cell *1*, 1035-1042.

Rudi, K., Skulberg, O.M., and Jakobsen, K.S. (1998). Evolution of cyanobacteria by exchange of genetic material among phyletically related strains. Journal of bacteriology *180*, 3453-3461.

Rudiger, S., Schneider-Mergener, J., and Bukau, B. (2001). Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. Embo J *20*, 1042-1050.

Russell, R., Jordan, R., and McMacken, R. (1998). Kinetic characterization of the ATPase cycle of the DnaK molecular chaperone. Biochemistry *37*, 596-607.

Sambrook, J., Fritsch, E., Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition edn (NY, Cold Spring Harbor Press, Cold Spring Harbor).

Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F.U., and Hayer-Hartl, M. (2007). Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. Cell *129*, 1189-1200.

Schagger, H., and von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Analytical biochemistry *199*, 223-231.

Schenk, D. (2002). Opinion: Amyloid-ß immunotherapy for alzheimer's disease: The end of the beginning. Nature Reviews Neuroscience *3*, 824.

Schmid, F.X. (1990). Spectral methods of characterizing protein conformation and conformational changes (Oxford, IRL Press).

Schneider, G., Knight, S., Andersson, I., Branden, C.I., Lindqvist, Y., and Lundqvist, T. (1990). Comparison of the crystal structures of L2 and L8S8 Rubisco suggests a functional role for the small subunit. Embo J *9*, 2045-2050.

Schreuder, H.A., Knight, S., Curmi, P.M., Andersson, I., Cascio, D., Sweet, R.M., Branden, C.I., and Eisenberg, D. (1993). Crystal structure of activated tobacco rubisco complexed with the reaction-intermediate analogue 2-carboxy-arabinitol 1,5-bisphosphate. Protein Sci *2*, 1136-1146.

Sha, B., Lee, S., and Cyr, D.M. (2000). The crystal structure of the peptide-binding fragment from the yeast Hsp40 protein Sis1. Structure *8*, 799-807.

Sharkia, R., Bonshtien, A.L., Mizrahi, I., Weiss, C., Niv, A., Lustig, A., Viitanen, P.V., and Azem, A. (2003). On the oligomeric state of chloroplast chaperonin 10 and chaperonin 20. Biochimica et biophysica acta *1651*, 76-84.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Analytical chemistry *68*, 850-858.

Shi, S., Huang, Y., Zeng, K., Tan, F., He, H., Huang, J., and Fu, Y. (2005). Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. Molecular phylogenetics and evolution *34*, 159-166.

Shively, J.M., van Keulen, G., and Meijer, W.G. (1998). Something from almost nothing: carbon dioxide fixation in chemoautotrophs. Annu Rev Microbiol *52*, 191-230.

Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. Faseb J *10*, 27-34.

Shtilerman, M., Lorimer, G.H., and Englander, S.W. (1999). Chaperonin function: folding by forced unfolding. Science *284*, 822-825.

Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F.U. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. Embo J *18*, 75-84.

Siegert, R., Leroux, M.R., Scheufler, C., Hartl, F.U., and Moarefi, I. (2000a). Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. Cell *103*, 621-632.

Siegert, R., Leroux, M.R., Scheufler, C., Hartl, F.U., and Moarefi, I. (2000b). Structure of the molecular chaperone prefoldin: Unique interaction of multiple coiled coil tentacles with unfolded proteins. Cell *103*, 621-632.

Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A., and Horwich, A.L. (1998). Structure and function in GroEL-mediated protein folding. Annu Rev Biochem *67*, 581-608.

Sipe, J.D.C., A. S. (2000). Review: History of the amyloid fibril. Journal of Structural Biology *130* 88-98.

Slepenkov, S.V., and Witt, S.N. (2002). The unfolding story of the Escherichia coli Hsp70 DnaK: is DnaK a holdase or an unfoldase? Molecular microbiology *45*, 1197-1206.

Somerville, C.R., McIntosh, L., Fitchen, J., Gurevitz, M. (1986). The cloning and expression in Escherichia coli of RuBP carboxylase/oxygenase large subunit genes. Methods Enzymol *118*, 419-433.

Spiess, C., Meyer, A.S., Reissmann, S., and Frydman, J. (2004). Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. Trends in cell biology *14*, 598-604.

Spreitzer, R.J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Archives of biochemistry and biophysics *414*, 141-149.

Spreitzer, R.J., and Salvucci, M.E. (2002). Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. Annual review of plant biology *53*, 449-475.

Spreter, T., Pech, M., and Beatrix, B. (2005). The crystal structure of archaeal nascent polypeptide-associated complex (NAC) reveals a unique fold and the presence of a ubiquitin-associated domain. J Biol Chem *280*, 15849-15854.

Staehelin, L.A. (1986). Chloroplast structure and supermolecular organization of photosynthetic membranes. Encyclopedia of Plant Physiology New Series *19*, 1-84.

Straus, D., Walter, W., and Gross, C.A. (1990). DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. Genes & development *4*, 2202-2209.

Sugawara, H., Yamamoto, H., Shibata, N., Inoue, T., Okada, S., Miyake, C., Yokota, A., and Kai, Y. (1999). Crystal structure of carboxylase reaction-oriented ribulose 1, 5-bisphosphate carboxylase/oxygenase from a thermophilic red alga, Galdieria partita. J Biol Chem *274*, 15655-15661.

Sugiyama, T., Mizuno, M., and Hayashi, M. (1984). Partitioning of Nitrogen among Ribulose-1,5-bisphosphate Carboxylase/Oxygenase, Phosphoenolpyruvate Carboxylase, and Pyruvate Orthophosphate Dikinase as Related to Biomass Productivity in Maize Seedlings. Plant Physiol *75*, 665-669.

Tabita, F.R. (1999). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a different perspective. Photosynth Res *60*, 1-28.

Tabita, F.R., Hanson, T.E., Li, H., Satagopan, S., Singh, J., and Chan, S. (2007). Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. Microbiol Mol Biol Rev *71*, 576-599.

Tabita, F.R., Hanson, T.E., Satagopan, S., Witte, B.H., and Kreel, N.E. (2008). Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. Philosophical transactions of the Royal Society of London *363*, 2629-2640.

Tanaka, N., Hiyama, T., and Nakamoto, H. (1997). Cloning, characterization and functional analysis of groESL operon from thermophilic cyanobacterium Synechococcus vulcanus. Biochimica et biophysica acta *1343*, 335-348.

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J., and Andersson, I. (2001). First crystal structure of Rubisco from a green alga, Chlamydomonas reinhardtii. J Biol Chem *276*, 48159-48164.

Temussi, P.A.M., L.; Pastore, A. (2003). From alzheimer to huntington: Why is a structural understanding so difficult? . EMBO Journal *22*, 355-361.

Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F.U. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. Cell *97*, 755-765.

Thulasiraman, V., Yang, C.F., and Frydman, J. (1999). In vivo newly translated polypeptides are sequestered in a protected folding environment. Embo J *18*, 85-95.

Tsuprun, V.L., Boekema, E.J., Samsonidze, T.G., and Pushkin, A.V. (1991). Electron microscopy of the complexes of ribulose-1,5-bisphosphate carboxylase (Rubisco) and Rubisco subunit-binding protein from pea leaves. FEBS Lett *289*, 205-209.

Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., and Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. Cell *93*, 863-873.

Valpuesta, J.M., Martin-Benito, J., Gomez-Puertas, P., Carrascosa, J.L., and Willison, K.R. (2002). Structure and function of a protein folding machine: the eukaryotic cytosolic chaperonin CCT. FEBS Lett *529*, 11-16.

van den Berg, B., Ellis, R.J., and Dobson, C.M. (1999). Effects of macromolecular crowding on protein folding and aggregation. Embo J *18*, 6927-6933.

van der Vies, S.M., and Georgopoulos, C. (1996). Regulation of Chaperonin Gene Expression. In In The Chaperonins, R.J. Ellis, ed. (San Diego, CA, Academic Press), pp. 137-160.

van der Vies, S.M., Bradley, D., and Gatenby, A.A. (1986). Assembly of cyanobacterial and higher plant ribulose bisphosphate carboxylase subunits into functional homologous and heterologous enzyme molecules in Escherichia coli. Embo J *5*, 2439-2444.

Viitanen, P.V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G.H., Gatenby, A., and Soll, J. (1995). Functional characterization of the higher plant chloroplast chaperonins. J Biol Chem *270*, 18158-18164.

Viitanen, P.V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G. H., Gatenby, A., and Soll, J. (1995). Functional Characterization of the Higher Plant Chloroplast Chaperonins. J Biol Chem *270*, 18158-18164.

Voordouw, G., van der Vies, S.M., and Bouwmeister, P.P. (1984). Dissociation of ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach by urea. European journal of biochemistry / FEBS *141*, 313-318.

Vorderwulbecke, S., Kramer, G., Merz, F., Kurz, T.A., Rauch, T., Zachmann-Brand, B., Bukau, B., and Deuerling, E. (2004). Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK. FEBS Lett *559*, 181-187.

Walsh, D.M.H., D. M.; Selkoe, D. J. (2003). The many faces of Aß: Structures and activity. Current Medicinal Chemistry: Immunology, Endocrine & Metabolic Agents *3*.

Wastl, J., Fraunholz, M., Zauner, S., Douglas, S., and Maier, U.G. (1999). Ancient gene duplication and differential gene flow in plastid lineages: the GroEL/Cpn60 example. Journal of molecular evolution *48*, 112-117.

Weiner, M.P., Costa, G.L., Schoettlin, W., Cline, J., Mathur, E., and Bauer, J.C. (1994). Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. Gene *151*, 119-123.

Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A., and Horwich, A.L. (1995). Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. Cell *83*, 577-587.

Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., and Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. Cell *84*, 481-490.

Whitney, S.M., Shaw, D.C., and Yellowlees, D. (1995). Evidence that some dinoflagellates contain a ribulose-1,5-bisphosphate carboxylase/oxygenase related to that of the alpha-proteobacteria. Proc Biol Sci *259*, 271-275.

Wiedmann, B., Sakai, H., Davis, T.A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. Nature *370*, 434-440.

Wild, J., Rossmeissl, P., Walter, W.A., and Gross, C.A. (1996). Involvement of the DnaK-DnaJ-GrpE chaperone team in protein secretion in Escherichia coli. Journal of bacteriology *178*, 3608-3613.

Wyatt, P.J. (1993). Light scattering and the absolute characterization of macromolecules. Anal Chim Acta *272*, 1-40.

Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature *388*, 741-750.

Xu, Z., and Sigler, P.B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. Journal of structural biology *124*, 129-141.

Yang, T., and Poovaiah, B.W. (2000). Arabidopsis chloroplast chaperonin 10 is a calmodulin-binding protein. Biochem Biophys Res Commun *275*, 601-607.

Yasuko Koumoto, T.S., Maki Kondo, Toshifumi Takao, Yasutsugu Shimonishi, Ikuko Hara-Nishimura and Mikio Nishimura (1999). Chloroplast Cpn20 forms a tetrameric structure inArabidopsis thaliana. Plant journal *17*, 467-477.

Yoon, M., Putterill, J.J., Ross, G.S., and Laing, W.A. (2001). Determination of the relative expression levels of rubisco small subunit genes in Arabidopsis by rapid amplification of cDNA ends. Analytical biochemistry *291*, 237-244.

Young, J.C., Agashe, V.R., Siegers, K., and Hartl, F.U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. Nature reviews *5*, 781-791.

Zako, T., Murase, Y., Iizuka, R., Yoshida, T., Kanzaki, T., Ide, N., Maeda, M., Funatsu, T., and Yohda, M. (2006). Localization of prefoldin interaction sites in the hyperthermophilic group II

chaperonin and correlations between binding rate and protein transfer rate. J Mol Biol *364*, 110-120.

Zhu, G.H., Jensen, R.G., Bohnert, H.J., Wildner, G.F., and Schlitter, J. (1998). Dependence of catalysis and CO2/O2 specificity of Rubisco on the carboxy-terminus of the large subunit at different temperatures. Photosynth Res *57*, 71-79.

Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E., and Hendrickson, W.A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. Science *272*, 1606-1614.

Zietkiewicz, S., Krzewska, J., and Liberek, K. (2004). Successive and synergistic action of the Hsp70 and Hsp100 chaperones in protein disaggregation. J Biol Chem *279*, 44376-44383.

8 Appendices

8.1 Primers and vectors used for cloning *rbcL, rbcS1A* and *rbcX* from *A. thaliana*

Gene	Primer	Sequence	Remarks
At-*rbcL*	AtRbcL_Fw	5'-ggagatata**catatg**tcaccacaaacagagac-3'	Vector: pET11a Cloned between NdeI and BamHI
	AtRbcL_RevNew	5'-gcagcc**ggatcc**gctagc<u>tta</u>ctactcttggccatc-3'	
At-*rbcS1A* and At-RbcS3B	AtRbcS1A_Fw	5'-ggagatata**catatg**gcttcctctatgctctc-3'	Vector: pET11a Cloned between NdeI and NheI-BamHI; with transit peptide
	AtRbcS1A_Rev	5'- gcagcc**ggatcc**<u>tta</u>ttaaccggtgaagcttgg-3'	
	AtRbcS3B_Rev	5'-gcagcc**ggatcc**<u>tta</u>ttaagcttcggtgaagct-3'	Vector: pET11a Fw primer: AtRbcS1A_Fw Cloned between NdeI and BamHI; with transit peptide
	AtRbcS1A-TP_Fw	5'-ggagatata**catatg**caggtgtggcct -3'	Vector: pET11a Cloned between NdeI and BamHI; transit peptide sequence removed
	AtRbcS3B-TP_Fw	5'-ggagatata**catatg**aaggtgtggcca -3'	Cloned between NdeI and BamHI; transit peptide sequence removed
	AtS1ANHis6_Fw	5'- ggagatata**catatg**catcaccatcaccatcacgaa aacctgtattttcagcaggtgtggcctccg -3'	Vector: pET11a Cloned between NdeI and BamHI; N-His6 tag transit peptide sequence removed
	AtS1ANHis6_Rev	5'-cggaggccacacctgctgaaaatacaggttttcgtgat ggtgatggtgatg**catatg**tatatctcc-3'	
	AtS3BNHis6_Fw	5'- ggagatata**catatg**catcaccatcaccatcacgaaa acctgtattttcagaaggtgtggccacca -3'	Vector: pET11a Cloned between NdeI and BamHI; N-His6 tag transit peptide sequence removed
	AtS3BNHis6_Rev	5'-aattggtggccacaccttctgaaaatacaggttttcgtgat ggtgatggtgatg**catatg**tatatc-3'	
At-*rbcX*	AtRbcX_Fw	5'-gatata**catatg**ggtgagtgcttgg-3'	Vector: pET11a Cloned between NdeI and BamHI; with transit peptide
	AtRbcX_Rev	5'-gcagcc**ggatcc**<u>tta</u>ttaccttgagtttgtg-3'	
	AtRbcX-TP82_Fw	5'-ggagatata**catatg**ggtaattacgacgat-3'	Vector: pET11a Cloned between NdeI and BamHI; transit peptide sequence removed
	AtX-82pHUE_Fw1	5'-ct**ccgcgg**tggt<u>atg</u>ggtaattacgacgat-3'	Vector: pHUE Cloned between SacII and BamHI; transit peptide sequence removed
	AtX-82pHUE_Fw2	5'-ct**ccgcgg**tggtggtaattacgacgat-3'	
	pHueAtX82NFlag_Fw	5'-cgcct**ccgcgg**tggtgattacaaagacgatgacgataa agcgggcggtaattacgacgatacc-3'	Vector: pHUE Cloned between SacII and BamHI; transit peptide sequence removed N-FLAG tag
	pHueAtX82NFlag_Rev	5'- ggtatcgtcgtaattaccgcccgctttatcgtcatcgtcttg taatcacca**ccgcgg**aggcg - 3'	
	AtxQ29A_Fw AtxQ29A_Rev	5'-acggttcttcatgctctgtacgaaatgaat-3' 5'-attcatttcgtacagagcatgaagaaccgt-3'	Vector: pHUE Cloned between SacII and BamHI; transit peptide sequence removed N-FLAG tag, Q29A

Table 8.1 Primers and vectors used for cloning *rbcL, rbcS* and *rbcX* from *A. thaliana*. Introduced restriction sites are mentioned in bold; start and stop codons are underlined.

8.2 Abbreviations

AA	acrylamide
aa	amino acid
ADP	adenosine 5'-diphosphate
Amp	ampicillin
Amp^R	ampicillin resistance
AMP-PNP	adenosine 5′-(β,γ-imido)triphosphate
Ana	Anabaena
ANS	1-anilino-8-naphthalene-sulphonate
APS	ammonium peroxodisulfate
At	Arabidopsis thaliana
ATP	adenosine 5'-triphosphate
AU	arbitrary unit
BLAST	Basic Local Alignment Search Tool
β-NADH	β-nicotinamide adenine dinucleotide
bp	base pairs
BSA	bovine serum albumin
^{o}C	degree Celsius
^{14}C	carbon-14
Cam	chloramphenicol
Cam^R	chloramphenicol resistance
CAPB	carboxy-arabinitol-1,5-bisphosphate
cDNA	copy DNA
CDTA	trans-1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid
Ch	chloroplast
CIP	calf intestinal alkaline phosphatase
Cpn	chaperonin
cryo-EM	cryoelectron microscopy
CTAB	cetyltrimethylammoniumbromid
Δ (delta)	deletion
DE	diethylaminoethyl
DNA	deoxyribonucleic acid
DnaJ	bacterial Hsp40 chaperone
DnaK	bacterial Hsp70 chaperone
dNTP	2`-desoxyribonucleotide-triphosphate
DSP	dithiobis[succinimidylpropionate]
DTT	dithiothreitol
ECL	enhance chemiluminescence
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
ESI	elctro spray ionization
FFF	field flow fractionation
Fig.	Figure
fmet-tRNA	N-formyl-mehtionyl-tRNA
FPLC	fast protein liquid chromatography
FT	flowthrough
g	acceleration of gravity, 9.81 m/s2

GdnHCl	guanidinium hydrochloride
GFP	green fluorescent protein
Gp31	bacteriophage T4 co-chaperone
GroEL	bacterial Hsp60 chaperonin
GroES	bacterial Hsp10 co-chaperone
GrpE	bacterial nucleotide exchange factor of DnaK
GSH	reduced glutathione
GSSG	oxidized glutathione
h	hour
HEPES	N-(2-hydroxyethyl)piperazin-N`-2-ethanesulfonic acid
His6	histitin-tag
HRP	horseradish peroxidase
Hsp	heat shock protein
IPTG	isopropyl-β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
Kan	kanamycin
Kan^R	kanamycin
KD	dissociation constant
kDa	kilodalton
LB	Luria Bertani
MDH	malate dehydrogenase
MES	2-morpholinoethanesulfonic acid
MLS	multiangle light scattering
MOPS	3-(N-morpholino)propanesulfonic acid
MS	mass spectrometry
mt	mitochondrial
MW	molecular weight
MWCO	molecular weight cut off
N-	N-terminal
NADPH	β-nicotinamide adenine dinucleotide 2'-phosphate
Ni-NTA	nickel-nitrilotriacetic acid
OAc	acetate
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pfu	Pyrococcus furiosus
pH	reverse logarithm of relative hydrogen proton (H+) concentration
Pi	inorganic phosphate
PK	Proteinase K
PMSF	phenyl-methyl-sulfonyl fluoride
RbcL	RuBisCO large subunit
RbcS	RuBisCO small subunit
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-1,5-bisphosphate
PVDF	polyvinylidenfluorid
RNA	ribonucleic acid
RNase A	ribonuclease A
rpm	revolutions per minute

Rr	Rhodospirillum rubrum
RT	room temperature
RTS	rapid transcription/translation system
^{35}S	sulphur-35
SDS	sodiumdodecylsulfate
SeMet	selenomethionine
So	Spinacia oleracea
SeMet	Selenomethionine
sp.	species
Syn	Synechococcus
TAE	Tris-acetate-EDTA (EDTA; Tris)
TBS	Tris-buffered saline (Tris)
TBS-T	TBS containing Tween20 (TBS; Tween20)
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
TEV	Tobacco etch virus
TF	trigger factor
Tris	trishydroxymethylaminomethane
TritonX-100	octyl phenol ethoxylate
Tween20	polyoxyethylen-sorbitan-monolaurate
tRNA	transfer-RNA
UV/VIS	ultraviolet/visible
w/o	without
v/v	volume per volume
w/v	weight per volume

8.3 Publication

Saschenbrecker, S., Bracher, A., **Rao, K.V.**, Rao, B.V., Hartl, F.U., Hayer-Hartl, M. (2007). Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. Cell 129, 1189-1200.

Prasad, V., **Rao, K.V.**, Charng, Y.Y., and Chan, M.T. (2006). Tapping Molecular Cross-talk of Temperature Stress Tolerance in Plants - A Functional Genomic Approach. Global Impact of Plant Genetic Engineering on Food and the Environment. Jesse M. Jaynes, Dick Sawyer and Luis Destefano (Ed)

8.4 *Curriculum Vitae*

Personal Information

Name:	Karnam Vasudeva Rao
Date of Birth:	29 June 1973
Place of Birth:	Chintrapalli-Karnataka, India
Nationality:	Indian
Gender:	Male
Marital Status:	Married
Spouse name:	Bharathi Vasudeva Rao

Academic Qualification

Aug 2004-Till date	PhD student, Department of Cellular Biochemistry, Max Planck Institute for Biochemistry, Martinsried, Munich, Germany (Research area: Chaperone assisted RuBisCO folding and assembly)
1995-1996	M. Biotech in Biotechnology Gulbarga University, Gulbarga, India
1993-1995	M.Sc. (Master of Science) in Botany, specialization Biotechnology Gulbarga University, Gulbarga, India
1990-1993	B.Sc.(Bachelor of Science) in Chemistry, Botany, and Zoology Gulbarga University, Gulbarga, India

Research Experience

Dec 2002-Sep 2003	Worked as Research Assistant in the laboratory of Dr. Ming Tsair Chan, Institute of Bioagricultural Sciences, Academia Sinica, Taipei, Taiwan (R.O.C). **Research area:** Functional genomics study of tomato plants resistant to water deficit-and drought stress and orchid transformation.
Feb 2000-Nov 2002	Worked as Research officer in the Department of Molecular Biology Namdhari Seeds Pvt. Ltd., Bidadi, Bangalore, India.

	Research area: Molecular marker assisted plant breeding (RAPD, SCAR and Microsatellite (SSR) markers)-tomato, watermelon and muskmelon.
Nov 1997-Feb 2000	Worked as Research Officer in the Department of Biotechnology, Rallis Research Center, Peenya Industrial Area, Bangalore-560058, India. **Research area**: *Agrobacterium* mediated transformation of cabbage for Diamond Back Moth resistance.
June 1996-Oct 1997	Worked as Project Assistant in the group of Prof. Shashidhar H. E. and Prof. Shailaja Hittalamani, Dept. of Genetics and Plant Breeding, University of Agricultural Sciences (GKVK), Bangalore, India. **Research area**: Molecular marker assisted breeding: to evaluate genetic distance among locally available drought resistant rice varieties using molecular markers.