

Ludwig-Maximilians-Universität München  
Max-Planck-Institut für Biochemie, Martinsried

**Folding and assembly of RuBisCO /  
Structural and functional characterization  
of the RuBisCO assembly chaperone RbcX**

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

vorgelegt von  
Sandra Saschenbrecker  
aus Schwerin

2007

## **Erklärung**

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

## **Ehrenwörtliche Versicherung**

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfsmittel erarbeitet.

München, am

Dissertation eingereicht am 25. Juni 2007

1. Gutacher Prof. Dr. F. Ulrich Hartl

2. Gutachter PD Dr. Konstanze F. Winklhofer

Mündliche Prüfung am 22. Oktober 2007

## **Danksagung**

Die vorliegende Arbeit wurde in der Zeit von Juni 2003 bis Juni 2006 in der Abteilung Zelluläre Biochemie des Max-Planck-Instituts für Biochemie in Martinsried angefertigt.

Mein besonderer Dank gilt Prof. Dr. F. Ulrich Hartl für die Bereitstellung des interessanten Themas, seine intensive Förderung und Unterstützung sowie die hervorragenden Arbeitsbedingungen.

Des Weiteren möchte ich mich herzlich bei Dr. Manajit Hayer-Hartl für ihre hervorragende Betreuung, konstruktive Zusammenarbeit, ihr lebendiges Interesse an meiner Arbeit und ihren stetigen Optimismus bedanken.

Ebenfalls gilt mein Dank Dr. Andreas Bracher. Seine Hilfe, Bemühungen und fachliche Kompetenz insbesondere bei den kristallographischen und strukturanalytischen Arbeiten sowie seine methodische Kreativität haben maßgeblich zum Gelingen dieser Arbeit beigetragen.

Besonders danken möchte ich auch Dr. Enrico Schleiff and Dr. Thomas Becker für ihre Kooperation und Hilfe sowie für norddeutsche Reminiszenzen.

Allen Mitarbeitern der Abteilung Zelluläre Biochemie danke ich für ihre kollegiale Hilfe und die gute Arbeitsatmosphäre.

Mein herzlichster und größter Dank ist meiner Familie gewidmet. Ohne ihre uneingeschränkte Unterstützung und ihren Rückhalt wäre diese Arbeit nicht möglich gewesen.

**CONTENTS**

<b>1</b>	<b>SUMMARY</b>	<b>1</b>
<b>2</b>	<b>INTRODUCTION</b>	<b>3</b>
2.1	Protein structure	3
2.2	Protein folding	4
2.3	Protein folding and aggregation <i>in vivo</i>	5
2.4	Molecular chaperones	8
2.4.1	Chaperones involved in <i>de novo</i> protein folding	9
2.4.2	Ribosome-associated chaperones	10
2.4.3	The Hsp70 system	11
2.4.4	Prefoldin/GimC	13
2.4.5	The chaperonins	14
2.4.5.1	Structure and function of the <i>E. coli</i> chaperonin system	14
2.4.5.2	Mechanism of GroEL/ES mediated protein folding	16
2.4.5.3	Chloroplast (and cyanobacterial) chaperonins	19
2.5	Photosynthesis	22
2.5.1	Light-dependent reactions	23
2.5.2	Light-independent reactions / The Calvin cycle	25
2.6	RuBisCO	26
2.6.1	Reactions catalyzed by RuBisCO	26
2.6.2	Regulation of RuBisCO activity	28
2.6.3	Structure of RuBisCO	29
2.6.4	Folding and assembly of RuBisCO and the role of chaperones	33
2.6.5	RbcX	35
2.7	Aim of the study	36
<b>3</b>	<b>MATERIALS AND METHODS</b>	<b>38</b>
3.1	Materials	38
3.1.1	Chemicals	38
3.1.2	Reagent and purification kits	38
3.1.3	Strains	39
3.1.4	Plasmids, DNA and oligonucleotides	39
3.1.5	Enzymes, proteins, peptides and antibodies	40
3.1.6	Media	40
3.2	Instruments	41
3.3	Molecular biological methods	42
3.3.1	DNA analytical methods	42
3.3.2	Preparation and transformation of competent <i>E. coli</i> cells	43
3.3.2.1	Chemocompetent <i>E. coli</i> cells and chemical transformation	43
3.3.2.2	Electrocompetent <i>E. coli</i> cells and electroporation	43

3.3.2.3	TSS-transformation	44
3.3.3	Isolation of chromosomal DNA from <i>Synechococcus</i> sp. PCC7002	44
3.3.4	Purification of plasmid DNA and DNA-fragments	45
3.3.5	PCR (polymerase chain reaction)	45
3.3.6	Restriction digest and ligation	46
3.3.7	Cloning strategies	47
<b>3.4</b>	<b>Protein biochemical and biophysical methods</b>	<b>49</b>
3.4.1	Protein analytical methods	49
3.4.1.1	Protein quantification	49
3.4.1.2	SDS-PAGE	49
3.4.1.3	Tricine-PAGE	50
3.4.1.4	Native PAGE	50
3.4.1.5	Bis-Tris Native PAGE	51
3.4.1.6	SDS-PAGE analysis of Native PAGE protein bands	51
3.4.1.7	Coomassie blue staining of polyacrylamide gels	51
3.4.1.8	Silver staining of polyacrylamide gels	52
3.4.1.9	Phosphoimaging	52
3.4.1.10	Western blotting and immunodetection	52
3.4.1.11	Generation of antiserum	53
3.4.1.12	TCA precipitation	53
3.4.1.13	FFF-MALS	54
3.4.1.14	N-terminal sequencing of proteins	54
3.4.1.15	Sequence alignments	54
3.4.2	Protein expression and purification	55
3.4.2.1	AtCpn60 $\alpha\beta$	55
3.4.2.2	AtCpn60 $\beta$	56
3.4.2.3	AtCpn60 $\alpha$	56
3.4.2.4	AtCpn20, AtCpn10 and SoCpn20	57
3.4.2.5	PsCpn20	57
3.4.2.6	AtCpn20 <sub>N-His6</sub> , AtCpn20/N <sub>His6</sub> and AtCpn20/C <sub>His6</sub>	58
3.4.2.7	Syn6301-RbcL <sub>8</sub> S <sub>8</sub>	58
3.4.2.8	Syn6301-RbcL <sub>8</sub>	59
3.4.2.9	Syn6301-RbcS and Syn7002-RbcS <sub>FLAG</sub>	60
3.4.2.10	Syn7002-RbcX and AnaCA-RbcX	61
3.4.2.11	(SeMet)-Syn7002-RbcX and (SeMet)-Syn7002-RbcX( $\Delta$ C25)	61
3.4.2.12	Syn7002-RbcLX <sub>N-His6</sub>	62
3.4.2.13	Syn6301-RbcL/AnaCA-RbcX <sub>N-His6</sub>	63
3.4.3	Functional analyses	63
3.4.3.1	ATPase activity assay	63
3.4.3.2	Aggregation prevention assay	64
3.4.3.3	MDH refolding assay	64
3.4.3.4	RuBisCO refolding assay	65
3.4.3.5	Carboxylation activity assay for <i>E. coli</i> lysates	66
3.4.3.6	Proteinase K protection assay	66
3.4.3.7	Cycling of RbcL on GroEL	66
3.4.3.8	<i>In vivo</i> co-expression of RbcL or RbcLS with chaperones in <i>E. coli</i>	67

3.4.3.9	<i>In vitro</i> translation of RuBisCO	67
3.4.3.10	Co-immunoprecipitation	68
3.4.3.11	Analytical gel filtration of <i>E. coli</i> lysate or protein complexes	69
3.4.3.12	Crosslinking	69
3.4.3.13	Peptide binding assay	70
3.4.3.14	Isothermal Titration Calorimetry (ITC)	70
3.4.4	Crystallography and Structure Analysis	71
3.4.4.1	Analytical subtilisin digest of Syn7002-RbcX	71
3.4.4.2	Tryptophan-fluorescence spectroscopy	71
3.4.4.3	ANS-fluorescence spectroscopy	72
3.4.4.4	Protein crystallization	72
3.4.4.5	Structure determination	74
<b>4</b>	<b>RESULTS</b>	<b>76</b>
<b>4.1</b>	<b><i>In vitro</i> analysis on reconstitution of type I RuBisCO</b>	<b>76</b>
4.1.1	Characterization of RuBisCO as folding substrate	76
4.1.2	Structural characterization of chloroplast chaperonins and co-chaperones used for <i>in vitro</i> studies	78
4.1.3	Functional characterization of GroEL, AtCpn60 $\alpha\beta$ , AtCpn60 $\beta$ and their interaction with various co-chaperones	79
4.1.4	Type I RuBisCO cannot be refolded <i>in vitro</i>	84
4.1.5	Analysis on the impediment in reconstitution of type I RuBisCO <i>in vitro</i>	84
4.1.5.1	GroEL binds unfolded Syn6301-RbcL, thereby preventing aggregation	84
4.1.5.2	Syn6301-RbcL is encapsulated in the GroEL cage with properly cycling GroES	85
4.1.5.3	Conformational changes of Syn6301-RbcL upon interaction with the chaperonin system	87
4.1.5.4	Syn6301-RbcL is not released from GroEL as assembly-competent protein	90
4.1.5.5	Syn6301-RbcL cycles on GroEL	91
<b>4.2</b>	<b>Requirement of chaperonin for efficient production of assembled Syn6301-RuBisCO in <i>E. coli</i></b>	<b>92</b>
<b>4.3</b>	<b>Functional and structural characterization of RbcX</b>	<b>94</b>
4.3.1	Sequential operation of GroEL/GroES and RbcX in assembly of Syn7002-RuBisCO	95
4.3.2	Different requirement for RbcX in the assembly of RuBisCO from <i>Synechococcus</i> sp. PCC6301 and <i>Synechococcus</i> sp. PCC7002	98
4.3.3	Further chaperone activities of RbcX	99
4.3.4	Structural characterization of RbcX	101
4.3.5	Protein interaction surfaces on RbcX	108
4.3.6	Interaction of RbcX with the C-terminus of RbcL	113
4.3.6.1	The C-terminal RbcL motif EIKFEFD is recognized and bound by RbcX	113
4.3.6.2	Structure of the RbcX/RbcL-peptide complex	114

---

4.3.6.3	Low affinity of Syn7002-RbcX for C-terminal RbcL-peptides	115
4.3.6.4	The C-terminus (in particular its phenylalanines) is essential for interaction with RbcX and thus for proper assembly in <i>E. coli</i>	118
4.3.7	Dynamic interaction between Syn7002-RbcL and Syn7002-RbcX and its physiological importance	120
4.3.8	RbcX-mediated assembly of RbcL	126
<b>5</b>	<b>DISCUSSION</b>	<b>130</b>
<b>5.1</b>	<b>RbcL is a class III chaperonin substrate</b>	<b>131</b>
<b>5.2</b>	<b>Unproductive reconstitution of type I RuBisCO <i>in vitro</i></b>	<b>132</b>
<b>5.3</b>	<b>The RuBisCO assembly chaperone RbcX</b>	<b>134</b>
5.3.1	RbcX-mediated RuBisCO assembly	135
5.3.2	The structural foundation for RbcX-mediated RuBisCO assembly	138
5.3.3	The benefit of dynamic RbcL-RbcX interaction	140
5.3.4	Implications	142
<b>6</b>	<b>REFERENCES</b>	<b>143</b>
<b>7</b>	<b>APPENDICES</b>	<b>163</b>
<b>7.1</b>	<b>Restriction enzymes</b>	<b>163</b>
<b>7.2</b>	<b>Oligonucleotides</b>	<b>163</b>
<b>7.3</b>	<b>Expression of chloroplast chaperonins and co-chaperones</b>	<b>165</b>
<b>7.4</b>	<b>Standard amino acids</b>	<b>165</b>
<b>7.5</b>	<b>Protein accession numbers</b>	<b>166</b>
<b>7.6</b>	<b>Crystallographic data</b>	<b>167</b>
<b>7.7</b>	<b>Alignment of RbcX from green alga, moss and higher plants</b>	<b>168</b>
<b>7.8</b>	<b>Abbreviations</b>	<b>169</b>
<b>7.9</b>	<b>Publications</b>	<b>171</b>
<b>7.10</b>	<b>Lebenslauf</b>	<b>172</b>

## 1 SUMMARY

To become biologically active, proteins have to acquire their correct three-dimensional structure by folding, which is frequently followed by assembly into oligomeric complexes. Although all structure relevant information is contained in the amino acid sequence of a polypeptide, numerous proteins require the assistance of molecular chaperones which prevent the aggregation and promote the efficient folding and/or assembly of newly-synthesized proteins. The enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), which catalyzes carbon fixation in the Calvin-Benson-Bassham cycle, requires chaperones in order to acquire its active structure. In plants and cyanobacteria, RuBisCO (type I) is a complex of approximately 550 kDa composed of eight large (RbcL) and eight small (RbcS) subunits. Remarkably, despite the high abundance and importance of this enzyme, the characteristics and requirements for its folding and assembly pathway are only partly understood. It is known that folding of RbcL is accomplished by chaperonin and most likely supported by the Hsp70 system, whereas recent findings indicate the additional need of specific chaperones for assembly. Nevertheless, this knowledge is incomplete, reflected by the fact that *in vitro* reconstitution of hexadecameric RuBisCO or synthesis of functional plant RuBisCO in *E. coli* has not been accomplished thus far.

In this thesis, attempts to reconstitute type I RuBisCO *in vitro* did not result in production of active enzyme although a variety of reaction conditions and additives as well as chaperones of different kind, origin and combination were applied. The major obstacle for reconstitution was found to be the incapability to produce RbcL<sub>8</sub> cores competent to form RbcL<sub>8</sub>S<sub>8</sub> holoenzyme. It could be shown that the RbcL subunits interact properly with the chaperonin GroEL in terms of binding, encapsulation and cycling. However, they are not released from GroEL in an assembly-competent state, leading to the conclusion that a yet undefined condition or (assembly) factor is required to shift the reaction equilibrium from GroEL-bound RbcL to properly folded and released RbcL assembling to RbcL<sub>8</sub> and RbcL<sub>8</sub>S<sub>8</sub>, respectively.

Cyanobacterial RbcX was found to promote the production of cyanobacterial RbcL<sub>8</sub> core complexes downstream of chaperonin-assisted RbcL folding, both in *E. coli* and in an *in vitro* translation system. Structural and functional analysis defined RbcX as a

homodimeric, arc-shaped complex of approximately 30 kDa, which interacts with RbcL *via* two distinct but cooperating binding regions. A central hydrophobic groove recognizes and binds a specific motif in the exposed C-terminus of unassembled RbcL, thereby preventing the latter from uncontrolled misassembly and establishing further contacts with the polar peripheral surface of RbcX. These interactions allow optimal positioning and interconnection of the RbcL subunits, resulting in efficient assembly of RbcL<sub>8</sub> core complexes. As a result of the highly dynamic RbcL-RbcX interaction, RbcS can displace RbcX from the core-complexes to produce active RbcL<sub>8</sub>S<sub>8</sub> holoenzyme. Species-specific co-evolution of RbcX with RbcL and RbcS accounts for limited interspecies exchangeability of RbcX and for RbcX-supported or -dependent assembly modes, respectively.

In summary, this study helped to specify the problem causing prevention of proper *in vitro* reconstitution of type I RuBisCO. Moreover, the structural and mechanistic properties of RbcX were analyzed, demonstrating its function as specific assembly chaperone for cyanobacterial RuBisCO. Since the latter is very similar to RuBisCO of higher plants, this work may not only augment the general understanding of type I RuBisCO synthesis, but it might also contribute to advancing the engineering of catalytically more efficient crop plant RuBisCO both in heterologous systems and *in planta*.

## 2 INTRODUCTION

Proteins are organic macromolecules involved in virtually every aspect of the biochemistry and physiology of living organisms. They play key roles in metabolism, transport, immune response and signal transduction, in which they can serve as structural elements, catalysts, adaptors, messengers, transporters or regulators.

### 2.1 Protein structure

A protein's biological function derives from its unique overall structure, whose organization can be divided into four different levels: primary, secondary, tertiary and quaternary structure.

As macromolecules, proteins consist of one or more polypeptides, which in turn are linear, unbranched polymers of up to 20 different covalently bonded amino acids. The sequence of amino acids in a polypeptide is referred to as its primary structure and is determined by the gene that encodes the respective protein.

The secondary structure of a polypeptide describes the local three-dimensional arrangement of its backbone atoms. This arrangement is primarily stabilized by hydrogen bonding and stereochemically restricted by the side chains. The most commonly observed secondary structural elements are the  $\alpha$ -helix,  $\beta$ -sheet and  $\beta$ -turn.

A polypeptide's tertiary structure refers to its exact three-dimensional structure, regarding the relative spatial positioning of all the secondary structure elements, their side chains and atoms. By this tertiary arrangement, residues separated in the primary structure can be brought in close proximity in the folded protein. In soluble proteins, hydrophobic residues usually become buried inside, forming a hydrophobic core, whereas polar and charged sidegroups frequently point outward, interacting with the solvent, substrates or participating in surface salt bridges.

Large proteins often consist of several folded polypeptide chains forming distinct structural and often functionally discernable subunits or domains. The spatial arrangement of these subunits in the complex represents the quaternary structure. Normally, weak hydrophobic, charged or polar interactions are the stabilizing forces between the subunits.

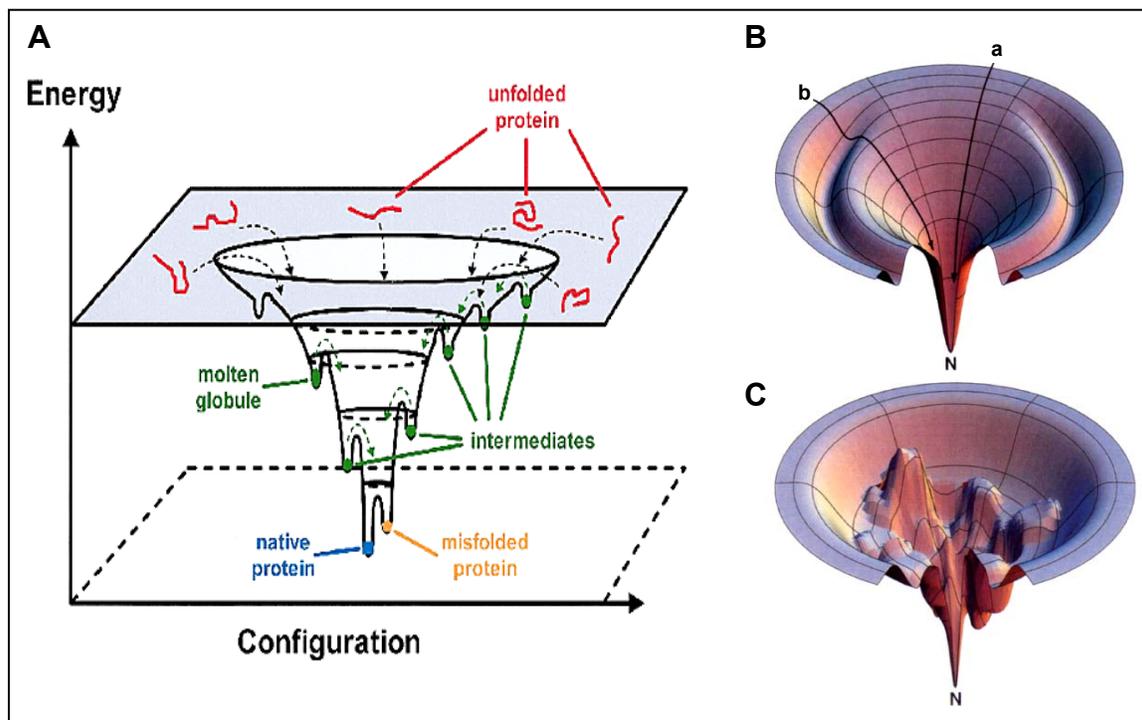
In general, the conformational stability of proteins can be provided by hydrophobic interactions, electrostatic interactions (van der Waals forces, hydrogen bonds, ionic interactions), covalent linkages (disulphide bridges) or by coordination of metals.

## 2.2 Protein folding

The folding of a protein is a complex process, which is based on a multitude of non-covalent interactions among the amino acids of the folding polypeptide resulting in the formation of the native protein.

As demonstrated by Christian Anfinsen, some proteins can adopt their native functional structure spontaneously in free solution. This observation suggests that all structure relevant information is contained in the amino acid sequence of a polypeptide (Anfinsen 1973). The possibility that proteins undergo a multitude of random searches to find their native conformation was led *ad absurdum* by the Levinthal paradox, which regards the enormous time consumption of random trials (Levinthal 1969). In fact, the considerable speed of folding reactions (with time constants far below one second for some proteins) implies the existence of directed and kinetically driven folding pathways. Whereas some proteins have a single folding roadmap, others may reach their final conformation *via* multiple pathways. Thermodynamically, a protein folds from a higher energy unfolded state to a lower energy folded (native) state, which is usually the most stable state under biological conditions. The energy landscape of a folding process can be described as a funnel. The roughness of the funnel surface and thus the formation of the native state can vary considerably between different proteins (Fig. 1) (Dill and Chan 1997, Schultz 2000). The major driving force in protein folding is considered to be the hydrophobic effect, which forces hydrophobic amino acids into the interior of a protein. At the beginning of folding events, local amino acid sequences initiate folding, followed by the rapid formation of transient, partially folded and more compact intermediates. Such an intermediate can be the more stabilized “molten globule” state, which shows already most elements of the secondary structure, but lacks tertiary interactions. Subsequently, in a stepwise manner and at a slower pace, higher structure elements are formed until the native state is accomplished. A folding process that follows a pathway like this, reduces the amount of possible conformations and thus allows folding to occur within a

biologically relevant period of time (Baldwin 1995, Baldwin and Rose 1999, Dill and Chan 1997, Dobson 1995, Privalov 1996).



**Figure 1. Folding energy landscapes.**

(A) The multiple states of the unfolded protein located at the top fall into a folding funnel consisting of an almost infinite number of local energy minima, each of which describes possible folding arrangements in the protein. Most of these states represent transient folding intermediates in the process of attaining the correct native fold. Some of these intermediates retain a more stable structure such as the molten globule, whereas other local minima act as folding traps, irreversibly capturing the protein in a misfolded state (Schultz 2000). (B) Moat landscape for a fast-folding throughwary process (a) and a slow-folding process with a kinetic trap (b); (C) Rugged energy landscape with kinetic traps, energy barriers and some narrow throughwary paths; N is the native state with a global thermodynamic energy minimum (Dill and Chan 1997).

### 2.3 Protein folding and aggregation *in vivo*

In contrast to *in vitro* conditions, where especially small and simply-structured proteins can accomplish folding spontaneously, protein folding in the cell is often more intricate and has to compete with numerous side reactions.

The interior of a cell is not only characterized by pH, ionic strength, redox potential or temperature, but also by the crowding with macromolecules (proteins, nucleic acids, polysaccharides etc.), that occupy approximately 20-30 % of the total cell volume. This occupied fraction is physically unavailable to other macromolecules, wherefore

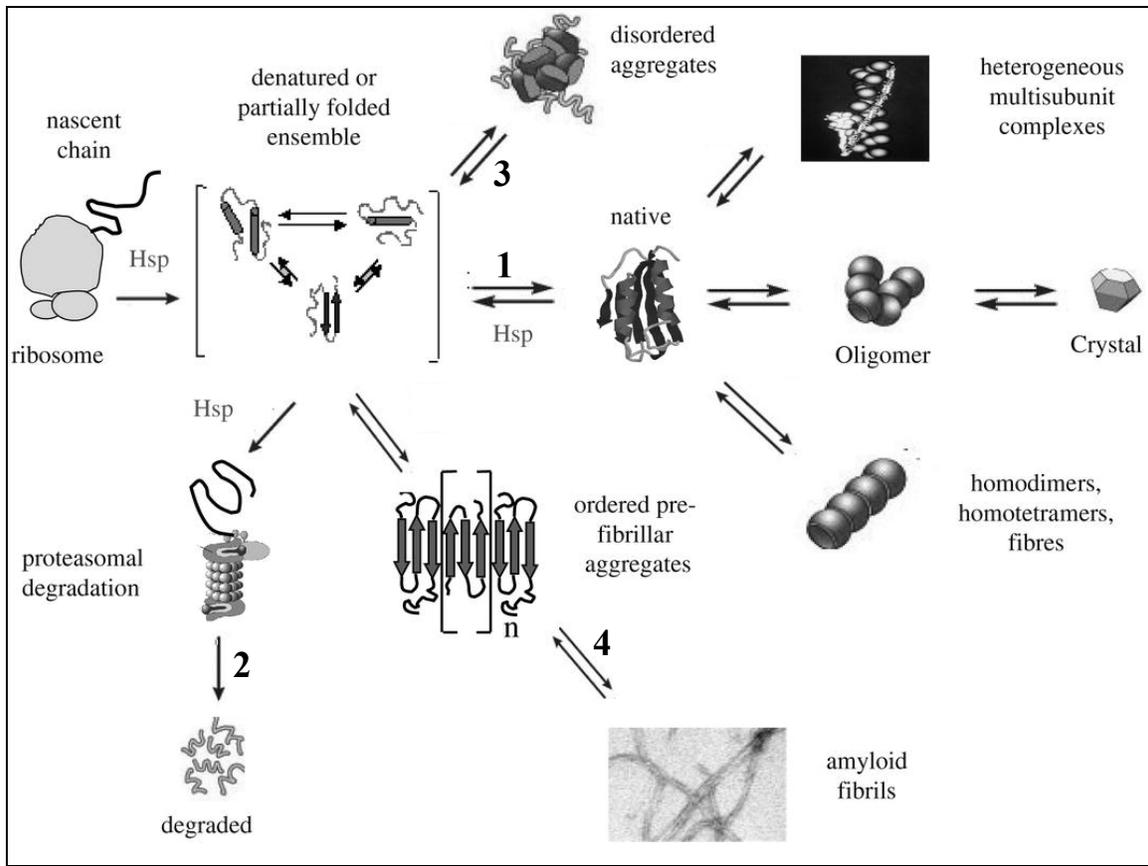
molecular crowding is also referred to as the excluded volume effect (Ellis 2001). Crowding constitutes a nonspecific force for protein association, folding and aggregation, causing small structures (e.g. compact aggregates or the globular native state) to be favored above large structures (e.g. unfolded or partially folded chains). In a crowded milieu it is more likely that structurally promiscuous protein species diffuse into one another and aggregate.

Aggregation is one of the common consequences of a polypeptide failing to reach or maintain its functional native structure. Enhanced by macromolecular crowding, aggregation severely competes with productive protein folding in the cytosolic environment. This problem is of particular relevance during *de novo* protein synthesis. Since translation by ribosomes proceeds at a relatively low speed and in a vectorial manner, the elongating polypeptide chains are not available for folding at once. The chains emerging at the ribosomal exit tunnel, transiently expose hydrophobic residues, which later become hidden in the interior of the native protein. However, as long as folding is not completed, they are prone to trigger aggregation upon intra- and intermolecular hydrophobic interactions.

Throughout the folding process, contacts between amino acid patches that are separate in the native state, can result in intermediate conformations that might expose hydrophobic side chains and unstructured backbone regions. Consequently, upon hydrophobic interactions and interchain hydrogen bonding, respective intermediates can also undergo misfolding and (dis)ordered aggregation (Dobson 2003, Hartl and Hayer-Hartl 2002).

Not only *de novo* synthesis produces partially folded molecules. Destabilizing conditions (thermal stress, pH etc.), specific mutations, misprocessing or chemical modification of a protein can result in non-native structures, that expose hydrophobic patches and therefore tend to aggregate, as well (Stefani and Dobson 2003).

The equilibrium between the partially folded polypeptides and the natively folded proteins is usually strongly in favor of the latter, but if it is shifted to the former, formation of aggregates and/or degradation become likely (Fig. 2).



**Figure 2. States accessible to a polypeptide chain following its biosynthesis.**

Proteins fold from their highly disordered unfolded state through partially structured intermediates to a globular native state (1). The latter might assemble into functional oligomeric complexes or/and form aggregated species, the most ordered of which are three-dimensional crystals. Enhanced by macromolecular crowding, unfolded or partially folded polypeptides might be degraded by the proteasome (2) or they can form highly disordered aggregates (3) or amyloid fibril formation occurs by a nucleation and growth mechanism (4) (Stefani and Dobson 2003).

Several human diseases, such as Alzheimer's, Parkinson's, Huntington's or Creutzfeldt-Jakob disease, can be caused by proteins that adopt non-native conformations, resulting in misfolding and deposition of ordered fibrillar aggregates, so-called amyloids. The latter display characteristic "cross- $\beta$ " structures in X-ray diffraction analyses in addition to a typical fibril morphology in electron micrographs and they are remarkably stable to degradation (Dobson 2003, Lansbury 1999, Perutz 1999, Prusiner 1997). However, the ability to form fibrils is not restricted to disease related proteins, but instead it appears to be a property of many proteins under appropriate conditions (Chiti *et al.* 1999).

As aggregation is often an irreversible process leading to a loss of proteins, the need of cellular protection mechanisms is obvious.

## 2.4 Molecular chaperones

Cells provide a sophisticated machinery of proteins which play a fundamental role in preventing the aggregation and in assisting the folding and assembly of proteins: the molecular chaperones. A variety of molecular chaperones have been found and characterized up to date. Based on differences regarding size, structure, function, mechanism or cellular compartmentalization, different families of chaperones can be classified: Hsp100/Clp, Hsp90, Hsp70, chaperonins (Hsp60), small Hsps and calnexin/calreticulin (Miernyk 1999, Walter and Buchner 2002).

A common property of molecular chaperones is their ability to recognize and to reversibly bind to exposed aggregation prone hydrophobic residues and unstructured backbone patches of non- or partially folded polypeptides that emerge at the ribosome exit tunnel during translation or have to remain unfolded during transport processes or result from unfolding events caused by stress conditions. The chaperones stabilize these non-native proteins, they can prevent them from misfolding or aggregation and direct them into productive folding and assembly pathways. If chaperones assist folding, they are not themselves components of the final structure and they do not impose structural information directing the folding of substrates. Their potential to prevent aggregation and to mediate protein folding is particularly relevant for *de novo* protein biosynthesis, but also for the reduction of damage under stress conditions, such as elevated temperatures. The latter usually result in elevated quantities of many chaperones, accounting for their designation as “stress proteins” or “Hsps/heat shock proteins”.

But molecular chaperones play also an essential role in intercompartmental protein transport processes, in the resolubilization and refolding of aggregated proteins or they can transfer non-native, damaged or irreversibly misfolded proteins to the cells' degradation machinery. In selected cases, chaperones are also involved in controlling the activity of regulatory proteins, in modulating cellular signal transduction or in processing antigens in immunological pathways (Ben-Zvi and Goloubinoff 2001, Bukau *et al.* 2006, Hartl 1996, Jackson-Constan *et al.* 2001, Kunisawa and Shastri 2006, Miernyk 1999, Young *et al.* 2001). Notably, several specific chaperones have been found to be required for the assembly of some oligomeric proteins, such as nucleosomes or the mammalian proteasome (Ellis 2006).

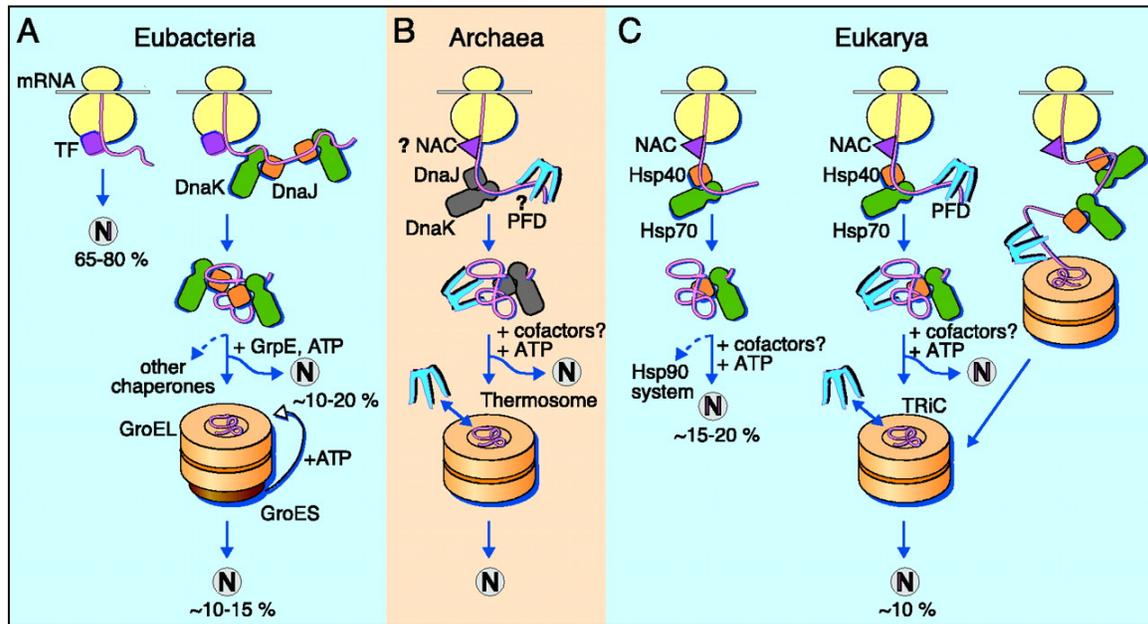
Below, some representatives of molecular chaperones will be characterized both structurally and functionally in more detail.

### **2.4.1 Chaperones involved in *de novo* protein folding**

Cells provide a complex network of molecular chaperones that assist proteins to adopt their functional three-dimensional structures. In general, there are two major classes of cytosolic chaperones mediating the folding of newly synthesized polypeptides in a sequential manner (Fig. 3) (Ellis and Hartl 1999, Hartl 1996, Hartl and Hayer-Hartl 2002, Netzer and Hartl 1998, Young *et al.* 2004).

The first class includes chaperones that bind directly to nascent polypeptide chains. Thus, their main role is to maintain the folding competence of the translating polypeptide chains during translation by shielding hydrophobic regions, which might otherwise become involved in inappropriate intra- or intermolecular interactions, leading to misfolding and aggregation. Some representatives of this class bind directly to the ribosome and start to interact with the nascent chains as soon as they exit the ribosome peptide tunnel, for example trigger factor (TF) or the nascent chain associated complex (NAC). Other members of this class also interact with the nascent chain, but without binding to the ribosome, e.g. Hsp70. A considerable proportion of mainly small newly translated proteins appears to fold rapidly and unassisted upon release from the ribosome-bound chaperones. However, numerous longer polypeptides require further chaperone assistance for stabilization and folding. A domain-wise co-translational folding mechanism has been suggested for TF. In addition, the Hsp70 system binds and stabilizes the elongating chains, but it can also support the co- or post-translational folding of some proteins by repeated binding/release cycles.

In some cases, non-native proteins have to be transferred to the second, more downstream class of chaperones in the *de novo* folding pathway: the cylindrical chaperonin complexes (GroEL/GroES, Thermosome, TRiC). The latter provide a sequestered environment in which post-translational folding can proceed unimpaired by intermolecular interactions between non-native polypeptides. Alternatively, some eukaryotic polypeptides need to be passed from the Hsp70 system to members of the Hsp90 family, which assist their folding.



**Figure 3. Model for chaperone-assisted *de novo* protein folding in the bacterial, archaeal and eukaryotic cytosol.**

(A) In the bacterial cytosol, the majority (65-80 %) of all newly synthesized polypeptides, predominantly small proteins, fold fast and spontaneously upon release from the ribosome and the ribosome-bound TF. The remainder (10-20 %) of mainly larger proteins accomplishes folding upon repeated binding/release cycles by DnaK/DnaJ and 10-15 % are subsequently transferred to the GroEL/ES chaperonin machinery. (B) Only some archaea contain DnaK/DnaJ. The existence of NAC and the interaction of nascent chains with PFD have not yet been experimentally confirmed. Chaperonin-mediated folding is accomplished by thermosomes. (C) In the eukaryotic cytosol, the majority of newly translated proteins folds upon release from NAC or upon subsequent interaction with Hsp70/40 and partially Hsp90. Approximately 10 % of proteins associate with PFD to be transferred to the chaperonin TRiC. N: native protein, TF: trigger factor, NAC: nascent chain-associated complex, PFD: prefoldin (Hartl and Hayer-Hartl 2002).

### 2.4.2 Ribosome-associated chaperones

The first molecular chaperones that interact with newly translated polypeptides emerging from the ribosome exit tunnel, are trigger factor (TF) in bacteria or the nascent chain associated complex (NAC) in archaea and eukaryotes (Deuerling and Bukau 2004). TF and NAC do not only bind to the nascent peptide chains, but they are also associated with the ribosome. Trigger factor is a 48 kDa protein in the cytosol of *E. coli*, with peptidyl-prolyl *cis/trans* isomerase activity (PPIase) and chaperone-like function (Hesterkamp and Bukau 1996). It can be found both free in the cytosol and interacting with the two adjacent proteins L23 and L29, located at the exit of the ribosome peptide tunnel. From the latter position, TF can co-translationally establish contacts to stretches of the growing translated polypeptides that are enriched with hydrophobic residues (Ferbitz *et al.* 2004,

Kramer *et al.* 2002). TF interacts with its substrates in a nucleotide-independent binding reaction and stabilizes them in a folding competent state, thus preventing their aggregation (Kaiser *et al.* 2006, Nishihara *et al.* 2000). More recent structural and biochemical findings suggest that TF even provides a hydrophobic cradle as both binding site and shielded folding environment for nascent chains which emerge from the ribosome exit tunnel (Deuerling and Bukau 2004, Ferbitz *et al.* 2004). The synthetic lethality combined with retarded folding and increased aggregation, observed when the genes for both Hsp70 (DnaK) and TF are deleted in *E. coli*, results from functional parallels between the two of them (Deuerling *et al.* 1999, Teter *et al.* 1999). Additionally, it has been shown that the cooperation of TF and DnaK can increase the folding yields of multidomain proteins by arresting the folding of single domains and allowing folding to occur delayed upon completion of translation (Agashe *et al.* 2004).

### 2.4.3 The Hsp70 system

Hsp70 chaperones and their co-factors comprise a group of abundant cellular devices that fulfill a variety of functions in almost all cellular compartments.

Under stress conditions, they prevent aggregation and support refolding of denatured or misfolded proteins. Under moderate conditions, Hsp70 is constitutively essential in many ways as well (Bukau and Horwich 1998, Hartl 1996, Hartl and Hayer-Hartl 2002, Walter and Buchner 2002). For example, it can bind to, shield and stabilize newly synthesized nascent polypeptide chains during *de novo* protein biosynthesis. Throughout the time of association, the substrates are prevented from aggregation and kept in a folding competent state. In this respect, Hsp70 functionally overlaps with trigger factor (TF), although it does not appear to bind directly to the ribosome, but rather to act on the nascent chains subsequent to TF. Consequently, a double deletion of both TF and Hsp70 (DnaK) in *E. coli* is conditionally lethal and enhances aggregation of many large newly translated polypeptides (Deuerling *et al.* 1999, Teter *et al.* 1999). By repeated binding/release cycles, several newly translated proteins can even become folded co- or post-translationally upon interaction with the Hsp70 chaperones.

Moreover, Hsp70s are often found to be involved in cellular processes, in which they are part of chaperone networks. For example, the sequential action of Hsp70 and chaperonins

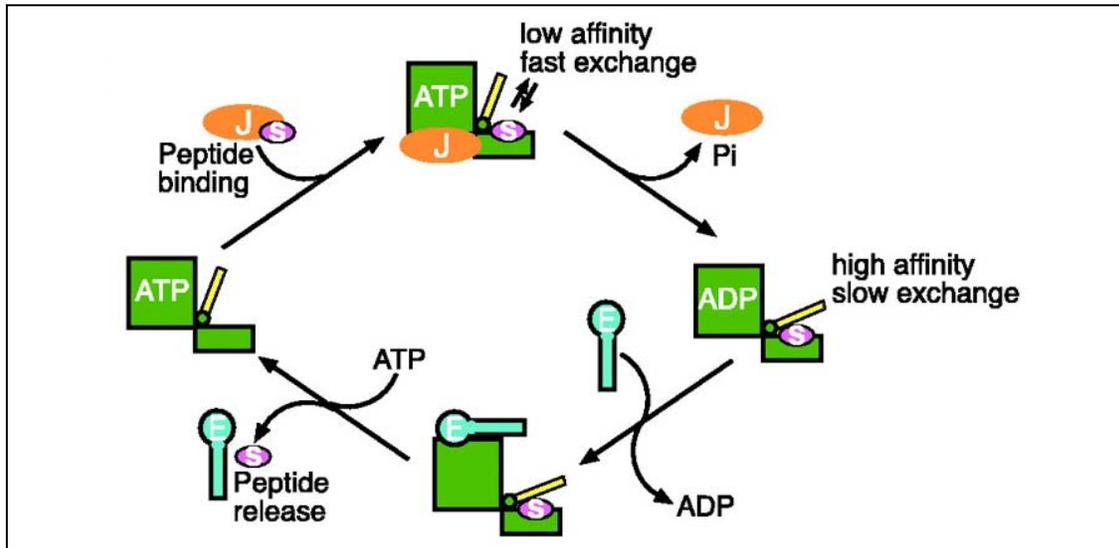
is essential in the folding pathway and for aggregation prevention of numerous proteins (Agashe and Hartl 2000, Frydman *et al.* 1994, Langer *et al.* 1992b). The cooperation of the Hsp70 system and the Hsp90 system has been found to be important for folding events or for the activation of steroid hormone receptors (Young *et al.* 2001). In concert with the Hsp100/Clp family, the Hsp70 system is implicated in pathways mediating protein degradation (Bukau *et al.* 2006, Wickner *et al.* 1999) or the resolubilization and refolding of aggregated proteins (Ben-Zvi and Goloubinoff 2001, Glover and Lindquist 1998). Additionally, a complex set of chaperones including the Hsp70 system is interacting in a multitude of protein translocation processes (Jackson-Constan *et al.* 2001, Neupert 1997).

The most extensively studied representative of Hsp70 systems is the DnaK/DnaJ/GrpE ensemble of *E. coli*, in which DnaJ (chaperone-activating protein) and GrpE (nucleotide exchange factor) are required as co-factors for the functional regulation of DnaK. DnaK itself is composed of an N-terminal ATPase domain and a C-terminal peptide binding domain. The latter binds preferentially to short hydrophobic peptide sequences of unfolded proteins, thereby preventing them from aggregation. The nature of bound nucleotide determines the peptide binding properties of DnaK and ATP hydrolysis serves as a molecular switch between the highly dynamic open ATP state (low substrate affinity) and the poorly dynamic ADP state (high substrate affinity).

In the DnaK reaction cycle an unfolded substrate initially binds to the co-factor DnaJ, before it is transferred to the open ATP form of DnaK. Stimulated by DnaJ and the substrate protein, ATP is hydrolyzed by DnaK, resulting in the closure of the binding site and the locking-in of the substrate. Subsequently, DnaJ leaves this stable DnaK/substrate complex and GrpE associates with DnaK to trigger nucleotide exchange, i.e. the release of ADP and binding of ATP. The concomitant release of GrpE induces a conformational change in DnaK, the peptide binding site opens and the substrate can dissociate (Fig. 4).

Unlike GroEL, Hsp70 does not provide a folding cavity, but it appears to function by ATP-dependent binding/release cycles of non-native polypeptides, that were generated by *de novo* synthesis or by unfolding processes. Hsp70 does not appear to induce global conformational changes in the substrates, but it rather acts locally. After release from Hsp70, some substrates might accomplish folding spontaneously. In contrast, certain

slow-folding and/or multidomain proteins might require repeated binding to Hsp70 in order to prevent intramolecular misfolding and hence to promote proper folding to the native state. Other proteins need to be transferred from Hsp70 to other chaperone systems that mediate further folding, transport or degradation – as described above (Bukau and Horwich 1998, Hartl and Hayer-Hartl 2002).



**Figure 4. Hsp70 reaction cycle exemplified by the *E. coli* DnaK/DnaJ/GrpE system.**

The non-native protein or elongating polypeptide chain at the ribosome exit tunnel is initially bound to DnaJ (orange) and subsequently transferred to ATP-binding DnaK (green). DnaJ and the substrate stimulate ATP hydrolysis by DnaK, resulting in locking of the substrate in the binding site. Dissociation of DnaJ is followed by binding of GrpE (blue), which triggers ADP release and ATP binding. Upon release of GrpE, the peptide binding site opens and the substrate can dissociate either to fold spontaneously or to rebind to the chaperone system (Hartl and Hayer-Hartl 2002).

#### 2.4.4 Prefoldin/GimC

Prefoldin (PFD) represents an additional chaperone in the folding pathway of newly translated polypeptides in archaea and in the eukaryotic cytosol.

PFD is a 90 kDa jellyfish-shaped complex, which is composed of two subunit types,  $\alpha$  and  $\beta$ , with a stoichiometry of  $\alpha_2\beta_4$  (Leroux *et al.* 1999, Siegert *et al.* 2000). By binding to non-native proteins *via* hydrophobic interactions, PFD allows ATP-independent stabilization of intermediates prior to folding by downstream chaperonins. In this respect, it appears to function similarly to TF or DnaK. Since eukaryotic PFD has been found to bind to newly synthesized actin and tubulin, it is also referred to as Gim complex (genes involved in microtubule biogenesis) (Siegers *et al.* 1999, Vainberg *et al.* 1998).

### 2.4.5 The chaperonins

The chaperonins are a highly conserved class of molecular chaperones, which are abundant and essential in all prokaryotic and eukaryotic organisms and increase to varying extents after stress. They usually form large multimeric double-ring complexes (800-1000 kDa) with ATPase activity and two integrated folding cavities. Newly translated, imported or stress-denatured non-native substrate proteins are bound by the chaperonins *via* hydrophobic interactions and become subsequently encapsulated inside the central cavity, in which the polypeptides can adopt native conformations while they are well protected from aggregation. Based on sequence and mechanistic deviations, the chaperonins are classified into two subgroups.

Group I chaperonins occur in eubacteria and eukaryotic organelles with endosymbiotic origin, i.e. mitochondria and plastids (Viale and Arakaki 1994). They are composed of 60 kDa subunits, resulting in their designation as Hsp60s, and they execute their ATP-dependent activity in concert with co-factors of the GroES or Hsp10 family. A more detailed structural and functional characterization of bacterial and chloroplast group I chaperonins is given below.

Representatives of group II chaperonins are TRiC (TCP-1 ring complex, also called CCT for chaperonin-containing TCP-1) in the eukaryotic cytosol or the thermosome in *Archaea*. They are composed of different but homologous subunits of 50-60 kDa, which assemble to complexes of two stacked octameric rings that function in an ATP-dependent manner without a GroES-like co-factor (Kubota *et al.* 1995, Waldmann *et al.* 1995). Unlike GroEL, their apical domains appear to have an  $\alpha$ -helical protrusion, which is assumed to operate as lid of the folding cavity and accounts for the independence from a co-factor (Klumpp *et al.* 1997). TRiC has been shown to interact transiently with about 9-15 % of newly synthesized proteins of 30-120 kDa, among them the cytoskeletal proteins actin and tubulin as major substrates (Thulasiraman *et al.* 1999).

#### 2.4.5.1 Structure and function of the *E. coli* chaperonin system

GroEL and GroES from *E. coli* represent one of the best characterized, physiologically indispensable chaperone system. Electron microscopy (Chen *et al.* 1994, Langer *et al.* 1992a, Ranson *et al.* 2001, Roseman *et al.* 1996 and 2001), the crystal structures of e.g.

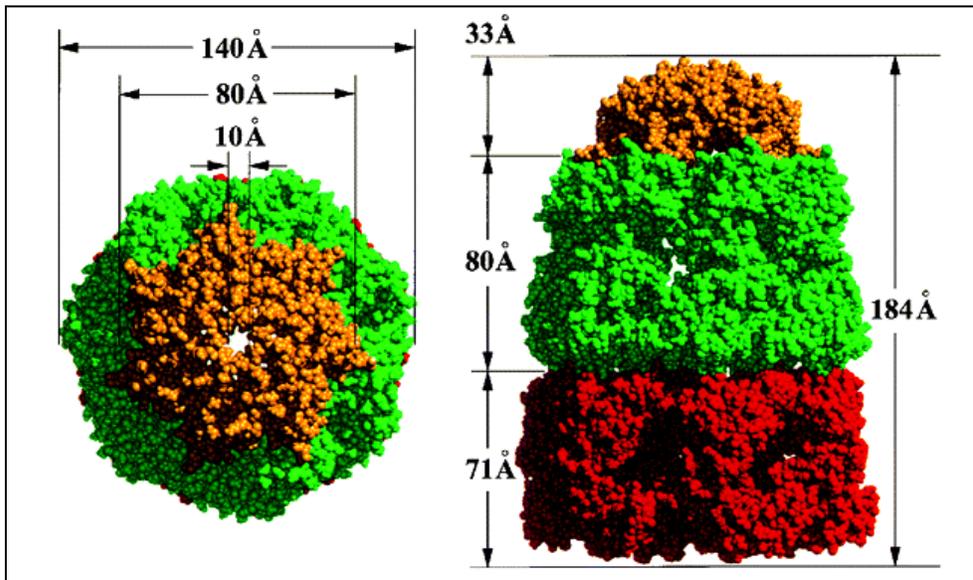
GroEL, ATP-bound GroEL or GroEL-GroES-(ADP)<sub>7</sub> (Boisvert *et al.* 1996, Braig *et al.* 1994, Xu *et al.* 1997) as well as numerous *in vivo* and *in vitro* functional analyses gave insight into structural, mechanistic and dynamic features of the GroEL/GroES system and the folding processes mediated by them (reviewed in Bukau and Horwich 1998, Hartl and Hayer-Hartl 2002, Sigler *et al.* 1998).

The chaperonin GroEL is a large cylindrical complex of about 800 kDa. It is composed of fourteen identical 57 kDa subunits, which form two heptameric rings stacked back-to-back and each of them enclosing a central cavity. Each GroEL subunit is composed of three structurally distinct domains: an apical, an intermediate and an equatorial domain (Fig. 5 and 6). The apical domain surrounds the opening ends of the central cavity and it exposes various hydrophobic residues towards the central cavity for binding of unfolded substrate protein and for subsequent association with GroES. The equatorial domain contributes most of the residues that form the nucleotide binding and ATPase sites and it establishes the majority of intra- and inter-ring contacts within the GroEL oligomer. The hingelike intermediate domain connects apical and equatorial domain and conveys exchange of allosteric information between them (Sigler *et al.* 1998).

GroES, the co-factor of GroEL, is a heptameric dome-shaped single ring of identical 10 kDa subunits, which can bind to the apical domains of GroEL, resulting in an asymmetric, bullet-shaped GroEL/GroES complex and encapsulation of the non-native substrate protein in the folding *cis* cavity (Fig. 5-7) (Hunt *et al.* 1996, Xu *et al.* 1997). The opposing ring is referred to as the *trans* cavity. With its so-called “mobile loops”, GroES can interact partially with the hydrophobic substrate binding sites of the apical domains of adenine nucleotide-bound GroEL (Landry *et al.* 1993). The transient interaction between GroEL and GroES is dynamically regulated by the GroEL ATPase activity. Negative allostery between the two GroEL rings causes binding of ATP and thus also of GroES preferentially to only one ring, but football-like structures with GroES on both ends have been observed as well (Schmidt *et al.* 1994).

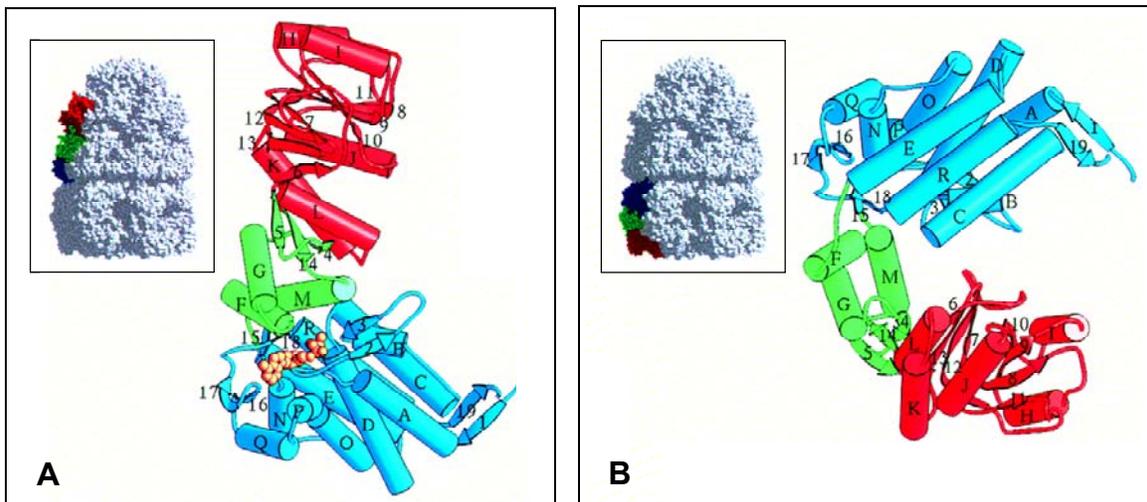
GroEL can prevent aggregation and accomplish folding of many different unfolded proteins *in vitro* (Viitanen *et al.* 1992). However, *in vivo*, where a number of proteins are folded by upstream chaperones, interaction analyses of GroEL with the *E. coli* proteome revealed a limited set of GroEL substrates, which represent only about 10-15 % of the

cytosolic proteins and which can be divided into three different classes. Class I substrates are non-stringent and class II substrates can alternatively be folded by other chaperones than GroEL, for example by the KJE-system. Only for folding of class III substrates, which often display  $(\alpha\beta)_8$  TIM barrel folds, GroEL is essential (Ewalt *et al.* 1997, Houry *et al.* 1999, Kerner *et al.* 2005)



**Figure 5. The asymmetric GroEL/GroES complex.**

Space filling model with the complex dimensions in a top view (left) and side view (right). The GroEL rings are colored in green and red, GroES is shown in yellow (Xu *et al.* 1997).



**Figure 6. Structural rearrangements in GroEL upon binding of GroES**

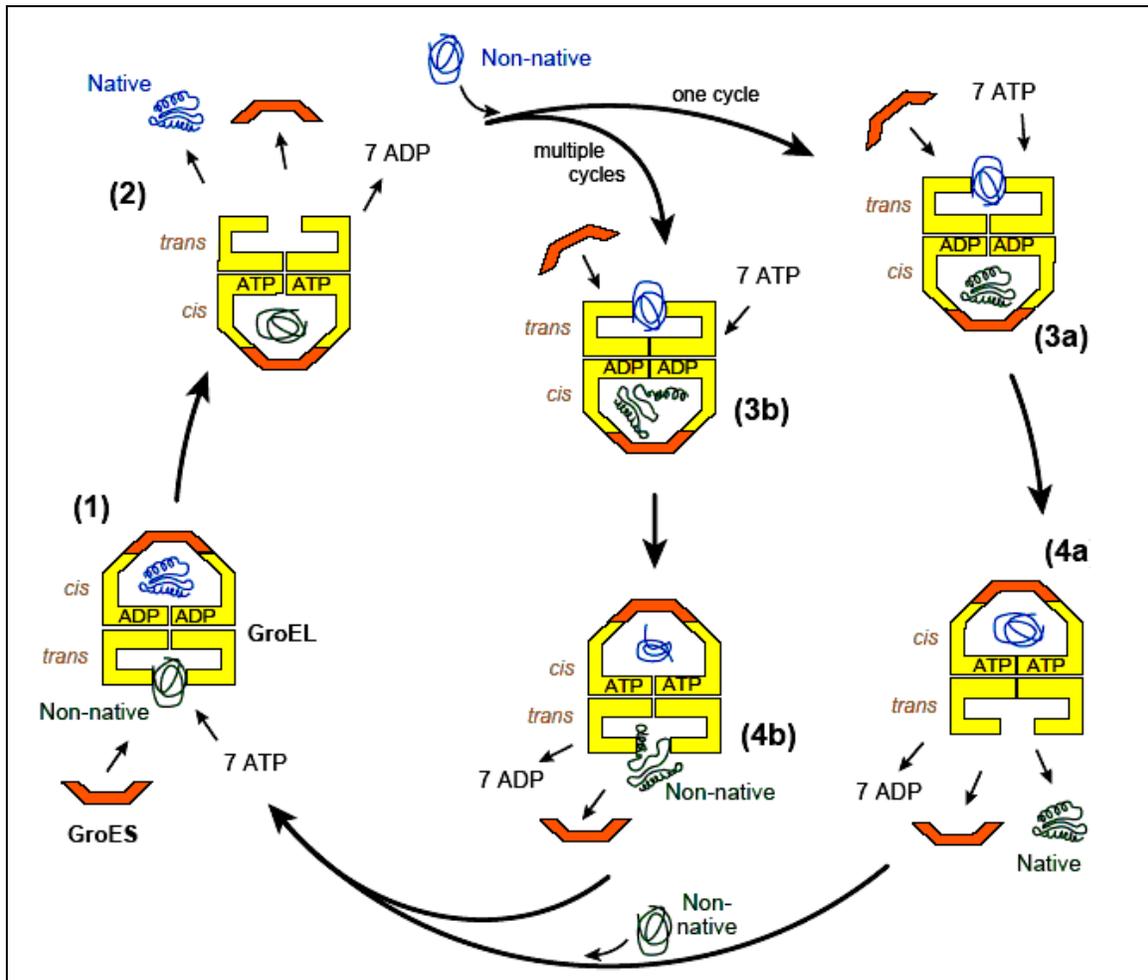
Small space filling models of the GroEL/GroES complex and enlarged ribbon diagrams of a single GroEL subunit belonging either to a GroEL ring that binds GroES (A) or that does not bind GroES (B). The equatorial domain of GroEL is depicted in blue, the intermediate domain in green and the apical domain in red (Xu *et al.* 1997).

#### 2.4.5.2 Mechanism of GroEL/ES mediated protein folding

Folding of non-native polypeptides is accomplished alternately by the two GroEL rings, in cooperation with GroES and upon hydrolysis of ATP (Fig. 7). At the beginning of a reaction cycle, a single unfolded substrate protein or folding intermediate binds to the apical domains of one GroEL ring *via* hydrophobic interactions (Chatellier *et al.* 1999, Chen and Sigler 1999, Fenton *et al.* 1994). Cooperative binding of seven ATP molecules to the equatorial domains of this *cis* ring causes rapid conformational changes, i.e. the apical domains open out and twist upward, resulting in an elongation of the GroEL cylinder (Boisvert *et al.* 1996, Roseman *et al.* 1996). Subsequently, the *cis* apical subunits can bind GroES and release the non-native substrate into the central cavity of the now bullet-shaped asymmetrical complex (Langer *et al.* 1992a, Mayhew *et al.* 1996, Weissman *et al.* 1996). Due to additional structural rearrangements in both the apical and the intermediate domains, the *cis* cavity becomes twofold enlarged and hydrophobic residues on the inner cavity walls are buried and replaced by mostly polar residues (Chen *et al.* 1994, Roseman *et al.* 1996, Xu *et al.* 1997). In this environment, polypeptides can initiate and optimally complete folding to their native state in the *cis* complex. This process takes place in about 10-15 sec and is accompanied by the hydrolysis of ATP (Hayer-Hartl *et al.* 1995, Rye *et al.* 1997). In the meantime, binding of ATP and GroES to the *cis* ring have induced the dissociation of ADP, GroES and previously folded substrate protein from the opposite *trans* ring. Finally, binding of a new non-native substrate, ATP and GroES to the *trans* ring triggers the opening of the *cis* cage and thus the release of ADP, GroES and the respective folding intermediate or product (Horovitz *et al.* 2001). Intermediates can rebind to GroEL for accomplishment of folding in one or more subsequent cycles (Mayhew *et al.* 1996).

The GroEL/ES folding cage (also called “Anfinsen cage”) can encapsulate polypeptides up to approximately 60 kDa and provides a hydrophilic milieu, which prevents interactions of the substrate with other aggregation prone molecules in the highly crowded intracellular environment. Moreover, spatial confinement in the cavity or forced unfolding-folding-cycles might promote or accelerate folding reactions and may smooth the energy landscape throughout folding pathways, allowing even polypeptides, which are kinetically trapped in incorrect conformations, to fold or refold correctly (Brinker *et*

*al.* 2001, Hartl and Hayer-Hartl 2002, Shtilerman *et al.* 1999, Tang *et al.* 2006, Todd *et al.* 1996). The steric information for the correct folding of a protein is encoded in its primary amino acid sequence (Anfinsen 1973). With regard to this, the role of GroEL is not to provide steric information and to actively fold its substrates, but it rather offers an appropriate isolated environment and efficient support for the folding process.



**Figure 7. Model for the GroEL/GroES reaction cycle.**

The two GroEL rings support protein folding in an alternating fashion and in cooperation with GroES. (1) Non-native substrate binds *via* hydrophobic interactions to the apical domains of the unoccupied GroEL ring. (2) Cooperative binding of 7 ATP to this *cis* ring induces conformational changes. As a consequence, the apical domains open out and twist upward, resulting in an elongation of the *cis* GroEL cylinder. GroES can bind to the apical subunits and the substrate is released into the central cavity of the *cis* ring. Concomitantly, the *trans* ring releases bound GroES, ADP and the formerly folded substrate protein. (3a,b) The substrate protein can fold inside the *cis* cavity, accompanied by ATP hydrolysis. (4a) Binding of new substrate, ATP and GroES to the *trans* ring induces the release of the natively folded protein, ADP and GroES from the *cis* ring. (4b) If the substrate is not yet completely folded, it can rebind to GroEL in one or more subsequent cycles (Naylor and Hartl 2001).

### 2.4.5.3 Chloroplast and cyanobacterial chaperonins

The chloroplast chaperonin was discovered in the early 1980s, in studies on the assembly of RuBisCO enzyme, where newly-synthesized large subunits of RuBisCO were found to be associated with a 60 kDa protein, named “large subunit binding protein” (LSU-BP) or later ch-Cpn60 (Barraclough and Ellis 1980). Striking parallels between sequence, structure and proposed function of LSU-BP and GroEL from *E. coli* revealed that they are evolutionary homologs, which were then termed “chaperonins” (Ellis 1990, Ellis 1996, Hemmingsen *et al.* 1988).

Whereas the bacterial and mitochondrial chaperonins are composed of 14 identical subunits, the chloroplast Cpn60 comprises two different subunit isoforms,  $\alpha$  (61 kDa) and  $\beta$  (60 kDa). Both subunits are nuclear encoded and need to be imported post-translationally into the chloroplast stroma or thylakoid lumen, before they are processed and assembled to double-ring complexes (Fig. 8 A) (Boston *et al.* 1996, Hill and Hemmingsen 2001, Levy-Rimler *et al.* 2002, Pushkin *et al.* 1982, Tsuprun *et al.* 1991). Interestingly, there are several genes encoding chloroplast Cpn60 proteins, for example two genes encoding  $\alpha$ -subunits and 4 genes encoding  $\beta$ -subunits in *Arabidopsis thaliana*. The amino acid sequences of ch-Cpn60 $\alpha$  and ch-Cpn60 $\beta$  from *A. thaliana* share only approximately 51 % identity with each other, 45 % identity with the mitochondrial Cpn60 and about 50 % identity with *E. coli* GroEL. Within each subunit type of *A. thaliana*, sequence identities amount to 57 % between the two ch-Cpn60 $\alpha$  types and to 60-92 % between the four ch-Cpn60 $\beta$  isoforms. Due to available EST clones, at least one of the two ch-Cpn60 $\alpha$  genes and three of the four  $\beta$  genes seem to be expressed (Cloney *et al.* 1994, Hill and Hemmingsen 2001, Martel *et al.* 1990, Zabaleta *et al.* 1992).

Recombinant expression of ch-Cpn60 subunits in *E. coli* as well as assembly studies *in vitro* revealed that ch-Cpn60 $\alpha$  does not assemble into tetradecameric complexes on its own, whereas ch-Cpn60 $\beta$  alone or the combination of both (ch-Cpn60 $\alpha\beta$ ) yielded functional double-ring species, indicating that  $\beta$ -subunits are required for efficient assembly (Cloney *et al.* 1992a, b, Dickson *et al.* 2000). However, homo-oligomers of ch-Cpn60 $\beta$ <sub>14</sub> have never been observed in chloroplasts. Whenever ch-Cpn60 was purified from plants, both  $\alpha$  and  $\beta$  subunits rather occur in roughly equal amounts in mixed oligomers, suggesting a stoichiometry of  $\alpha_7\beta_7$  in the chaperonin complex (Nishio *et al.*

1999). If the subunits are alternately arranged in both rings or if they form one  $\alpha$ - and one  $\beta$ -ring is still beyond knowledge. In an ATP-dependent manner, ch-Cpn60 can accomplish *in vitro* folding of different substrate proteins in presence of the chloroplast co-chaperones, but also in combination with *E. coli* GroES or mouse mitochondrial Cpn10 (Dickson *et al.* 2000, Viitanen *et al.* 1995).

A variety of proteins has been found to be associated with ch-Cpn60 in chloroplasts, implying a broad substrate spectrum for this chaperonin. Among them are the RuBisCO large and small subunits, the Rieske iron-sulphur protein, Ferredoxin-NADP<sup>+</sup> reductase, glutamine synthetase, the light-harvesting chlorophyll a/b binding protein or the  $\beta$ -subunit of ATP synthase (Jackson-Constan *et al.* 2001, Kessler and Blobel 1996, Landry and Bartlett 1989, Lubben *et al.* 1989, Madueno *et al.* 1993, Tsugeki and Nishimura 1993). The essentiality of chloroplast chaperonins for normal plant development has been shown by analyses of plants or embryos bearing ch-Cpn60 mutations or by RNA interference (Apuya *et al.* 2001, Ishikawa *et al.* 2003, Zabaleta *et al.* 1994).

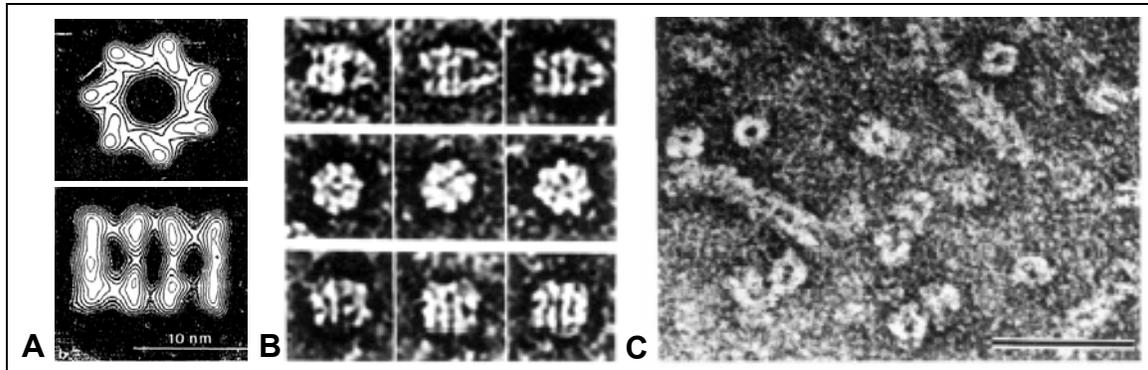
With regard to the origin of the various *ch-Cpn60* genes in plant cells, it might be of interest that some photosynthetic prokaryotes (e.g. *Synechocystis* sp. PCC6803) contain two different, functionally distinguishable species of Cpn60 (Kovacs *et al.* 2001, Lehel *et al.* 1993). One of them (GroEL1), which lacks the C-terminal GGM repeats, is encoded in a bicistronic operon with *groES*, whereas the other (GroEL2) is expressed from a monocistronic gene. Phylogenetic studies suggest that in the evolution of green algae and higher plants, cyanobacteria were initially established as primary endosymbionts in eukaryotic cells, where they underwent a metamorphosis to chloroplasts. During the genesis of chloroplasts, gene flow of *groEL2* (and *groES*) occurred from the symbiont into the host nucleus. In the latter, gene duplication of the cyanobacterial *groEL2* gene resulted in several genes for Cpn60 $\alpha$  and Cpn60 $\beta$ , while the chloroplastic *groEL1* was most likely eliminated (Wastl *et al.* 1999).

The chloroplast chaperonin system is also unique because it is provided with two different nuclear encoded co-factors, namely ch-Cpn20 and ch-Cpn10. Interestingly, the ch-Cpn20 monomer is a protein of approximately 21 kDa which is composed of two GroES-like domains fused together in tandem and sharing approximately 40 % sequence identity or 70 % homology, respectively. Most likely, this structure has arisen either

through gene duplication or by fusion of two distinct Cpn10 genes after the endosymbiotic event in chloroplast genesis, and it might reflect a response to the occurrence of the two divergent Cpn60 isoforms in plants. Indeed, the N-terminal domain of ch-Cpn20 resembles Cpn10 from chlamydia whereas the C-terminal half shows a higher homology to cyanobacterial Cpn10 (Bertsch *et al.* 1992, Boston *et al.* 1996). When expressed separately, both domains of Cpn20 have been shown to perform co-factor function for GroEL in *E. coli*, but not *in vitro* (Baneyx *et al.* 1995, Bertsch and Soll 1995). In *A. thaliana* ch-Cpn20 is nuclear encoded by only one gene and post-translationally transferred into the chloroplast stroma, where it becomes processed and assembled (Hill and Hemmingsen 2001). Although an equilibrium of multiple oligomers was observed, the functional ch-Cpn20 is most likely a homotetramer with a ring-like toroidal structure (Fig. 9 C) (Baneyx *et al.* 1995, Hirohashi *et al.* 1999, Koumoto *et al.* 1999, Sharkia *et al.* 2003). Together with chloroplast Cpn60, it forms asymmetric complexes which bear considerable resemblance to those observed for GroEL and GroES (Fig. 8 B) (Viitanen *et al.* 1995, Xu *et al.* 1997). Interestingly, it was observed that ch-Cpn20 can bind in a calcium-dependent manner to Calmodulin, a ubiquitous  $\text{Ca}^{2+}$  signal transducer in eukaryotes (Yang and Poovaiah 2000). However, evidence for the physiological relevance of this finding *in planta* has not been provided yet.

Next to ch-Cpn20, chloroplasts accommodate ch-Cpn10, which appears to be divergent from other co-chaperones in several respects. Two sequences encoding this protein including a chloroplast transit peptide have been found in the nuclear genome of *A. thaliana* (Hill and Hemmingsen 2001). Like GroES, ch-Cpn10 forms heptameric rings of 10 kDa subunits, but it seems to have a smaller or no roof structure compared to GroES (Koumoto *et al.* 2001, Sharkia *et al.* 2003). This structural feature might contribute to a larger folding cavity in the complex of ch-Cpn10 and Cpn60, similar to the bacteriophage T4 co-factor Gp31, which also misses a roof structure and thus allows folding of the T4 capsid protein Gp23 by GroEL (Hunt *et al.* 1997). Both ch-Cpn20 and ch-Cpn10 have been shown to interact with ch-Cpn60 and GroEL in the refolding of various substrate proteins *in vitro* (Baneyx *et al.* 1995, Bertsch *et al.* 1992, Dickson *et al.* 2000, Koumoto *et al.* 1999, Koumoto *et al.* 2001, Sharkia *et al.* 2003). However, ch-Cpn10 causes deviating folding efficiencies compared to ch-Cpn20, and it also displays a different

expression pattern *in planta*. Hence, both co-factors might have evolved with specialized functions and may assist folding of specific sets of substrates *in organello* (Koumoto *et al.* 2001, Schlicher and Soll 1996).



**Figure 8. Chloroplast Cpn60 and Cpn20.**

(A) Correlation averaging of electron micrographs of Cpn60 from pea leaves, top view (top) and side view (bottom) (Tsuprun *et al.* 1991). (B) Electron micrographs showing side views of asymmetric complexes of ch-Cpn60 from pea chloroplasts and recombinant spinach ch-Cpn20 (top line), end views of asymmetric complexes and/or unliganded ch-Cpn60 (middle line) and side views of unliganded pea ch-Cpn60 (bottom line) (Viitanen *et al.* 1995). (C) Electron micrographs of purified recombinant spinach ch-Cpn20 (Baneyx *et al.* 1995).

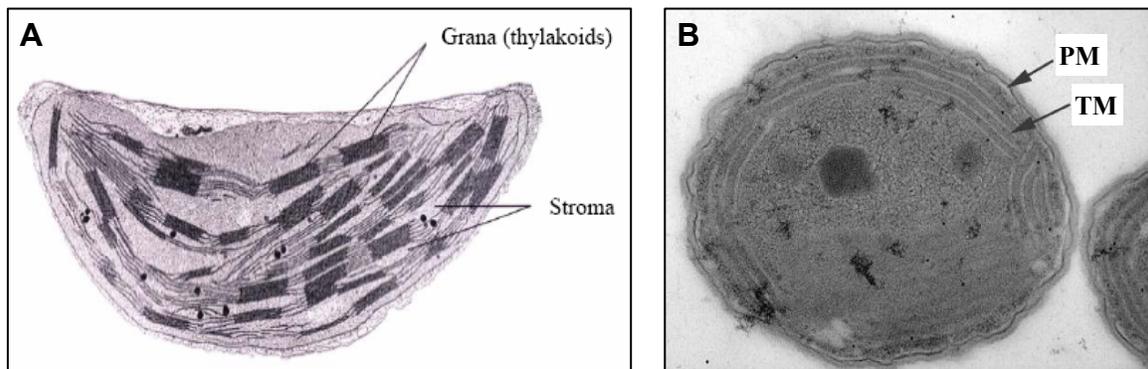
## 2.5 Photosynthesis

Photosynthesis is a fundamental biochemical process, in which photosynthetic organisms (plants, algae, some bacteria and archaea) convert the energy of sunlight into chemical energy, resulting in the formation of organic compounds (carbohydrates) from usually  $\text{CO}_2$  and  $\text{H}_2\text{O}$  ( $6\text{CO}_2 + 6\text{H}_2\text{O} \rightarrow \text{O}_2 + \text{C}_6\text{H}_{12}\text{O}_6$ ).

In photosynthetic eukaryotic cells of plants or algae, photosynthesis takes place in chloroplasts (Fig. 9 A). These organelles are surrounded by two membranes and they contain many flattened, membrane-surrounded vesicles (thylakoids), which are usually arranged in stacks (grana) and whose interior is termed lumen. The thylakoid membranes accommodate all components required for the light-dependent reactions of photosynthesis, such as the light harvesting complexes (containing light-capturing pigments, e.g. chlorophyll, carotene, xanthophyll), the pigment-rich photosystems (PS) I and II with the photosynthetic reaction centers P680 and P700, electron carrier chains, ATP-synthases and further enzyme complexes. In contrast, the interior aqueous phase of

chloroplasts (stroma) is equipped with most enzymes that are involved in the light-independent carbon assimilation reactions.

Unlike eukaryotes, photosynthetic bacteria are not provided with chloroplasts. Instead, their photosynthesis takes place directly within the cell. In gram-negative cyanobacteria, the photosynthetic electron transport machinery, with both photosystems I and II, is embedded into folds of the cell membrane, also called thylakoids (Fig. 9 B). Attached to these thylakoid membranes, phycobilisomes contain pigments (e.g. phycocyanobilin, carotenoids, phycoerythrobilin) and act as light harvesting antennae for PSII. Other photosynthetic bacteria have a variety of different pigments (so-called bacteriochlorophylls) and are often equipped with only one type of specialized photosystem. Under anaerobic conditions, many of these bacteria are able to use electron donors like  $\text{H}_2\text{S}$ ,  $\text{H}_2\text{SO}_4$ ,  $\text{H}_2\text{SO}_3$  or  $\text{H}_2$  instead of water (Schlegel 1992).



**Figure 9. Localization of photosynthesis.**

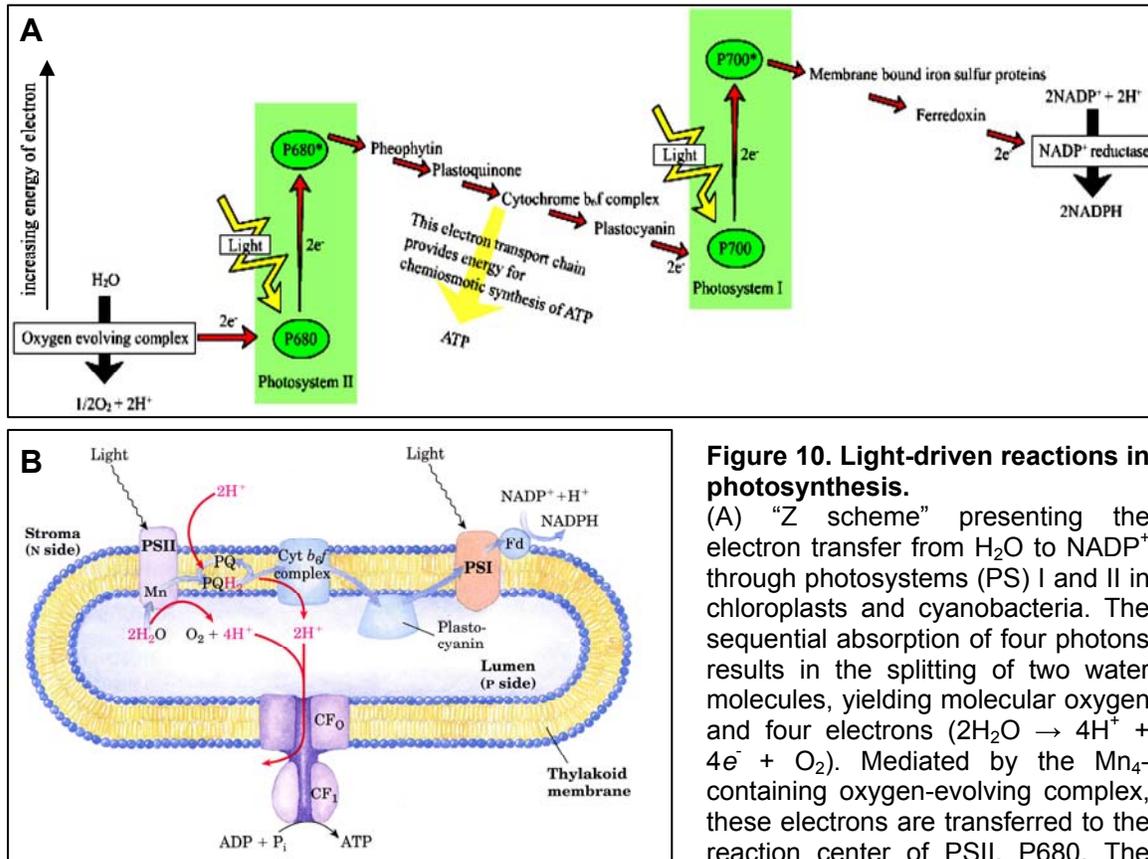
(A) Electron micrograph showing a chloroplast at high magnification. The aqueous phase of the chloroplast interior is called stroma. Flattened, membrane-surrounded vesicles (thylakoids) are forming stacks, the so-called grana (Nelson and Cox 2005). (B) Electron micrograph of the cyanobacterium *Synechocystis* sp. PCC6803. PM: Plasma membrane, TM: thylakoid membranes (Schneider 2000).

### 2.5.1 Light-dependent reactions

In general, photosynthesis can be divided into two processes: the light-dependent reactions and the light-independent carbon-fixation/assimilation reactions.

In the light-driven reactions, chlorophyll and other pigments in the light harvesting complexes (LHC) and in the photosystems absorb light of definite wavelengths, leading to an electron flow through the photosystems and electron carrier chains. As shown in Fig. 10, the electron flow produces on the one hand NADPH and on the other hand it

causes an electrochemical proton gradient between the thylakoid lumen and the chloroplast stroma, which allows ATP-synthesis (Mitchell 1966). Both, the energy carrier ATP and the reductant NADPH, are then provided for the light-independent carbon fixation reactions.



**Figure 10. Light-driven reactions in photosynthesis.**

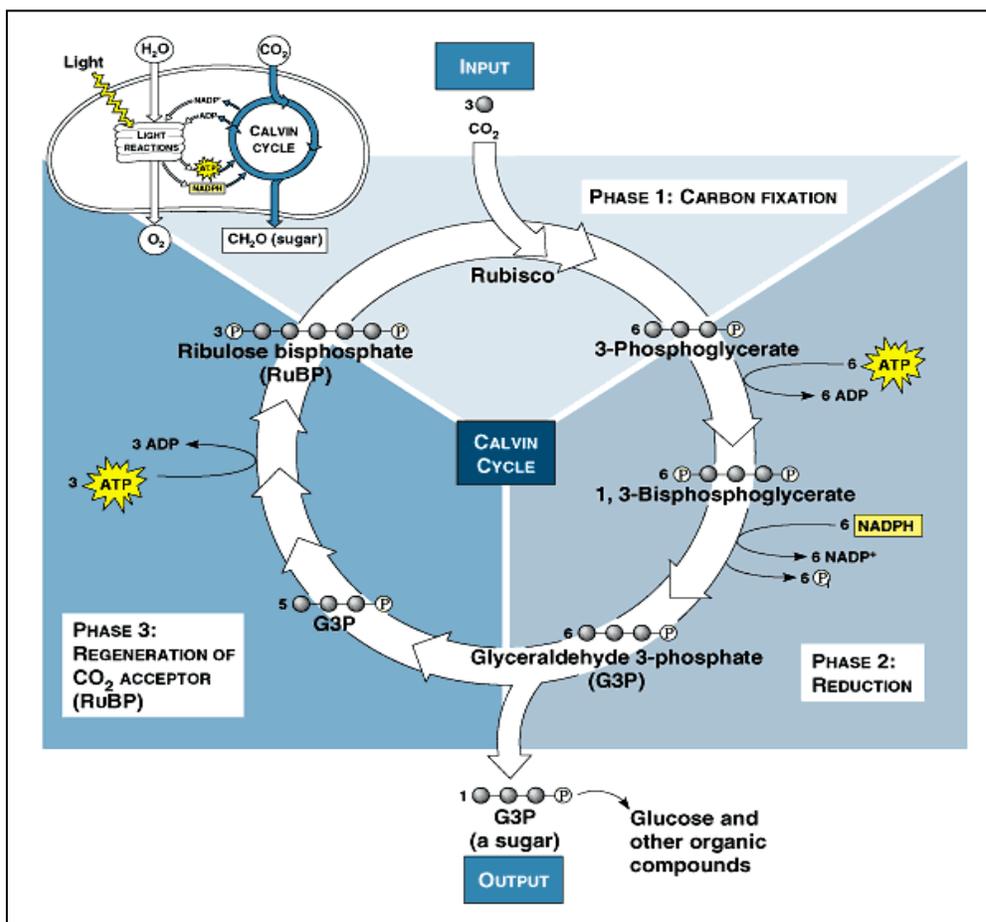
(A) "Z scheme" presenting the electron transfer from  $\text{H}_2\text{O}$  to  $\text{NADP}^+$  through photosystems (PS) I and II in chloroplasts and cyanobacteria. The sequential absorption of four photons results in the splitting of two water molecules, yielding molecular oxygen and four electrons ( $2\text{H}_2\text{O} \rightarrow 4\text{H}^+ + 4e^- + \text{O}_2$ ). Mediated by the  $\text{Mn}_4$ -containing oxygen-evolving complex, these electrons are transferred to the reaction center of PSII, P680. The

resulting  $\text{P680}^*$ , excited by absorbed light energy, transfers electrons over pheophytin to plastoquinone. From here, the electrons pass over a chain of intermediate carriers (cytochrome  $b_6/f$  complex and plastocyanin), to the reaction center of PSI, P700. Upon light excitation, the electrons flow from P700 through a series of carriers to ferredoxin and finally they reduce  $\text{NADP}^+$  to  $\text{NADPH}$ . Between PSI and PSII, the electron flow through plastoquinone and the cytochrome  $b_6/f$  complex drives protons from the stroma into the thylakoid lumen, creating a proton-motive force that provides the energy for ATP synthesis by an ATP synthase, as shown in (B).

(B) Proton and electron circuits in thylakoids. Electrons (blue arrows) move from  $\text{H}_2\text{O}$  through PSII, PSI and intermediate carrier chains to  $\text{NADP}^+$ , yielding  $\text{NADPH}$  ( $\text{NADP}^+ + \text{H}^+ + 2e^- \rightarrow \text{NADPH}$ ). Protons (red arrows) are pumped into the thylakoid lumen, driven by the cyclic and non-cyclic electron flow through the carriers linking PSII and PSI (see A). Additionally, protons are generated in the thylakoid lumen by the light-driven splitting of  $\text{H}_2\text{O}$  in the oxygen-evolving complex. Along the generated gradient, protons pass across the thylakoid membrane and re-enter the stroma through the  $\text{CF}_0$  subunits of ATP-synthases, which in turn utilize the proton motive force for the synthesis of ATP by the  $\text{CF}_1$  subunit ( $\text{ADP} + \text{P}_i + 3\text{H}^+_{\text{thylakoid lumen}} \rightarrow \text{ATP} + 3\text{H}^+_{\text{stroma}}$ ), according to the Mitchell hypothesis (Mitchell 1966). Both  $\text{NADPH}$  and  $\text{ATP}$  now conserve some of the absorbed light energy and can be used for the photosynthetic carbohydrate synthesis in the Calvin cycle (Nelson and Cox 2005).

### 2.5.2 Light-independent reactions / The Calvin cycle

The second, light-independent phase of photosynthesis is the reductive conversion of  $\text{CO}_2$  and  $\text{H}_2\text{O}$  into carbohydrates at the expense of energy and reducing power of ATP and NADPH, which are generated in the light-dependent reactions. The assimilation of carbon dioxide is achieved *via* a cyclic pathway, the Calvin-Benson-Bassham cycle (also known as Calvin cycle, photosynthetic carbon reduction cycle or reductive pentose phosphate pathway). This pathway takes place in the chloroplast stroma or bacterial cytosol and encompasses essentially three phases (Fig. 11).



**Figure 11. The Calvin-Benson-Bassham cycle.**

In the first phase of the cycle, RuBisCO catalyzes the incorporation of  $\text{CO}_2$  into the acceptor ribulose-1,5-bisphosphate, resulting in intermediates that are cleaved into 3-phosphoglycerate. In the subsequent second phase, 3-phosphoglycerate is phosphorylated and reduced in a process that consumes ATP and NADPH from the light-dependent reactions of photosynthesis (see Fig. 10). The resulting glyceraldehyde-3-phosphate can either be metabolized to transport- or storage-competent sugars (e.g. sucrose, starch) or it is utilized for recycling of the acceptor, ribulose-1,5-bisphosphate, in an ATP-dependent multi-step regeneration phase.

Initially, ribulose-1,5-bisphosphate (RuBP) is carboxylated and concomitantly cleaved into two molecules of 3-phosphoglycerate by the enzyme RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase). In the subsequent ATP- and NADPH-consuming second phase, phosphorylation and reduction reactions result in the conversion of 3-phosphoglycerate into glyceraldehyde-3-phosphate. The majority of the latter is utilized to regenerate the acceptor, RuBP, in the third phase of the cycle. This phase requires ATP and comprises sequential transaldolase and transketolase reactions including several intermediates of three-, four-, five-, six- and seven-carbon sugars. A small fraction of glyceraldehyde-3-phosphate is not used for recycling, but it can be converted either to starch for storage in chloroplasts or to sucrose for transport purposes or it might be degraded in the cytosol by glycolysis to provide energy. As a result of the Calvin cycle, the carboxylation of 6 C<sub>5</sub> molecules yields 1 C<sub>6</sub> molecule (glucose-P or fructose-P) and allows reconstitution of 6 C<sub>5</sub> molecules at the expense of 18 ATP and 12 NADPH.

## 2.6 RuBisCO

### 2.6.1 Reactions catalyzed by RuBisCO

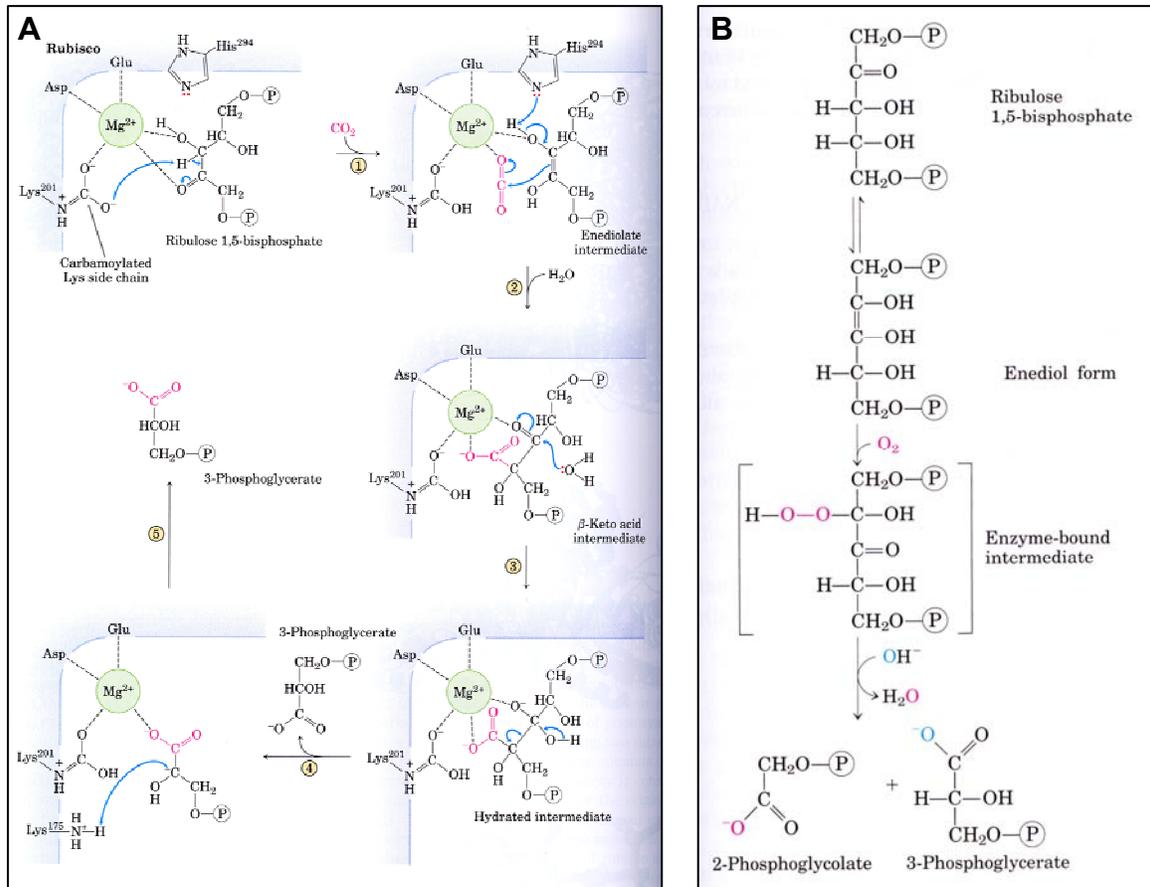
The initial step in the Calvin-Benson-Bassham cycle is catalyzed by the RuBisCO enzyme. RuBisCO makes up almost 50 % of soluble chloroplast protein and is one of the most abundant enzymes in the biosphere. As a carboxylase, it promotes the covalent attachment of CO<sub>2</sub> to the five-carbon sugar RuBP as well as the concomitant cleavage of the unstable six-carbon intermediate into two molecules of 3-phosphoglycerate (Fig. 12 A).

Unfortunately, RuBisCO is not absolutely specific for CO<sub>2</sub>. Instead, it can also utilize O<sub>2</sub> as substrate and thus catalyze the oxygenation of RuBP, resulting in 3-phosphoglycerate and 2-phosphoglycolate (Fig. 12 B). Since 2-phosphoglycolate is a metabolically useless compound, it usually enters the glycolate pathway, which includes sequential reactions in the chloroplast stroma, the peroxisomes and in the mitochondrial matrix. In this pathway, two molecules of 2-phosphoglycolate are converted to one molecule of CO<sub>2</sub> and one molecule of 3-phosphoglycerate. The latter can re-enter the Calvin cycle and therefore some fixed carbon is salvaged. But nevertheless, a considerable proportion of it becomes released as CO<sub>2</sub> and, even worse, significant amounts of cellular energy are consumed by

the whole process. Hence, the RuBisCO-catalyzed oxygenation reaction is a pretty expensive process for the cell – it results in no fixation of carbon and causes both energetic and metabolic losses. Since this pathway consumes oxygen and produces carbon dioxide, similar to mitochondrial respiration, it is also called photorespiration (Nelson and Cox 2005, Spreitzer and Salvucci 2002). Photorespiration principally reduces the photosynthetic potential of plants and consequently limits their growth rate.

Both the carboxylation and oxygenation reactions are competitive. The apparent affinity of plant RuBisCO for CO<sub>2</sub> (K<sub>m</sub> ca. 9 μM) is much greater than for O<sub>2</sub> (K<sub>m</sub> ca. 350 μM). However, the atmospheric composition (20 % O<sub>2</sub>, 0.04 % CO<sub>2</sub>) and the solubility of both compounds in aqueous solutions (250 μM O<sub>2</sub>, 11 μM CO<sub>2</sub>) allow a significant proportion of oxygenation to occur (Nelson and Cox 2005). Under moderate conditions, O<sub>2</sub> binds to RuBP in every fourth or fifth turnover. This rate increases further at low CO<sub>2</sub> concentrations at the site of synthesis especially at higher temperatures, when the ratio of O<sub>2</sub> to CO<sub>2</sub> increases, resulting in a decrease of the RuBisCO affinity for CO<sub>2</sub>. There are several plants (especially found in hot and dry climates) that have developed mechanisms, which guarantee sufficient CO<sub>2</sub> fixation under these conditions (Hatch-Slack-cycle of C<sub>4</sub> plants or crassulaceae acid metabolism of CAM plants).

In addition, the enzymatic activity of RuBisCO is relatively low – only 3 carbon dioxide molecules are fixed per second and per molecule of RuBisCO at 25 °C. Thus, when light is not limiting photosynthesis, it is RuBisCO itself which constitutes the primary rate-limiting enzyme of the Calvin cycle. Consequently, a lot of efforts have been made to engineer RuBisCO catalyzing predominantly the carboxylation reaction or displaying higher catalytic rates (Andrews and Whitney 2003, Parry *et al.* 2003, Spreitzer and Salvucci 2002). So far, all attempts have only resulted in modest improvements. It is postulated that photorespiration might be of physiological importance and that evolution has already nearly perfectly optimized RuBisCO to the different gaseous and thermal environments of the plants (Smith and Tabita 2003, Tcherkez *et al.* 2006).



**Figure 12. Reactions catalyzed by RuBisCO.**

(A) Carboxylase activity of RuBisCO. RuBisCO is activated after carbamoylation of the  $\epsilon$ -amino group of Lys<sup>201</sup> and binding of Mg<sup>2+</sup> to the active site (shaded in grey). Mg<sup>2+</sup> brings together and orients the reactants at the active site. One molecule CO<sub>2</sub> and one molecule RuBP form two molecules of 3-phosphoglycerate, one of which contains the carbon atom of the fixed CO<sub>2</sub> (red). (1) RuBP forms an enediolate. (2) CO<sub>2</sub> is polarized by the proximity of the Mg<sup>2+</sup> ion and undergoes nucleophilic attack by the enediolate, producing a branched six-carbon intermediate. (3) C-3 of this intermediate is hydroxylated. (4) Upon aldol cleavage, one molecule of 3-phosphoglycerate is formed, which leaves the enzyme active site. (5) The carbanion of the remaining three-carbon fragment is protonated by the nearby site chain of Lys<sup>175</sup>, generating a second molecule of 3-phosphoglycerate. (B) Oxygenase activity of RuBisCO. Incorporation of O<sub>2</sub> into RuBP results in an unstable intermediate that splits into 2-phosphoglycolate and 3-phosphoglycerate. The former can be recycled in the costly glycolate pathway, the latter can re-enter the Calvin cycle (Nelson and Cox 2005).

### 2.6.2 Regulation of RuBisCO activity

The Calvin-Benson-Bassham cycle begins to operate only when sufficient amounts of NADPH and ATP from the light-dependent photosynthetic reactions as well as the sugar substrate (RuBP) for the RuBisCO enzyme are available. Accordingly, the activity of RuBisCO and of several further enzymes in the Calvin cycle is regulated and coordinated

by light and expression of the RuBisCO subunits is induced upon illumination on both transcriptional and translational level (Berry *et al.* 1985, 1986, Dean *et al.* 1989).

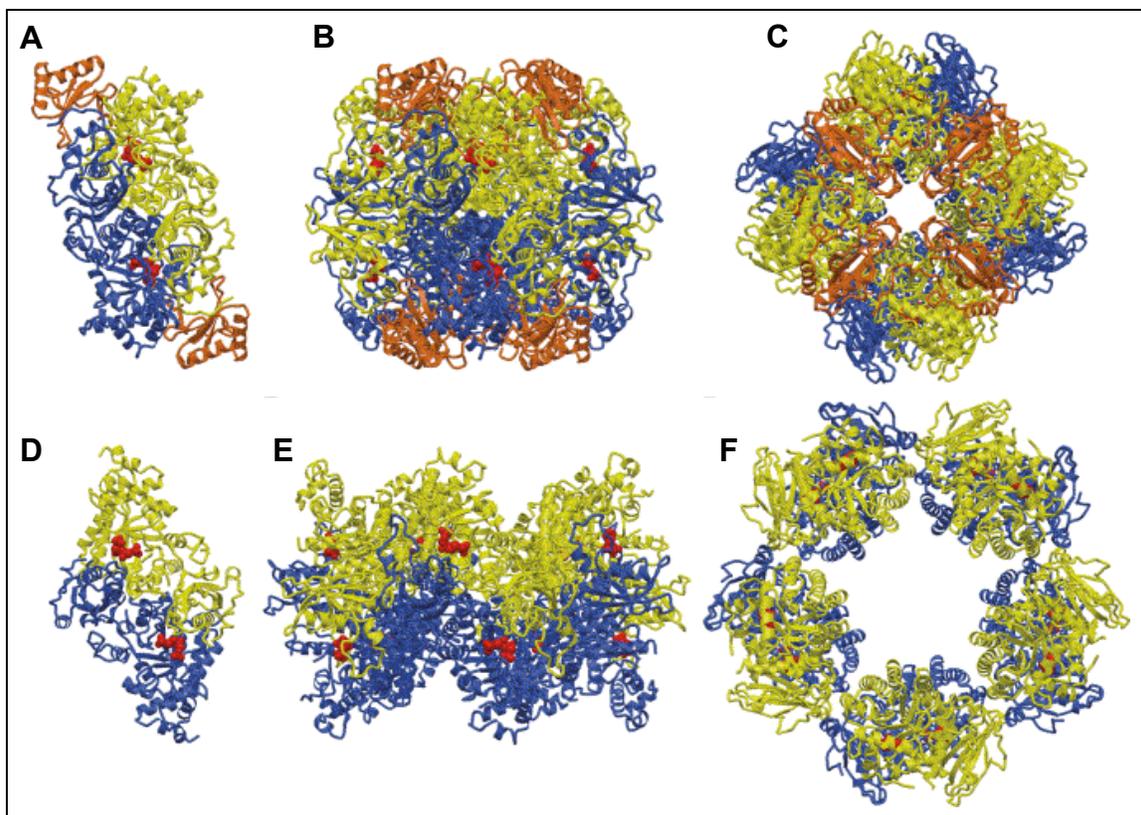
Moreover, illumination and light reaction cause the transfer of protons from the stroma into the thylakoid lumen (Fig. 10) and the movement of  $Mg^{2+}$  out of the thylakoids, resulting in an increased pH (about 8.0) and elevated  $Mg^{2+}$  concentrations in the stroma. Since RuBisCO has a high optimal pH and requires  $Mg^{2+}$  in the active site for catalytic activity, it becomes activated this way. At the same time, reduced ferredoxin/thioredoxin and NADPH are generated in the light-dependent reactions and in concert with the increased stromal pH and  $Mg^{2+}$  levels, they optimize conditions for several enzymes in the regeneration phase of the Calvin cycle (fructose biphosphatase, FBPase, or sedoheptulose biphosphatase, SBPase) (Wolosiuk *et al.* 1993).

For catalysis, RuBisCO first needs to become activated by carbamoylation at the  $\epsilon$ -amino group of Lys201 and by binding  $Mg^{2+}$  at its active site (Fig. 12) (Lorimer and Mizioroko 1980). However, the substrate RuBP binds more strongly to active sites lacking the carbamate, and the active site can also become occupied during darkness by inhibitory substrate analogs (e.g. 2-carboxy-D-arabinitol-1-phosphate or xylulose-1,5-bisphosphate), leading to the inhibition of RuBisCO. An enzyme termed RuBisCO activase promotes the release of these inhibitors from the active site, allowing proper activation and binding of RuBP. Since the reaction catalyzed by RuBisCO activase is ATP-dependent, inhibited by ADP and modified by the stromal redox state, the activity of the activase and thus also of RuBisCO is regulated by light intensity (Portis 1990 and 2003, Spreitzer and Salvucci 2002, Wolosiuk *et al.* 1993, Zhang *et al.* 2002).

Moreover, it has been shown that inorganic phosphate can modulate the coordinated regulation of RuBisCO by affecting the conformational transitions between the activated and the less active enzyme. In this respect, phosphate appears to have positive effects on cyanobacterial RuBisCO activity at suboptimal levels of  $CO_2$  and at high concentrations of RuBP (Marcus and Gurevitz 2000).

### 2.6.3 Structure of RuBisCO

Three major structural types of  $CO_2$  fixing RuBisCO have been classified so far (Fig. 13) (Andersson and Taylor 2003, Spreitzer and Salvucci 2002, Tabita 1999).



**Figure 13. Different arrangements of the quaternary structure of RuBisCO showing the molecular symmetry.**

(A-C) Type I RuBisCO from spinach; RbcL<sub>2</sub>S<sub>2</sub> unit viewed down the twofold symmetry axis (A) and RbcL<sub>8</sub>S<sub>8</sub> hexadecamer viewed down the twofold (B) and fourfold (C) axes. (D) Dimeric type II RuBisCO (RbcL<sub>2</sub>) from *Rhodospirillum rubrum* showing the twofold symmetry. (E and F) RbcL<sub>10</sub> RuBisCO from *Thermococcus kodakaraensis* viewed down the twofold (E) and fivefold (F) axes, respectively. Large subunits are depicted in yellow and blue, small subunits are shown in orange and active sites are colored in red (Andersson and Taylor 2003).

The dominant type I enzyme is a complex of tetragonal symmetry (RbcL<sub>8</sub>S<sub>8</sub>, MW ~550000), composed of eight large subunits (RbcL, 50-55 kDa) and eight small subunits (RbcS, 12-18 kDa). It can be found in eukaryotic and most prokaryotic photosynthetic organisms, i.e. plants, algae and most cyanobacteria (Curmi *et al.* 1992, Hartman and Harpel 1994, Knight *et al.* 1990, Newman *et al.* 1993, Schreuder *et al.* 1993). In contrast, type II RuBisCO is usually a homodimer of large subunits (RbcL<sub>2</sub>, MW ~110000) and occurs e.g. in the purple non-sulfur photosynthetic bacterium *Rhodospirillum rubrum* (Schneider *et al.* 1986, 1990a, Tabita and McFadden 1974a,b). Several chemoautotrophic bacteria have been shown to possess both type I and II enzymes or particular species of type II RuBisCO (Chory *et al.* 1985, English *et al.* 1992, Hernandez *et al.* 1996, Shively

*et al.* 1998). More recently, a new type III of RuBisCO from the archaeon *Thermococcus kodakaraensis* KOD1 was described as a ring-shaped complex of five large subunit dimers, (RbcL<sub>2</sub>)<sub>5</sub> (MW ~497000; Kitano *et al.* 2001, Maeda *et al.* 1999).

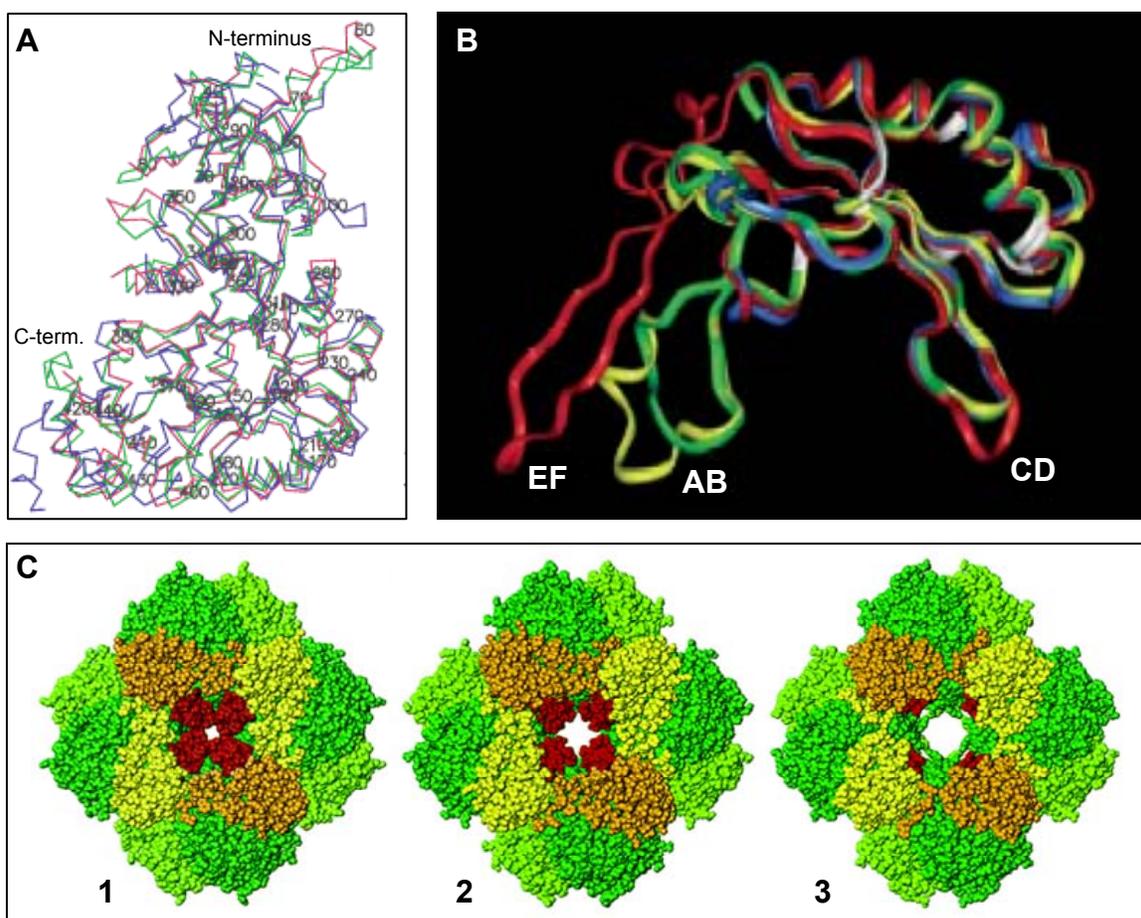
The large subunit monomers consist of two domains (Fig. 14 A). The smaller N-terminal domain is composed of a five-stranded mixed  $\beta$ -sheet with two  $\alpha$  helices on both sides of the sheet. The larger C-terminal domain forms an eight-stranded parallel  $\alpha/\beta$ -barrel. The active site residues are situated in the loops that connect the  $\beta$ -strands with the  $\alpha$ -helices on the C-terminal side of the barrel. Hence, there are two active sites per RbcL<sub>2</sub> dimer located in the interface between the N-terminal domain of one subunit and the C-terminal domain of the second subunit. The interactions between the two subunits of a dimer are tight and extensive (Schneider *et al.* 1986).

Although organisms with the type I RbcL<sub>8</sub>S<sub>8</sub> enzyme have a very diverse phylogenetic background, their large RuBisCO subunits display a considerable sequence homology of 70-90 %. Large subunits of all three enzyme types are only 30-35 % identical in amino acid sequence, but they share a striking similarity in the three-dimensional structure of their RbcL<sub>2</sub> units as well as highly conserved active sites (Fig. 13 and Fig. 14A) (Kitano *et al.* 2001, Schneider *et al.* 1990b).

The small subunits in RbcL<sub>8</sub>S<sub>8</sub> are required for maximal catalytic rates, but hardly contribute to the enzyme's CO<sub>2</sub>/O<sub>2</sub> specificity. They show a higher divergence than their large counterparts as well as functionality even in interspecific hybrid enzymes. For the assembly of RbcL<sub>8</sub> core complexes, small subunits are not obligatory, but they appear to maintain the structural integrity of RbcL<sub>8</sub>S<sub>8</sub> holoenzyme by establishing dominant contacts between the four RbcL<sub>2</sub> dimers (Andrews 1988, Andrews and Ballment 1983, Curmi *et al.* 1992, Gutteridge 1991, Lee and Tabita 1990, Lee *et al.* 1991, Spreitzer 2003, Spreitzer and Salvucci 2002, van der Vies *et al.* 1986).

Some major structural variations in small subunits of different phylogenetic origin have been found in the  $\beta$ A- $\beta$ B loop, which reside at each end of the solvent channel that traverses the holoenzyme and appears to make interactions between two adjacent small subunits and three large subunits (Andersson and Taylor 2003, Knight *et al.* 1990, Newman and Gutteridge 1993). Whereas cyanobacteria and non-green algae have only 10 residues in this  $\beta$ A- $\beta$ B loop, the loop of higher plants is composed of 22 amino acids and

that of green algae contains 28 residues (Fig. 14 B). This extension narrows the entrance to the holoenzyme's central solvent channel (Fig. 14 C). Moreover, some prokaryotes, non-green algae and dinoflagellates display additional residues in the  $\beta\text{C}$ – $\beta\text{D}$  loop as well as an extension of approximately 30 residues at the C-terminus that forms an additional  $\beta\text{E}$ – $\beta\text{F}$  loop in the central solvent channel of the holoenzyme, which fills the space arising from the smaller  $\beta\text{A}$ – $\beta\text{B}$  loop (Andersson and Taylor 2003, Karkehabadi *et al.* 2005, Spreitzer 2003, Spreitzer and Salvucci 2002).



**Figure 14. Comparison of the X-ray crystal structures of RuBisCO subunits.**

(A) Comparison of monomeric large subunits of RuBisCO from spinach, type I (green), *Rhodospirillum rubrum*, type II (blue) and *Thermococcus kodakaraensis*, type III (red) (Kitano *et al.* 2001). (B) Comparison of small RuBisCO subunits from *Chlamydomonas* (yellow), spinach (green), *Synechococcus* (blue) and *Galdieria partita* (red). Loops are labeled relative to their flanking  $\beta$ -strands. Residues that are more than 95 % conserved among all known small subunit sequences are colored white (Spreitzer 2003). (C) Comparison of RuBisCO holoenzyme structure top views from *Chlamydomonas* (1), spinach (2) and *Synechococcus* (3). Large subunits are depicted in green and light-green, small subunits are shown in yellow and orange. The  $\beta\text{A}$ – $\beta\text{B}$  loops (red) of four small subunits reside at each end of the solvent channel that traverses the holoenzyme (Karkehabadi *et al.* 2005).

#### 2.6.4 Folding and assembly of RuBisCO and the role of chaperones

Prokaryotic organisms synthesize and assemble their RuBisCO in the cytosol after transcription of a single gene for large subunits (type II RuBisCO, RbcL<sub>2</sub>) or of an operon encoding both the large and small subunits (type I RuBisCO, RbcL<sub>8</sub>S<sub>8</sub>). In eukaryotic cells, functional type I RuBisCO is only found in the chloroplasts. However, the small subunit *rbcS* genes are generally encoded in the nucleus as a small multi-gene family of up to more than ten members. Therefore, after translation on cytosolic ribosomes, the small subunit precursor proteins have to be imported into the chloroplasts, where they are processed and fold to the native state (Dean *et al.* 1989, Gatenby and Ellis 1990). In contrast, higher plants' large subunit *rbcL* genes occur as a single copy per chloroplast genome, but since each chloroplast contains several copies of this genome, a higher number of the same *rbcL* gene can be found in each chloroplast. These *rbcL* genes are structurally similar to prokaryotic genes (and therefore also amenable to expression in *E. coli*). For example, they lack introns, have 5'-sequences that resemble bacterial promoters and produce mRNAs with Shine-Dalgarno-like sequence stretches. Moreover, their translation is initiated by fmet-tRNA and carried out by 70S ribosomes in the chloroplast stroma (Gutteridge and Gatenby 1995, Somerville *et al.* 1986).

Recombinantly, the dimeric RbcL<sub>2</sub> enzyme of *Rhododspirillum rubrum* can be expressed efficiently in *E. coli* yielding considerable amounts of soluble and active enzyme. It can also be refolded from the denatured state *in vitro* with the *E. coli* chaperonin system GroEL/GroES (Goloubinoff *et al.* 1989b, Viitanen *et al.* 1990). In contrast, the type I RuBisCO (RbcL<sub>8</sub>S<sub>8</sub>) has never been refolded *in vitro* and results in the formation of insoluble aggregates. Recombinant expression of cyanobacterial type I RuBisCO resulted in the production of ~ 2-5 % soluble active enzyme. Higher plant RuBisCO, on the other hand, has never been produced as an active enzyme in *E. coli*, even not upon co-expression of the chloroplast chaperonin system (Cloney *et al.* 1993, Gatenby 1984, Gatenby *et al.* 1987, Somerville *et al.* 1986).

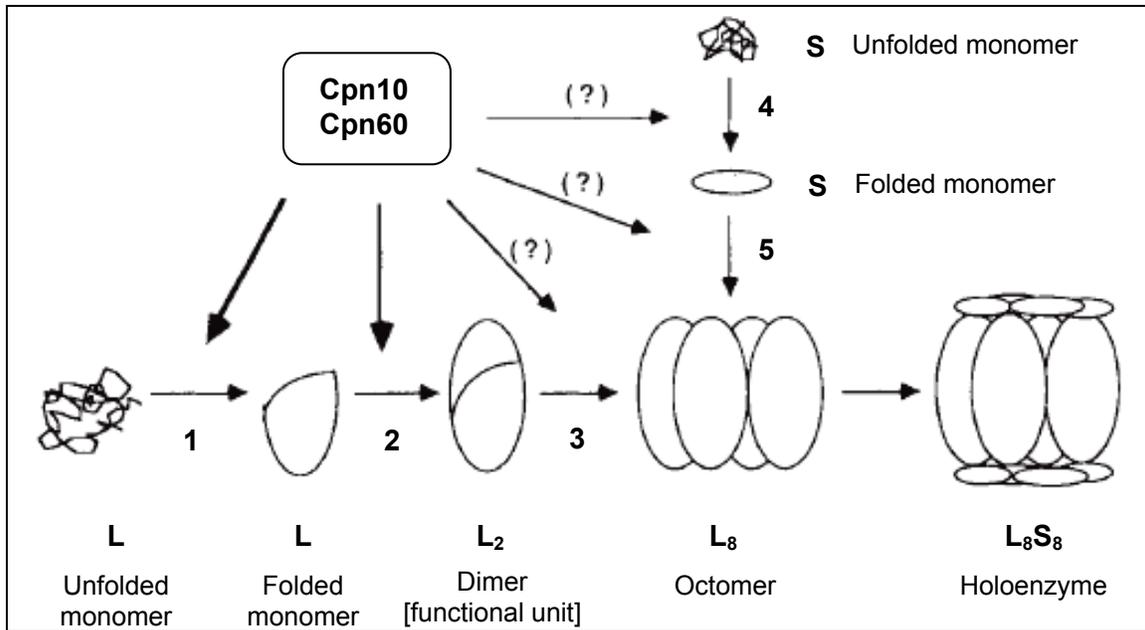
Since some cyanobacterial type I RuBisCOs can be expressed in *E. coli*, yielding small amounts of soluble and active enzyme, they provide an opportunity to investigate folding and assembly of the RbcL<sub>8</sub>S<sub>8</sub> complex. RuBisCO from the cyanobacterial species *Synechococcus* sp. PCC6301 (*Anacystis nidulans*) is the most frequently utilized

representative (Goloubinoff *et al.* 1989a, Gatenby *et al.* 1985, Tabita and Small 1985). Assembled RbcL<sub>8</sub> core complexes of this enzyme have been recombinantly produced in *E. coli*, purified and assembled to the active holoenzyme upon addition of small subunits (Andrews 1988, Lee and Tabita 1990). Furthermore, minute amounts of soluble and active RuBisCO from *Anabaena* sp. PCC7120 (Fitchen *et al.* 1990, Gurevitz *et al.* 1985, Larimer and Soper 1993) or *Chromatium vinosum* (Viale *et al.* 1985) could be synthesized in *E. coli*.

It is unclear, why RuBisCO of some cyanobacteria can assemble in *E. coli*, whereas that of higher plants cannot - despite the high degree of sequence and structure similarity (Knight *et al.* 1990, Newman *et al.* 1993). It has been suggested, that the large subunits undergo some special post-translational modifications in chloroplasts, which might be necessary for proper folding and assembly, but maybe not implemented in *E. coli*. Indeed, several RuBisCO modifications have been found, e.g. deformylations, acetylations, hydroxylations, methylations or carbamoylations. However, some of these modifications affect the enzyme's activity or stability more than its assembly, others occur only in selected but not all plant or algal species, few appear only after assembly and several modifications can also be found in prokaryotes (Houtz and Portis 2003). Consequently, it is less likely that missing post-translational modifications cause the failure to produce or reconstitute functional plant RuBisCO in *E. coli* or *in vitro*. Instead, it might be rather one or several specific factors, which support the folding and assembly pathway in chloroplasts, but which are absent in *E. coli* or in *in vitro* systems.

Beyond doubt, the chaperonin Cpn60/GroEL is one fundamental factor in the folding and assembly pathway of RuBisCO. Both type I and II of RuBisCO have been shown to be a substrate of GroEL/GroES, whose overproduction positively affects the production of active RuBisCO in *E. coli* (Goloubinoff *et al.* 1989a,b, Larimer and Soper 1993). In chloroplasts of higher plants, newly synthesized large subunits were found to be associated with Cpn60 (LSU-BP, large subunit binding protein) (Barracough and Ellis 1980, Cannon *et al.* 1986, Gatenby and Ellis 1990, Gutteridge and Gatenby 1995, Roy 1989, Roy and Cannon 1988). Involvement of the DnaK/DnaJ/GrpE chaperone system in folding of recombinantly expressed cyanobacterial type I RuBisCO has been suggested as well (Checa and Viale 1997).

In the current view of the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme maturation pathway, large subunits are folded by and released from Cpn60 in a manner dependent on Mg<sup>2+</sup> and ATP, followed by the formation of RbcL<sub>2</sub> dimers. The latter subsequently tetramerize around a fourfold axis to RbcL<sub>8</sub> cores, before eight small subunits spontaneously associate with the RbcL<sub>8</sub> core on top and bottom, resulting in RbcL<sub>8</sub>S<sub>8</sub> (Fig. 15) (Gatenby and Ellis 1990, Goloubinoff *et al.* 1989a, Hubbs and Roy 1993b, Roy *et al.* 1988).



**Figure 15. Assembly pathway for type I RuBisCO (RbcL<sub>8</sub>S<sub>8</sub>).**

Large subunit monomers are folded (1) and assemble subsequently into RbcL<sub>2</sub> dimers (2). Four dimers tetramerize into the RbcL<sub>8</sub> core (3). Folding of small subunits (4) and association of small subunits with the RbcL<sub>8</sub> core (5) to form the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme. (Gatenby and Ellis 1990, Goloubinoff *et al.* 1989a).

### 2.6.5 RbcX

In the search for further factors promoting the formation of productive type I RuBisCO, a factor termed RbcX has recently attracted some attention, since in numerous cyanobacteria (e.g. *Anabaena* sp. PCC7120, *Anabaena* sp. CA, *Synechocystis* sp. PCC6803, *Synechococcus* sp. PCC7002) the *rbcX* gene is located in the intergenic space between the large (*rbcL*) and the small (*rbcS*) subunit genes of the *rbc* operon (Larimer and Soper 1993, Li and Tabita 1994, Onizuka *et al.* 2002, Onizuka *et al.* 2004). Assuming that the position of *rbcX* in the *rbcLXS* operon and the resulting co-expression of RbcX with both RuBisCO subunits is not fortuitous, a functional relation or interaction

between RbcX and RbcLS is a reasonable presumption, even though in the genomes of some other cyanobacterial or higher plant species *rbcX* homologues can be found encoded outside the *rbc* operon.

When RuBisCO from *Anabaena* sp. CA was produced in *E. coli*, co-expression of GroEL/GroES ameliorated the recovery of active enzyme. Overproduction of the conspecific RbcX imitated and even enhanced the positive effect of GroEL/GroES, suggesting a chaperone-like function of RbcX (Li and Tabita 1997, 1994).

Furthermore, it was reported that in *Synechococcus* sp. PCC7002 (*Agmenellum quadruplicatum* PR-6) a translational frameshift in the *rbcX* gene (ORF134) resulted in a significant decrease in both the recovery of soluble RuBisCO subunits and measurable RuBisCO activity, whereas a complete loss of the *rbcX* gene impaired the viability of cyanobacterial cells. Expression of the *rbc* genes from *Synechococcus* sp. PCC7002 in the cellular environment of *E. coli* revealed that RbcX promotes the synthesis of RbcLS and their post-translational assembly into the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme structure. RbcX was found not to be involved in transcriptional regulation and its ability to positively affect RbcLS synthesis suggested a chaperone-like function.

In the case of *Synechococcus* sp. PCC7942, a species encoding *rbcX* outside the *rbcLS* operon, *rbcX* deletion did not affect the cyanobacterial growth rate or RuBisCO activity and production of active RuBisCO in *E. coli* was observed in the absence of RbcX co-expression (Emlyn-Jones *et al.* 2006). However, the involvement of RbcX in RuBisCO translation, folding, assembly and protection from degradation or aggregation is still under discussion.

## 2.7 Aim of the study

For engineering plant RuBisCO with improved catalytic properties, which could be advantageous regarding agricultural productivity, a fast and efficient screening system providing sufficient amounts of active RuBisCO would be extremely beneficial in order to overcome the currently applied time-consuming and species-limited methods for functional mutant analysis (e.g. plastid transformation, Andrews and Whitney 2003). But both the mechanism and the molecular requirements for the folding and assembly pathway of the hexadecameric type I RuBisCO are not yet understood entirely,

presenting an obstacle in the establishment of appropriate heterologous *in vivo* or *in vitro* expression and screening systems. So far, *in vitro* synthesis and assembly of functional plant RuBisCO requires rather homologous systems containing chloroplast extracts or chloroplasts themselves (Cannon *et al.* 1986, Hubbs and Roy 1992, 1993a, b, Roy 1989, Roy *et al.* 1982). It has already been shown that folding of the enzyme requires the assistance of the Cpn60 chaperonin machinery, and involvement of the Hsp70 system has been proposed as well. However, *in vitro* reconstitution of this complex enzyme has not been accomplished thus far.

More recently, RbcX was suggested to promote the production of assembled cyanobacterial type I RuBisCO, but structural information on RbcX and a concrete mechanistic conception of its involvement in the folding and assembly pathway of type I RuBisCO are still missing.

Hence, the aim of this study was to analyze the requirements and impediments for the *in vitro* reconstitution of type I RuBisCO. Moreover, the structural peculiarities of the putative assembly chaperone RbcX as well as its functional significance and mechanistic attributes in the RuBisCO folding and assembly pathway were to be studied in detail. Based on the solution of the RbcX crystal structure, site directed mutagenesis as well as biochemical *in vivo* and *in vitro* assays, a model for the RbcX-mediated folding and assembly pathway of type I RuBisCO was to be developed.

### 3 MATERIAL AND METHODS

#### 3.1 Materials

##### 3.1.1 Chemicals

All chemicals were of *pro analysi* quality and were purchased from **Sigma-Aldrich** (Steinheim, Germany) or **Merck** (Darmstadt, Germany) unless stated otherwise.

**Amersham Biosciences** (Freiburg, Germany): chromatogr. resins, NaH<sup>14</sup>CO<sub>3</sub>, <sup>35</sup>S-Met

**BioMol** (Hamburg, Germany): IPTG, HEPES

**BioRad** (Munich, Germany): Bradford protein assay dye reagent, ethidiumbromide

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

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

**Fluka** (Deisenhofen, Germany): acetic acid, acetone, Bis-Tris, H<sub>2</sub>O<sub>2</sub>, luminol, PEG

**Invitrogen** (Karlsruhe, Germany): dNTP set

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

**Pierce** (Rockford / IL, USA): DSP

**Qiagen** (Hilden, Germany): Ni-NTA-Agarose

**Roche** (Basel, Switzerland): Complete protease inhibitor cocktail, DTT

**Roth** (Karlsruhe, Germany): ampicillin, glycine, scintillation fluid

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

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

##### 3.1.2 Reagent and purification kits

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

**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): QIAquick PCR Purification Kit, QIAprep Spin Miniprep Kit, QIAGEN Plasmid Midi Kit

**Roche** (Basel, Switzerland): RTS 100 *E. coli* HY Kit

### 3.1.3 Strains

**American Type Culture Collection**, ATCC (Manassas / VA, USA): *Synechococcus* sp. PCC7002 [PR-6], ATCC # 27264

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

**Novagen** (Darmstadt, Germany): *E. coli* DH5a

### 3.1.4 Plasmids, DNA and oligonucleotides

Chromosomal DNA from *Synechococcus* sp. PCC7002 was purified from cyanobacterial cultures as described below (chapter 3.3.3).

An *Arabidopsis thaliana* cDNA library and pSOCPN70/8.1 (encoding Cpn20 from *Spinacia oleracea*, Bertsch *et al.* 1992) were both friendly gifts of Prof. Dr. J. Soll.

The plasmids pUC-*rbcLS* (Smith and Tabita 2003) and pLS32 (Li and Tabita 1997) encode the *rbc*-operons from *Synechococcus* sp. PCC6301 and *Anabaena* sp. CA, respectively, and were a kind gift from F.R. Tabita.

Plasmid pBAD33*ES/EL* encodes the arabinose inducible *groEL/ES* operon of *E. coli* (Ewalt *et al.* 2000). pG-*KJE8* contains the operons for both *groEL/ES* and *dnaK/dnaJ/grpE* from *E. coli*, which are inducible by tetracycline or arabinose, respectively (Nishihara *et al.* 2000). pG-*Tf2* can be used for tetracycline inducible expression of *groEL/ES* and *trigger factor* (TF) from *E. coli* (Nishihara *et al.* 2000).

Plasmids that were generated throughout this study, are listed in the appendices and will be described below (chapter 3.3.7). They encode the IPTG-inducible *rbc*-genes from *Synechococcus* sp. PCC7002, *Synechococcus* sp. PCC6301 and *Anabaena* sp. CA as well as the genes for chaperonins and their co-factors from *Arabidopsis thaliana* or *Spinacia oleracea*.

**Invitrogen** (Karlsruhe, Germany): pProEX-HTa

**Novagen** (Darmstadt, Germany): pET11a, pET15b, pET28b, pET30b

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

### 3.1.5 Enzymes, proteins, peptides and antibodies

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

*Amersham Biosciences* (Freiburg, Germany): porcine RNAGuard ribonuclease inhibitor, Protein A sepharose beads

*Dr. E. Schleiff and Dr. T. Becker*: pea chloroplast stroma

*Merck* (Darmstadt, Germany): Benzonase

*New England Biolabs* (NEB, Frankfurt/Main, Germany): Vent DNA polymerase, 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-Cpn60 $\alpha$  antibody, anti Cpn60 $\beta$ -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

### 3.1.6 Media

Media were prepared with deionized, double-distilled water and subsequently autoclaved.

LB medium: 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl,  
(15 g/l agar for solid medium), pH 7.1 adjusted with NaOH

SOB medium: 20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 0.186 g/l KCl,  
pH 7.0 adjusted with NaOH, 10 mM MgCl<sub>2</sub> added before use

M9 minimal medium: 800 ml autoclaved H<sub>2</sub>O, 200 ml 5 x M9 salts, 2 ml 1M MgSO<sub>4</sub>,  
0.1 ml 1 M CaCl<sub>2</sub>, 20 ml 20 % sterile filtered glucose,  
1 ml 20 mg/ml thiamine

5 x M9 salts: 43 g/l Na<sub>2</sub>HPO<sub>4</sub> x 2 H<sub>2</sub>O, 15 g/l KH<sub>2</sub>PO<sub>4</sub>, 2.5 g/l NaCl,  
5 g/l NH<sub>4</sub>Cl

MN marine medium: 40 mg/l MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 20 mg/l CaCl<sub>2</sub> x 2 H<sub>2</sub>O, 750 mg/l  
NaNO<sub>3</sub>, 20 mg/l K<sub>2</sub>HPO<sub>4</sub> x 3 H<sub>2</sub>O, 3 mg/l citric acid, 3 mg/l ferric  
ammonium citrate, 0.5 mg/l EDTA, 20 mg/l Na<sub>2</sub>CO<sub>3</sub>, 1 ml/l trace  
metals A-5, 750 ml/l artificial seawater, (for plates 10 g/l agar),  
after autoclaving: pH 8.5, addition of 20 ug/l sterile Vitamin B<sub>12</sub>

trace metals A-5: 2.86 g/l H<sub>3</sub>BO<sub>3</sub> 1.81 g/l MnCl<sub>2</sub> x 4 H<sub>2</sub>O, 0.222 g/l ZnSO<sub>4</sub> x 7 H<sub>2</sub>O,  
0.039 g/l Na<sub>2</sub>MoO<sub>4</sub> x 2 H<sub>2</sub>O, 0.079 g/l CuSO<sub>4</sub> x 5 H<sub>2</sub>O,  
0.049 g/l Co(NO<sub>3</sub>)<sub>2</sub> x 6 H<sub>2</sub>O

artificial sea water: 28.13 g/l NaCl, 0.77 g/l KCl, 1.6 g/l CaCl<sub>2</sub> x 2 H<sub>2</sub>O,  
4.8 g/l MgCl<sub>2</sub> x 6 H<sub>2</sub>O, 0.11 g/l NaHCO<sub>3</sub>, 3.5 g/l MgSO<sub>4</sub> x 7 H<sub>2</sub>O

### 3.2 Instruments

**Amersham Biosciences** (Freiburg, Germany): FPLC systems, prepacked chromatography columns

**Aventis** (Mannheim, Germany): EmulsiFlex-C5 high pressure 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): Gene Pulser Xcell electroporation system, electroporation cuvettes, Mini-Protean electrophoresis chambers, Power Pac 300

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

**Eppendorf** (Hamburg, Germany): centrifuges (5415D, 5417R), Thermomixer comfort

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

*Fuji/Raytest* (Straubenhardt, Germany): Fuji-LAS3000 luminescence and densitometry system, Image Reader Fuji-FLA2000, software AIDA v.3.52.046

*Hampton Research* (Aliso Viejo / CA, USA): vDX Plates

*Hoefler Scientific Instruments* (San Francisco, USA): SEMI-PHOR blotting system

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

*MicroCal* (Northhampton / MA, USA): MicroCal VP-ITC MicroCalorimeter

*Millipore* (Bedford / MA, USA): Amicon Ultra centrifugal filters, Steritop vacuum filters

*MWG Biotech AG* (Ebersberg, Germany): gel documentation system

*New Brunswick Scientific* (Nürtingen, Germany): Innova 4430 incubator

*Pall* (Portsmouth, England): PVDF membrane

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

*Schleicher & Schuell* (Dassel, Germany): Protran nitrocellulose membrane

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

### 3.3 Molecular biological methods

#### 3.3.1 DNA analytical methods

DNA concentrations were determined by UV spectroscopy. At a wavelength of 260 nm, one absorption unit corresponds to 50 ng/μl double stranded DNA in water. 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 *et al.* 1989).

Agarose gel electrophoresis was performed in 1 % TAE-agarose gels at 7-10 V/cm (TAE: 40 mM Tris-acetate, pH 8.5, 2 mM EDTA, supplemented with ethidium bromide). Prior to loading samples on the gel, they were mixed with 5x DNA loading dye (0.1 M EDTA, 0.1 % (w/v) SDS, 0.025 % (w/v) bromphenolblue, 40 % (v/v) glycerol).

Authenticity of cloned constructs was confirmed by sequencing either by *Medigenomix GmbH* (Martinsried, Germany) or by the *MPI* sequencing service.

### **3.3.2 Preparation and transformation of competent *E. coli* cells**

#### **3.3.2.1 Chemocompetent *E. coli* cells and chemical transformation**

Chemically competent *E. coli* BL21(DE3) cells were prepared by the RbCl-method (Hanahan 1983). 50 ml SOB medium was inoculated (1:100) with an overnight culture of the respective strain and grown at 37 °C to an OD<sub>600</sub> 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 °C. Cells were carefully resuspended in 16 ml buffer I (100 mM RbCl, 50 mM MnCl<sub>2</sub>, 30 mM KOAc, 10 mM CaCl<sub>2</sub>, 15 % (v/v) glycerol, pH 5.8 adjusted with acetic acid) and incubated on ice for 15 min. Thereafter, cells were centrifuged again and resuspended in 4 ml buffer II (10 mM MOPS, 10 mM RbCl, 75 mM CaCl<sub>2</sub>, 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 °C.

For transformation, 50 µl of chemically competent *E. coli* BL21(DE3) cells were thawed on ice and mixed with 50 ng plasmid DNA or ≤ 5 µl ligation reaction. After incubation on ice for 30 min, cells were heat shocked at 42 °C for 50 sec and put back on ice for 2 min. Next, they were diluted in 500 µl LB medium and shaken for 1 h at 37 °C. Aliquots of the reaction were plated on selective LB plates and incubated overnight at 37 °C.

#### **3.3.2.2 Electrocompetent *E. coli* cells and electroporation**

Electrocompetent *E. coli* DH5α cells were prepared as described (Ausubel *et al.* 1997). 500 ml LB medium were inoculated (1:200) with an overnight culture of the respective strain and grown at 37 °C to an OD<sub>600</sub> of 0.5-0.6. Cells were chilled on ice for 15 min and pelleted by centrifugation for 20 min at 4000 rpm and 4 °C. The pellet was resuspended in 500 µl ice-cold sterile water and centrifuged again. This washing step was repeated and cells were resuspended in 40 µl of 10 % (v/v) ice-cold glycerol. After centrifugation for 10 min at 3500 rpm and 4 °C, cells were resuspended in 500-1000 µl of ice-cold 10 % (v/v) glycerol, divided into aliquots of 50 µl, frozen on dry ice and stored at -80 °C.

To transform electrocompetent *E. coli* DH5α cells, 50 µl cells were thawed on ice, mixed with 50 ng of plasmid DNA or ≤ 5 µl ligation reaction and transferred to a chilled cuvette (0.1 cm). Immediately after applying the pulse (1250 V, 25 µF, 200 Ω; Gene Pulser

Xcell), cells were taken up in 800 µl LB medium, shaken for 1 h at 37 °C and partly spread on selective LB plates, which were incubated overnight at 37 °C.

### 3.3.2.3 TSS-transformation

For fast transformation purposes (Chung *et al.* 1990), 5 ml LB medium were inoculated with a single colony of the respective *E. coli* strain and grown at 37 °C to an OD<sub>600</sub> of ~0.4. 2 ml of the cell culture were pelleted (20800 x g for 1 min at RT) and resuspended in 100 µl TSS-solution (10 % (w/v) PEG6000, 5 % (v/v) DMSO, 50 mM MgCl<sub>2</sub>, in LB), containing 0.5 -1 µg of plasmid DNA, and incubated on ice for 30 min. Thereafter, 0.5 ml LB medium were added and the culture was shaken for 1 h at 37 °C. Aliquots were plated on selective LB plates and incubated overnight at 37 °C.

### 3.3.3 Isolation of chromosomal DNA from *Synechococcus* sp. PCC7002

*Synechococcus* sp. PCC7002 [PR-6] (ATCC Nr. 27264) was purchased from the American Type Culture Collection (ATCC). According to the ATCC product information, the cyanobacteria were plated on MN marine medium agar plates and grown at RT under permanent light. After four weeks, colonies were transferred into MN marine medium, which was continuously stirred at RT under permanent light.

After 8 days, 12 ml of the culture were harvested by centrifugation, the cell pellet was washed in 1 ml TES-buffer I (10 mM Tris-HCl, pH 8, 1 mM EDTA, 200 mM NaCl) and resuspended in 0.5 ml TE-buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA). After addition of lysozyme to a final concentration of 5 mg/ml and incubation for 50 min at 37 °C, Proteinase K and SDS were added to final concentrations of 100 µg/ml and 0.5 %, respectively, followed by incubation at 55 °C for 1 h. After two consecutive extractions with one sample volume of phenol/chloroform/isoamyl-alcohol (25:24:1) and chloroform/isoamyl-alcohol (24:1), nucleic acids were precipitated by addition of 1/10 volume of 3 M sodium acetate, pH 6 and 2 volumes of ice-cold ethanol for 30 min at -80 °C. After centrifugation, the pellet was washed with 70 % ethanol, air-dried and resuspended in 400 µl TE-buffer (20 min at 65 °C). One volume of CTAB-precipitation solution (50 mM Tris-HCl, pH 8, 100 mM EDTA, 1 % (w/v) CTAB) was added and after centrifugation the pellet was resuspended in 500 µl TES-buffer II (10 mM Tris-HCl, pH

8, 1 mM EDTA, 1 M NaCl). Nucleic acids were precipitated by addition of 0.6 volumes isopropanol and centrifugation. The pellet was washed with 70 % ethanol, air-dried and resuspended in TE-buffer (Brahamsa 1996).

### 3.3.4 Purification of plasmid DNA and DNA-fragments

Amplification of plasmid DNA was performed in *E. coli* DH5 $\alpha$ , which were grown for 8-16 h in LB medium containing the appropriate antibiotics (Sambrook *et al.* 1989). Subsequently, plasmids were purified *via* anion exchange chromatography using the QIAprep Spin Miniprep Kit, Wizard Plus SV Miniprep DNA Purification System, QIAGEN Plasmid Midi Kit or PureYield Plasmid Midiprep System according to the manufacturer's instructions.

In order to purify or isolate DNA-fragments after agarose gel electrophoresis or enzymatic reactions, either the Wizard SV Gel and PCR Clean-Up System or the QIAquick PCR Purification Kit was applied for anion exchange chromatography following the producer's recommendations.

### 3.3.5 PCR (polymerase chain reaction)

Amplification of DNA fragments was generally performed according to protocol 1 (Table 1). Only if mutations were introduced *via* whole plasmid PCR (Weiner *et al.* 1994), reactions were designed with mutation-site specific primers as listed in protocol 2 (Table 1). The applied cycling conditions are summarized in Table 2.

	<b>Protocol 1</b>	<b>Protocol 2</b>
template DNA	100-200 ng plasmid or genomic DNA	100 ng plasmid DNA
buffer	1x ThermoPol reaction buffer	1x cloned <i>Pfu</i> buffer
dNTPs	400 $\mu$ M each	250 $\mu$ M each
primers	40 pmol each	40 pmol each
Polymerase	4 U Vent DNA Polymerase	2.5 U <i>Pfu</i> DNA Polymerase
total volume	100 $\mu$ l	50 $\mu$ l

**Table 1. Standard protocols for polymerase chain reactions (PCR).**

step	purpose	temperature	duration	number of cycles
1	initial denaturation	95 °C	5 min	1 cycle
2	cycle denaturation	95 °C	30 sec	30 cycles (steps 2-4)
3	cycle annealing	45-60 °C	45 sec	
4	cycle extension	72 °C	2 min/kb	
5	final extension	72 °C	7 min	1 cycle
6	soak	4 °C	∞	1 cycle

**Table 2. Thermal cycling conditions for PCR amplification.**

### 3.3.5 Restriction digest and ligation

5-20 µg DNA (PCR-product or plasmid DNA) was incubated in reactions of 50-100 µl with 20-60 units of the respective restriction enzyme/s in the absence or presence BSA (Table A1). Digests were performed in the recommended reaction buffer and at the optimal temperature for 3 h. Digested plasmid DNA was dephosphorylated with shrimp alkaline phosphatase according to the manufacturer's instructions prior to further use. Products of restriction digests were analyzed and isolated by agarose gel electrophoresis. For ligations, ~100 ng of digested and dephosphorylated plasmid DNA and excess of respectively digested DNA-inserts were incubated with T4 DNA ligase reaction buffer and 400 U of T4 DNA ligase in a final volume of 10 µl for at least 2 h at RT or overnight at 16 °C. Afterwards, aliquots of the ligation reactions were transformed into electrocompetent *E. coli* DH5α cells.

For processing of constructs resulting from whole plasmid PCR, the desired PCR-fragments were initially excised and purified from the agarose gel. A subsequent digest with *DpnI* 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 electrocompetent *E. coli* DH5α cells.

### 3.3.7 Cloning strategies

Syn7002-*rbcL*, Syn7002-*rbcS* and Syn7002-ORF134 (encoding RbcX) were amplified by PCR from genomic DNA of *Synechococcus sp.* PCC7002 [PR-6] (ATCC Nr. 27264) using primers containing flanking *NdeI* and *BamHI* restriction sites (Table A2). The reverse primer for *rbcL* contained an additional *NheI* site between the stop codon of *rbcL* and the *BamHI* site. PCR-products were cloned between the *NdeI/BamHI* restriction sites of plasmids pET11a (Amp<sup>R</sup>), pET28b (Kan<sup>R</sup>, N-His<sub>6</sub>) or pET30b (Kan<sup>R</sup>), resulting in the following constructs: Syn7002-*rbcL*-pET11a, Syn7002-*rbcS*-pET11a, Syn7002-*rbcX*-pET11a, Syn7002-*rbcX*<sub>N-His6</sub>-pET28b and Syn7002-*rbcX*-pET30b. For the co-expression plasmids Syn7002-*rbcLX*-CoEx-pET11a/28b<sub>N-His6</sub>/30b and Syn7002-*rbcLS*-CoEx-pET11a, the *rbcL* expression cassette was extracted from Syn7002-*rbcL*-pET11a with *SphI/NheI* and introduced between the *SphI/XbaI* sites of Syn7002-*rbcX*-pET11a/28b<sub>N-His6</sub>/30b or Syn7002-*rbcS*-pET11a, in front of the ribosomal binding site of *rbcS* or *rbcX*, respectively. Thus, the coexpressed genes share the same T7-promotor for co-transcription, but have their own ribosomal binding sites for translation.

Point mutations in *rbcX* or *rbcL* were introduced *via* whole plasmid PCR with mutation-site specific primers (Weiner *et al.* 1994) and C-terminally truncated *rbcX* or *rbcL*, were cloned with modified reverse primers. The FLAG-tagged constructs Syn7002-*rbcX*<sub>N-FLAG</sub>-pET30b, Syn7002-*rbcLX*<sub>N-FLAG</sub>-CoEx-pET30b (for expression of both wildtype or mutant RbcX) and Syn7002-*rbcS*<sub>N-FLAG</sub>-pET11a were obtained by introducing the sequences for FLAG-tag (MDYKDDDDK) and a connecting spacer (AG) in front of the start codon for *rbcX* or *rbcS* *via* the forward primer.

Plasmids pUC-*rbcLS* (Smith and Tabita 2003) and pLS32 (Li and Tabita 1997) were used as templates for PCR amplification of the *rbcL*, *rbcS* and *rbcLS* genes from *Synechococcus sp.* PCC6301 or of *rbcX* from *Anabaena sp.* CA, respectively. These genes were cloned between *NdeI/BamHI*-sites of pET-vectors, resulting in the constructs Syn6301-*rbcL*-pET11a, Syn6301-*rbcS*-pET11a, Syn6301-*rbcLS*-pET11a, AnaCA-*rbcX*-pET11a, AnaCA-*rbcX*-pET30b and AnaCA-*rbcX*<sub>N-His6</sub>-pET28b. Plasmids Syn6301-*rbcL*(mutant or  $\Delta$ C)-pET11a, AnaCA-*rbcX*<sub>N-FLAG</sub>-pET30b as well as Syn7002-*rbcL*/AnaCA-*rbcX*<sub>FLAG</sub>-CoEx-pET30b were cloned as described above.

A cDNA-library of *Arabidopsis thaliana* and plasmid pSOCPN70/8.1 (encoding Cpn20 from *Spinacia oleracea*, Bertsch *et al.* 1992) were used to amplify the genes of the chaperonins AtCpn60 $\alpha$  (aa 47-586, at2g28000) or AtCpn60 $\beta$  (aa 56-600, at1g55490) and of the co-factors AtCpn20 (aa 52-253, for N-domain aa 52-150, for C-Domain aa 158-253, at5g20720), AtCpn10 (aa 44-139, at2g44650) and SoCpn20 (aa 54-255, Q02073), lacking the N-terminal transit peptide sequences as indicated. Since several genes are encoding these factors in *A. thaliana*, those genes with highest expression levels *in planta* were chosen for this study (Table A3). The genes were cloned between the *NdeI/BamHI*-sites of plasmids pET11a (Amp<sup>R</sup>), pET15b<sub>N-His6</sub> (Amp<sup>R</sup>) or between the *EheI/HindIII*-sites of pProExHTa<sub>N-His6</sub>, resulting in *atCpn60 $\alpha$ -pET11a* and *atCpn60 $\beta$ -pET11a*, *atCpn20-pET11a*, *atCpn20<sub>N-His6</sub>-pProEx*, *atCpn20/N-domain<sub>N-His6</sub>-pET15b*, *atCpn20/C-domain<sub>N-His6</sub>-pET15b*, *atCpn10-pET11a* and *soCpn20-pET11a*. Since the reverse primer for *atCpn60 $\alpha$*  contained an additional *NheI* site between the stop codon and the *BamHI* site, the *atCpn60 $\alpha$*  expression cassette could be extracted from *atCpn60 $\alpha$ -pET11a* with *SphI/NheI* and introduced between the *SphI/XbaI* sites of *atCpn60 $\beta$ -pET11a*, in front of the ribosomal binding site of *atCpn60 $\beta$* . In the resulting co-expression plasmid *atCpn60 $\alpha$  $\beta$ -CoEx-pET11a*, both *atCpn60 $\alpha$*  and *atCpn60 $\beta$*  share the same T7-promotor for co-transcription, but have their own ribosomal binding sites for translation.

All 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](http://www.ncbi.nlm.nih.gov)) and MultAlin analysis program (Corpet 1988, <http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html>).

### 3.4 Protein biochemical and biophysical methods

#### 3.4.1 Protein analytical methods

##### 3.4.1.1 Protein quantification

Protein concentrations were determined spectrophotometrically using the theoretical extinction coefficient of proteins at a wavelength of 280 nm (Stoscheck 1990). Extinction coefficients were calculated by the ProtParam tool at the ExPASy proteomics server (<http://us.expasy.org/tools/protparam.html>).

Alternatively, protein concentrations were measured colorimetrically (595 nm) by Bradford assay using the Bio-Rad protein assay reagent according to the manufacturer's recommendations (Ausubel *et al.* 1997, Bradford *et al.* 1976).

For proteins present as complexes in their native state (e.g. chaperones or RuBisCo), molar concentrations will be expressed for oligomers, unless stated otherwise. As far as denatured substrates are concerned, concentrations will be expressed for the monomers, since this is the assumed state of substrate binding to chaperones.

##### 3.4.1.2 SDS-PAGE

Proteins were analyzed under denaturing, reducing conditions by SDS-PAGE (sodiumdodecylsulfate polyacrylamide gel electrophoresis). This method allows the separation of proteins primarily according to their molecular weight and the resulting electrophoretic mobility (Laemmli 1970). Gels were cast according to Table 3. Protein samples were mixed with 2 x SDS-loading dye (14 mM Tris-HCl, pH 6.8, 5 % SDS, 20 % (v/v) glycerol, 0.1 % (w/v) bromphenolblue, 2 % (v/v)  $\beta$ -mercaptoethanol) and boiled for 5 min at 95 °C before loading onto the gels. Electrophoresis was performed in Mini-Protean electrophoresis chambers in SDS-electrophoresis buffer (25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS) at a constant voltage of 200 V.

ingredients (amounts for 4 gels)	resolution gel				stacking gel
	8 %	12.5 %	16 %	17.5 %	5 %
30 % AA / 0.8 % BisAA	4.5 ml	6.8 ml	9.0 ml	9.7 ml	1.66ml
1.875 M Tris-HCl, pH 8.8	3.5 ml	3.5 ml	3.5 ml	3.5 ml	-
0.6 M Tris-HCl, pH 6.8	-	-	-	-	1.00 ml
ddH <sub>2</sub> O	8.7 ml	6.4 ml	4.2 ml	3.5 ml	7.20 ml
10 % (w/v) SDS	167 $\mu$ l	167 $\mu$ l	167 $\mu$ l	167 $\mu$ l	100 $\mu$ l
10 % (w/v) APS	100 $\mu$ l	100 $\mu$ l	100 $\mu$ l	100 $\mu$ l	50 $\mu$ l
TEMED	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l

**Table 3.** Preparation of SDS-PAGE gels.

### 3.4.1.3 Tricine-PAGE

Protein samples were analyzed on tricine gels, if a high resolution of small proteins (5-20 kDa) was required.

Samples were taken up in 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)  $\beta$ -mercaptoethanol) and boiled for 5 min at 95 °C, before they were applied to the gels, which were prepared according to Table 4. 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  $\leq$  130 V.

ingredients (amounts for 2 gels)	resolution gel		stacking gel
	16 %		4 %
30 % AA / 0.8 % BisAA	5.35 ml		0.65 ml
3 M Tris-HCl, pH 8.45	3.35 ml		1.24 ml
ddH <sub>2</sub> O	200 $\mu$ l		3.10 ml
10 % (w/v) SDS	100 $\mu$ l		50 $\mu$ l
glycerol	1.00 ml		-
10 % (w/v) APS	100 $\mu$ l		50 $\mu$ l
TEMED	10 $\mu$ l		5 $\mu$ l

**Table 4. Preparation of Tricine-PAGE gels.**

### 3.4.1.4 Native PAGE

In Native PAGE the mobility of a protein depends on both its charge and its hydrodynamic size.

Gels were cast according to Table 5 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 glycine, pH not adjusted) at 4 °C, employing a constant voltage of 100 V for the first 30 min and 200 V throughout the end of the run.

ingredients (amounts for 2 gels)	resolution gel		stacking gel
	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
H <sub>2</sub> O	5.00 ml	3.84 ml	3.65 ml
10 % (w/v) APS	50 $\mu$ l	50 $\mu$ l	25 $\mu$ l
TEMED	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l

**Table 5. Preparation of Native PAGE gels.**

### 3.4.1.5 Bis-Tris Native PAGE

In the discontinuous Bis-Tris Native PAGE system, protein migration correlates with molecular masses of proteins and was performed according to described procedures (Hansen *et al.* 1999, Schagger and von Jagow 1991).

Gels were prepared by pouring the stacking gel on top of the non-polymerized resolution gel (Table 6). Samples were adjusted with loading dye to final concentrations of 50 mM Bis-Tris, pH 7.0, 5 % glycerol and 0.1 % bromphenolblue. Electrophoresis was performed 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. Voltage was adjusted to 150 V for the first 40 min and to 300 V for the remaining running time.

ingredients (amounts for 4 gels)	resolution gel		stacking gel		
	12.5 %	13 %	5 %	6 %	7 %
30 % AA / 0.8 % BisAA	6.24 ml	6.51 ml	2.20 ml	2.40 ml	2.80 ml
3x gel buffer (150 mM Bis-Tris, pH 7.0, 1.5 M $\epsilon$ -amino n-caproic acid)	5.01 ml	5.01 ml	4.00 ml	4.00 ml	4.00 ml
glycerol	2.37 ml	2.37 ml	-	-	-
ddH <sub>2</sub> O	1.32 ml	1.05 ml	5.68 ml	5.48 ml	5.08 ml
10 % (w/v) APS	50.1 $\mu$ l	50.1 $\mu$ l	96 $\mu$ l	96 $\mu$ l	96 $\mu$ l
TEMED	5.1 $\mu$ l	5.1 $\mu$ l	9.6 $\mu$ l	9.6 $\mu$ l	9.6 $\mu$ l

**Table 6. Preparation of Bis-Tris Native PAGE gels.**

### 3.4.1.6 SDS-PAGE analysis of Native PAGE protein bands

For SDS-PAGE analysis of proteins that were separated by native PAGE, protein bands were extracted from vacuum dried Native PAGE gels by excision and incubation at RT for 4 h to overnight in 1x 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)  $\beta$ -mercaptoethanol). Samples were boiled for 10 min at 95  $^{\circ}$ C. After centrifugation, the supernatant was mixed with 2  $\mu$ l 1 M Tris-base and subjected to SDS-PAGE analysis.

### 3.4.1.7 Coomassie blue staining of polyacrylamide gels

In order to detect protein amounts of  $\geq 0.5 \mu$ g on SDS-PAGEs, Coomassie blue staining was carried out. 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 several washes in destaining solution (20 % (v/v) ethanol, 7 % (v/v) acetic acid) to remove background staining.

#### **3.4.1.8 Silver staining of polyacrylamide gels**

The silver staining method allows a more sensitive detection of very little protein amounts ( $\geq 10$  ng). Directly 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<sub>2</sub>O, 2 g/l sodium thiosulphate x 5 H<sub>2</sub>O) 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.

#### **3.4.1.9 Autoradiography**

Minute amounts of radioactively (<sup>35</sup>S) labeled proteins could be visualized by autoradiography. 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 paper with a radioactive dye. The radioactive pattern was transferred to a phosphoimaging plate (exposure time several hours or overnight) which was read by laser scanning (Image Reader Fuji-FLA2000) and processed using AIDA software.

#### **3.4.1.10 Western blotting and immunodetection**

Immunoblotting was performed following a modified described method (Towbin *et al.* 1979). Proteins were transferred from gels (SDS-PAGE or Native PAGE) to a

nitrocellulose membrane in a semi-dry Western blotting unit (SEMI-PHOR). Native PAGE gels required incubation for 30-60 min in SDS-buffer (62.5 mM Tris, 480 mM glycine, 0.25 % (w/v) SDS) prior to blotting, in order to bind negatively charged groups to the proteins. The transfer was performed in Western blot buffer (20 % (v/v) methanol, 50 mM Tris, 192 mM glycine) at a constant current of 150 mA for 60-90 min. If not stated otherwise, blots were subsequently blocked for 1 h with 1 % (w/v) milk powder in TBST-buffer (20 mM Tris-HCl, pH 7.5, 137 mM NaCl, 0.1 % (v/v) Tween 20), washed with TBST-buffer and incubated with the primary antibody (1:2000-1:10000) in TBST-buffer for 1 h. Thereafter, washed blots were incubated with the horseradish peroxidase (HRP)-conjugated secondary antibody (1:5000-1:10000) in TBST-buffer for 1 h, followed by several washing steps. 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-aminophthalhydrazid), 400  $\mu$ M p-coumaric acid) and ECL solution II (100 mM Tris-HCl, pH 8.5, 5.4 mM H<sub>2</sub>O<sub>2</sub>). Membranes were incubated in the resulting solution and protein bands were detected and documented with the Fuji-LAS3000 luminescence and densitometry system.

#### **3.4.1.11 Generation of antiserum**

Antiserum against Syn7002-RbcX was generated in the animal facilities of the *MPI for Biochemistry*. For initial immunization, equal volumes (500  $\mu$ l) of TiterMax Classic adjuvant and protein solution (1 mg Syn7002-RbcX in 20 mM Tris-HCl, 7.5, 50 mM NaCl, 5 % (v/v) glycerol) were mixed. This water in oil emulsion was injected subcutaneously into a rabbit. In the same way, two succeeding boosts were carried out at intervals of 5 weeks. Test bleeds were taken 10 days after each injection.

For preparation of the antiserum, the immunized blood was incubated at 37 °C for 3 h, before coagulated components were removed by two successive centrifugation steps. The serum was qualitatively analyzed by SDS-PAGE and immunoblotting.

#### **3.4.1.12 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  $\mu$ l 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  $\mu$ l 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)  $\beta$ -mercaptoethanol), mixed with 1  $\mu$ l 2 M Tris-base, boiled for 5 min at 95 °C and subjected to SDS-PAGE analysis.

#### **3.4.1.13 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  $\mu$ g) 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 and Kulicke 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.

#### **3.4.1.14 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<sub>2</sub>O and air-dried. Protein bands of interests were cut and analyzed *via* Edman degradation by the *MPI* protein sequencing service.

#### **3.4.1.15 Sequence alignments**

Multiple sequence alignment with hierarchical clustering was performed using MultAlin (Corpet 1988, <http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html>).

### 3.4.2 Protein expression and purification

All protein purification steps were performed at 4 °C unless stated otherwise.

#### 3.4.2.1 AtCpn60 $\alpha\beta$ (*Arabidopsis thaliana* Cpn60 $\alpha_7\beta_7$ )

*E. coli* BL21(DE3), transformed with *atCpn60 $\alpha\beta$ -CoEx-pET11a*, were grown at 37 °C in LB medium. After reaching mid-log phase, expression of AtCpn60 $\alpha\beta$  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 AtCpn60 $\alpha_7\beta_7$  from GroEL and from AtCpn60-monomers. Fractions containing AtCpn60 $\alpha\beta$  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 AtCpn60 $\alpha\beta$  containing fractions were subsequently applied to a Hi-Trap Heparin Sepharose column, equilibrated with 20 mM Tris-HCl, pH 7.5. AtCpn60 $\alpha\beta$  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 AtCpn60 $\alpha\beta$  were pooled, concentrated (MWCO 100 kDa), flash-frozen in liquid nitrogen and stored at -80 °C. Complex concentration was determined spectrophotometrically at 280 nm ( $192150 \text{ M}^{-1}$  complex AtCpn60 $\alpha_7\beta_7$ ). Native PAGE, light scattering and functional assays confirmed purification of active AtCpn60 $\alpha_7\beta_7$ -complexes. ESI-MS and SDS-PAGE (8 % resolution) verified the absence of GroEL/GroES and the presence of equal amounts of  $\alpha$ - and  $\beta$ -subunits in the purified complexes.

#### 3.4.2.2 AtCpn60 $\beta$ (*Arabidopsis thaliana* Cpn60 $\beta$ <sub>14</sub>)

AtCpn60 $\beta$  cell lysate was prepared from *E. coli* BL21(DE3) cells harboring plasmid *atCpn60 $\beta$ -pET11a* as described for AtCpn60 $\alpha$  $\beta$ . The soluble lysate fraction was applied to a DE52(diethylaminoethyl)-cellulose column, equilibrated with buffer A / 1 mM EDTA and developed with a linear gradient from 0.02 to 1 M mM NaCl. Fractions were analyzed as described for AtCpn60 $\alpha$  $\beta$ . Those containing AtCpn60 $\beta$  were pooled, supplemented with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (20 % saturation) and applied to a Phenyl Sepharose column, equilibrated with 20 mM Tris-HCl, pH 7.5, 20 % saturation (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The column was eluted with a linear gradient from 20 to 0 % saturation (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Fractions containing AtCpn60 $\beta$  were dialyzed against 20 mM MES-NaOH, pH 6, 20 mM NaCl and loaded onto a MonoQ column, equilibrated with the same buffer. The column was eluted with a linear gradient of 0-0.95 M NaCl. Pooled fractions were dialyzed against 20 mM Tris-HCl, pH 7.5, 50 mM NaCl and applied to a Hi-Trap Heparin Sepharose column, which was equilibrated and eluted with the latter buffer. AtCpn60 $\beta$  was collected in the flow through, concentrated (MWCO 100 kDa) and applied to a Sephacryl S-300 column in buffer B. Fractions with enriched AtCpn60 $\beta$  were pooled, concentrated (MWCO 100 kDa), flash-frozen in liquid nitrogen and stored at -80 °C. Complex concentration was determined spectrophotometrically at 280 nM (215040 M<sup>-1</sup> complex AtCpn60 $\beta$ <sub>14</sub>). Native PAGE, light scattering and functional assays confirmed purification of active AtCpn60 $\beta$ <sub>14</sub>-complexes. The identity of AtCpn60 $\beta$  and the absence of GroEL/GroES in the preparation were confirmed by 8 % SDS-PAGE and ESI-MS.

#### 3.4.2.3 AtCpn60 $\alpha$ (*Arabidopsis thaliana* Cpn60 $\alpha$ )

AtCpn60 $\alpha$  was expressed and purified from *E. coli* BL21(DE3) cells, transformed with *atCpn60 $\alpha$ -pET11a*, essentially in the same way as AtCpn60 $\beta$ . The fact that AtCpn60 $\alpha$  was not assembled in tetradecameric complexes but appeared predominantly as monomers, was taken into consideration when the protein was concentrated (MWCO 30 kDa) or when gel filtration chromatography was performed (Sephacryl S-200). Protein concentration was determined spectrophotometrically at 280 nM (12090 M<sup>-1</sup> monomer AtCpn60 $\alpha$ ).

#### 3.4.2.4 AtCpn20, AtCpn10 and SoCpn20 (*Arabidopsis thaliana* Cpn20 and Cpn10, *Spinacia oleracea* Cpn20)

*E. coli* BL21(DE3) cells, transformed with *atCpn20*-pET11a, *atCpn10*-pET11a or *soCpn20*-pET11a were grown to mid-log phase at 37 °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<sup>-1</sup> tetramer Cpn20, 57750 M<sup>-1</sup> heptamer Cpn10). N-terminal sequencing confirmed that the proteolysis product of AtCpn20, 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.

#### 3.4.2.5 PsCpn20 (Cpn20 from *Pisum sativum* chloroplast stroma)

Pea chloroplast stroma was a kind gift of Dr. E. Schleiff and Dr. T. Becker. In order to remove high molecular weight complexes, the stroma was subjected to ultracentrifugation (150000 x g, 18 h). The upper supernatant fraction was applied to a DE52 ion exchange column, equilibrated with **buffer D** (50 mM Tris-HCl, pH 7.7, 10 mM KCl, 10 mM MgCl<sub>2</sub>) / 1 mM EDTA, and eluted with a linear salt gradient from 0 to 1 M NaCl. Fractions containing Cpn20 were then subjected consecutively to a Mono Q and a Hi-Trap Heparin Sepharose column, both equilibrated and eluted as the previous DE52 column. The resulting Cpn20 pool was concentrated (MWCO 30 kDa) and

applied to Sephacryl S-200 gel filtration chromatography in buffer D / 20 mM NaCl. Fractions containing Cpn20 were concentrated, flash-frozen in liquid nitrogen and stored at -80 °C. Protein concentration was determined by Bradford assay.

#### 3.4.2.6 AtCpn20<sub>N-His6</sub>, AtCpn20/N-domain<sub>N-His6</sub>, AtCpn20/C-domain<sub>N-His6</sub>

For *E. coli* BL21(DE3) cells harboring plasmid *atCpn20*<sub>N-His6</sub>-pProEx, *atCpn20/N-domain*<sub>N-His6</sub>-pET15b or *atCpn20/C-domain*<sub>N-His6</sub>-pET15b, protein expression and cell lysis was performed as described for AtCpn60 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 AtCpn20<sub>N-His6</sub>, AtCpn20/N<sub>N-His6</sub> or AtCpn20/C<sub>N-His6</sub> 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 AtCpn20<sub>N-His6</sub> or the domains 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.

#### 3.4.2.7 Syn6301-RbcL<sub>8</sub>S<sub>8</sub> (*Synechococcus* sp. PCC6301 RbcL<sub>8</sub>S<sub>8</sub>)

To increase the amount of soluble Syn6301-RbcL<sub>8</sub>S<sub>8</sub> 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 pBAD33ES/EL and grown to OD<sub>600</sub> ~ 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<sub>8</sub>S<sub>8</sub> by SDS-PAGE and immunoblotting against RbcL. The Syn6301-RbcL<sub>8</sub>S<sub>8</sub> pool was supplemented with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (20 % saturation) and applied to a Phenyl-Sepharose column, equilibrated with 50 mM Tris-HCl, pH 7.5, 20 % saturation (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.5 mM DTT. Elution was performed with a linear gradient from 20 to 0 % saturation (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Syn6301-RbcL<sub>8</sub>S<sub>8</sub> 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<sub>8</sub>S<sub>8</sub> were concentrated (MWCO 100 kDa) and passed over a Superose 6 gel filtration column in buffer B. Eluted Syn6301-RbcL<sub>8</sub>S<sub>8</sub> 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<sup>-1</sup> complex RbcL<sub>8</sub>S<sub>8</sub>). Light scattering, Native PAGE and carboxylation activity confirmed the complex nature and activity of the purified Syn6301-RbcL<sub>8</sub>S<sub>8</sub>.

#### 3.4.2.8 Syn6301-RbcL<sub>8</sub> (*Synechococcus* sp. PCC6301 RbcL<sub>8</sub>)

*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 ca. 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<sub>8</sub> 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 AtCpn60, 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, equilibrated with **buffer F** (50 mM Tris-HCl, pH 8, 50 mM NaHCO<sub>3</sub>, 10 mM MgCl<sub>2</sub>) / 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<sub>8</sub> 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<sub>8</sub> were pooled and dialyzed against 20 mM Imidazol, pH 6.5, 50 mM NaHCO<sub>3</sub>, 10 mM MgCl<sub>2</sub>, 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<sub>8</sub> containing fractions were concentrated (MWCO 100 kDa) and subjected to Superdex 200 gel filtration chromatography in in buffer F. Fractions containing Syn6301-RbcL<sub>8</sub> 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<sup>-1</sup> complex RbcL<sub>8</sub>). Light scattering and Native PAGE confirmed that the purified protein was assembled Syn6301-RbcL<sub>8</sub>, which showed carboxylation activity upon addition of RbcS.

#### **3.4.2.9 Syn6301-RbcS and Syn7002-RbcS<sub>FLAG</sub> (*Synechococcus* sp. PCC6301 RbcS and *Synechococcus* sp. PC7002 RbcS<sub>FLAG</sub> from inclusion bodies)**

RbcS was purified from inclusion bodies by modification of previously described methods (Coligan *et al.* 2000, Somerville *et al.* 1986). *E. coli* BL21(DE3) cells, transformed with Syn6301-RbcS-pET11a or Syn7002-RbcS<sub>FLAG</sub>-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. The recovery of refolded protein was ca. 65 to 85 %. Protein concentration was determined spectrophotometrically at 280 nM (19060 M<sup>-1</sup> monomer Syn6301-RbcS, 24410 M<sup>-1</sup> monomer Syn7002-RbcS<sub>FLAG</sub>).

#### **3.4.2.10 Syn7002-RbcX, AnaCA-RbcX (*Synechococcus* sp. PC7002 and *Anabaena* sp. CA wildtype/mutant/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<sub>2</sub>) 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<sup>-1</sup> dimer Syn7002-RbcX, 30440 M<sup>-1</sup> dimer AnaCA-RbcX). Recombinant mutant and FLAG-tagged RbcX from *Synechococcus* sp. PC7002 or *Anabaena* sp. CA were purified in cooperation with Karnam Vasudeva Rao and Bharathi Vasudeva Rao.

#### **3.4.2.11 (SeMet)-Syn7002-RbcX and (SeMet)-Syn7002-RbcX( $\Delta$ C25)**

Selenomethionine (SeMet) incorporation into Syn7002-RbcX followed an established protocol (Van Duyne *et al.* 1993). *E. coli* BL21(DE3) were transformed with Syn7002-RbcX-pET11a and grown overnight in LB medium at 30 °C. The cell pellet of 1 ml

overnight culture was resuspended in 1 ml M9 minimal medium (containing 0.4 % (w/v) glucose and 20 mg/l thiamine), diluted into one liter of the latter medium and grown at 25 °C to an OD<sub>600</sub> of ~0.6. Then amino acids were added (lysine, phenylalanine and threonine 100 mg/l, L-selenomethionine 60 mg/l, isoleucine, leucine and valine 50 mg/l) and after 15 min growth, protein expression was induced with 1 mM IPTG for 10 h at 25 °C. Purification of (SeMet)-Syn7002-RbcX was performed essentially the same way as for Syn7002-RbcX, except that all buffers were augmented with 1 mM DTT and gel filtration was performed in 10 mM Tris-HCl, pH 7.5, 1 mM DTT.

(SeMet)-Syn7002-RbcX( $\Delta$ C25) was produced by incubating 1 mg/ml (SeMet)-Syn7002-RbcX with 25  $\mu$ g/ml subtilisin for 2 h in ice-chilled 20 mM HEPES-NaOH, pH 7.5, 50 mM NaCl. Subsequently, the sample was applied to a Superdex 200 gel filtration column in order to remove the subtilisin and to exchange the buffer to 10 mM Tris-HCl, pH 7.5. Truncated Syn7002-RbcX was identified as Syn7002-RbcX( $\Delta$ C25) by Edman degradation (N-terminal sequencing) and ESI-MS.

#### **3.4.2.12 Syn7002-RbcLX<sub>N-His6</sub> (*Synechococcus* sp. PCC7002 RbcLX<sub>N-His6</sub>-complex)**

*E. coli* BL21(DE3) cells were transformed with plasmid Syn7002-*rbcLX<sub>N-His6</sub>*-CoExpET28b. 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 AtCpn20<sub>N-His6</sub>. The complex of Syn7002-RbcLX<sub>N-His6</sub> as well as free dimers of Syn7002-RbcX<sub>N-His6</sub> 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<sub>N-His6</sub>-complexes, whereas a second peak contained only free Syn7002-X<sub>N-His6</sub> 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.

### 3.4.2.13 Syn6301-RbcL/AnaCA-RbcX<sub>N-His6</sub> (complex of *Synechococcus sp.* PCC 6301 RbcL and *Anabaena sp.* CA RbcX<sub>N-His6</sub>)

*E. coli* BL21(DE3) cells, transformed with pG-*KJE8*, Syn6301-*rbcL*-pET11a and AnaCA-*rbcX<sub>N-His6</sub>*-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<sub>N-His6</sub> was induced with 1 mM IPTG for ca. 3 h. Soluble cell lysate was produced and applied to Ni-NTA-Agarose as described above for Syn7002-RbcLX<sub>N-His6</sub>. The resulting pool of Syn6301-RbcL/AnaCA-RbcX<sub>N-His6</sub>-complexes and free dimers of AnaCA-RbcX<sub>N-His6</sub> (eluted at 250 mM imidazole), was dialyzed against 20 mM Tris-HCl, pH 9, 50 mM NaHCO<sub>3</sub>, 10 mM MgCl<sub>2</sub> 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<sub>N-His6</sub>-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.

## 3.4.3 Functional analyses

### 3.4.3.1 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<sub>2</sub>). If the influence of co-chaperones was analyzed, the latter were added at a final concentration of 1 µM. Given the case that the activity dependence on different salts or salt concentrations had to be tested, the buffer was modified as indicated. 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<sub>2</sub>O and 4.2 % (w/v) ammonium molybdate in 4

N HCl; supplemented before use with 0.1 % (v/v) Triton X-100) and 40  $\mu$ l 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  $\mu$ l of solutions containing 0, 10, 25, 50, 100, 250, 500 and 1000  $\mu$ M  $K_2HPO_4$  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.

#### **3.4.3.2 Aggregation prevention assay**

Proteins for aggregation analysis (Rr-RbcL, Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub>) were denatured in **assay-denaturation-buffer 1** (20 mM MOPS-KOH, pH 7.5, 6 M GdnHCl, 100 mM KCl, 5 mM  $MgCl_2$ , 1 mM EDTA, 10 mM DTT) for 60 min at 25 °C. Afterwards, they were diluted 100-fold (final concentration 400 nM, RbcL monomer) into assay-buffer 1 containing GroEL at the indicated concentrations. Immediately, protein aggregation was monitored spectrophotometrically at 320 nm for 20 min at 25 °C.

#### **3.4.3.3 MDH refolding assay**

50  $\mu$ M MDH was denatured in **assay-denaturation-buffer 2** (20 mM Tris-Cl, pH 7.6, 3 M GdnHCl, 100 mM KCl, 5 mM  $MgCl_2$ , 5 mM DTT) for 1 h at 37 °C. It was diluted 50-fold into **assay-buffer 2** (20 mM Tris-Cl, pH 7.6, 100 mM KCl, 10 mM  $MgCl_2$ , 10 mM DTT, 0.18 mg/ml BSA) containing 2  $\mu$ M chaperonins and 2  $\mu$ M co-chaperones as indicated. Reactions were started by addition of 2 mM ATP and incubated at 37 °C. At several time points, samples of 10  $\mu$ l were taken and immediately stopped with CDTA (65 mM). To determine MDH activity, a sample (12  $\mu$ l) was diluted into 446  $\mu$ l of reaction mixture (assay buffer 2 containing 1 mM CDTA, 1.1 mg/ml BSA, 0.56 mM oxalacetic acid, 0.22 mM  $NADH_2$ ). After fast but thorough mixing and transfer into cuvettes, the change of absorbance at 340 nm at 25 °C was measured over 30 sec. The slope of the resulting curve represented the relative MDH-activity (Hayer-Hartl 2000).

#### 3.4.3.4 RuBisCO refolding assay

RuBisCO was denatured in assay-denaturation-buffer 1 at 25 °C for 1 h, before it was diluted 100-fold into ice-cold assay-buffer 1, containing 5 mM DTT, 1 mg/ml BSA and chaperonin. Reactions were incubated for 5 min at 25 °C, before co-chaperone was added. Refolding was started by addition of 2 mM ATP and carried out at 25 °C. Samples were taken at indicated time points and immediately stopped with glucose (10 mM) and hexokinase (2.5 U). The applied substrate and chaperone concentrations are indicated in the respective figure legends.

Apart from this standard refolding assay, numerous experiments for refolding of Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> were performed including a range of concentrations and combinations of different chaperonins (GroEL, AtCpn60αβ, AtCpn60β), co-factors (GroES, mt-Hsp10, Gp31, AtCpn20, SoCpn20, PsCpn20, AtCpn10), DnaK/DnaJ/GrpE or/and different species of RbcS and RbcX. Moreover, refolding was tested with several concentrations of differently denatured substrate, including guanidinium-, urea- or acid-denatured RuBisCO. Additionally, the effect of different temperatures, buffers (e.g. Tris, MOPS, ethanolamine, RTS-buffer), pH-values (adjusted e.g. with NaOH, KOH, HCl, Acetic acid), salts (e.g. NaCl, KCl, NaOAc, KOAc, MgCl<sub>2</sub>, Mg(OAc)<sub>2</sub>, ammoniumsulfate), the presence of soluble *E. coli* lysate, BSA, casein, chaotropic agent arginine, RuBisCO substrate RuBP, crowding reagents (e.g. dextran, ficoll, PEGs, TMAO), redox reagents or redox pairs (DTT, β-mercaptoethanol, GSH/GSSG) in different concentrations and combinations was tested. Applied experimental details, which deviate from the standard assay, will be denoted in the figure legends.

If intermediates or products of refolding were analyzed by gel filtration chromatography, reaction samples were applied to a Superdex 200 column, equilibrated and eluted in assay- buffer 1. Fractions were analyzed by SDS-PAGE, followed by Coomassie blue staining and/or immunoblotting against RbcL.

For determination of carboxylation activity, samples resulting from refolding of Rr-RbcL or Syn6301-RbcLS could be used directly, whereas samples containing Syn6301-RbcL were supplemented with 7 μM Syn6301-RbcS and assembly allowed to proceed for 5 min at RT. Thereafter, a <sup>14</sup>C-Mix (in 100 mM Tris-HCl, pH 7.5, 10 mM KCl) was added

to give final concentrations of 60 mM NaHCO<sub>3</sub>, 0.5 μCi NaH<sup>14</sup>CO<sub>3</sub> and 10 mM MgCl<sub>2</sub>. 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, Viitanen *et al.* 1995, Goloubinoff *et al.* 1989a,b).

#### 3.4.3.5 Carboxylation activity assay for *E. coli* lysates

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)<sub>2</sub>) containing 1 mM DTT and 2 μM BSA. Samples containing only RbcL were supplemented with 7 μM Syn6301-RbcS or Syn7002-RbcS<sub>FLAG</sub> and assembly was allowed to proceed for 5 min at RT. The subsequent procedure was as described above (chapter 3.4.3.4).

#### 3.4.3.6 Proteinase K protection assay

Rr-RbcL or Syn6301-RbcL<sub>8</sub> was denatured in assay-denaturation-buffer 1 for 60 min at 25 °C. The denatured protein was diluted 100-fold (final concentration 300 nM, RbcL monomer) into ice-cold assay-buffer 1 containing GroEL (300 nM). Reactions were immediately incubated for 5 min at 25 °C, before aggregated material was removed by centrifugation. Where indicated, reactions were then supplemented with GroES (1 μM) and incubated for further 5 min, followed by supplementation with nucleotides (4 mM ATP or AMP-PNP). Digestion was initiated by addition of Proteinase K at a final concentration of 2.5 μg/ml at 25 °C. At indicated times, samples of 20 μl were removed and the digestion was stopped with 1 mM PMSF. Samples were subjected to SDS-PAGE and immunoblotting against RbcL.

#### 3.4.3.7 Cycling of RbcL on GroEL

Syn6301-RbcL<sub>8</sub> or Rr-RbcL were denatured in **low salt assay-denaturation-buffer** (20 mM MOPS-KOH, pH 7.5, 6 M GdnHCl, 10 mM KCl, 1 mM EDTA, 10 mM DTT) for 60 min at 25 °C. The denatured protein was diluted 100-fold (0.5 μM, RbcL monomer) into

ice-cold **low salt assay-buffer** (20 mM MOPS-KOH, pH 7.5, 10 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT) containing 0.5 μM GroEL. After incubation for 10 min at 25 °C, aggregated material was removed by centrifugation. To the resulting solution of GroEL-bound RbcL either ATP (2 mM), GroEL-SR1 (2 μM) and GroES (3 μM) were added with a time lag of 15 sec and 60 sec, respectively, or GroEL-SR1 (2 μM) was added alone. 50 μl of the each reaction were subjected to mass determination by FFF-MALS. Collected fractions were TCA-precipitated, analyzed by SDS-PAGE and consecutive immunoblotting against RbcL.

#### **3.4.3.8 *In vivo* co-expression of RbcL or RbcLS with chaperones in *E. coli***

*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 or with pG-*Tf2* for GroEL/GroES/TF 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 ca. 3.5 h at 30 °C with or without prior transient overexpression of GroEL/GroES or GroEL/GroES/TF (induced with 20 ng/ml tetracycline) and/or DnaK/DnaJ/GrpE (induced with 0.4 % (w/v) arabinose) for 2 h at 30 °C, followed by a shift 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<sub>2</sub>, 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. RuBisCO carboxylation activity in the in the soluble lysate fractions was measured as described above (chapter 3.4.3.5).

#### **3.4.3.9 *In vitro* translation of RuBisCO**

*In vitro* translation of Syn7002-RbcL was performed in reconstituted *E. coli* lysate (RTS100 *E. coli* HY Kit) in the presence of 8 ng/μl Syn7002-*RbcL*-pET11a, 0.5 U/μl

RNAguard ribonuclease inhibitor, Complete protease inhibitor cocktail, 6.7 % (v/v)  $^{35}\text{S}$ -methionine, 50  $\mu\text{M}$  unlabelled methionine, GroEL (0.5  $\mu\text{M}$ ), GroES (1  $\mu\text{M}$ ) and Syn7002-RbcX (40  $\mu\text{M}$  dimer). Translation was carried out at 30  $^{\circ}\text{C}$  and stopped by addition of Chloramphenicol (Cam, 200  $\mu\text{g}/\text{ml}$ ) on ice. Post-translational addition of proteins was performed after Cam addition, followed by transfer of reactions back to 30  $^{\circ}\text{C}$ . When indicated, immunodepletion of GroEL from the lysate was achieved by incubation with polyclonal GroEL antibody bound to Protein A sepharose beads by gentle shaking for 45 min at 4  $^{\circ}\text{C}$  and removal of beads by centrifugation. Successful depletion of GroEL was confirmed by immunoblotting. Reactions were separated into soluble and insoluble fractions by centrifugation (20800 x g for 30 min at 4  $^{\circ}\text{C}$ ). The former was analyzed by native PAGE and the latter by SDS-PAGE with subsequent autoradiography (Agashe *et al.* 2004).

In pulse-chase experiments translation was performed in presence of 0.5  $\mu\text{M}$  GroEL, 1  $\mu\text{M}$  GroES and 40  $\mu\text{M}$  Syn7002-RbcX or AnaCA-RbcX (as indicated) for 6 min at 30  $^{\circ}\text{C}$ , before  $^{35}\text{S}$ -methionine was added. After 6 min, reactions were chased by addition of 3 mM 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. Where indicated, *in vitro* translation studies were performed by Karnam Vasudeva Rao.

#### 3.4.3.10 Co-immunoprecipitation

Cell lysate for co-immunoprecipitation was produced by incubating cells in lysis buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 5 mM  $\text{MgCl}_2$ , 1 mM EDTA, 0.1 % Triton X-100, 0.05 mg/ml lysozyme) for 30 min on ice, followed by ultrasonication and removal of cell debris by centrifugation.

RbcX<sub>FLAG</sub> was immunoprecipitated using EZview Red ANTI-FLAG M2 Affinity Gel. The gel beads were constantly kept on ice or at 4  $^{\circ}\text{C}$ , unless stated otherwise, and all centrifugation steps were performed at 8200 x g for 30 sec. Before use, beads were washed twice by gentle vortexing with 500  $\mu\text{l}$  TBS (50 mM Tris-HCl, pH 8, 150 mM NaCl). In general, soluble *E. coli* lysate (250  $\mu\text{l}$ ) and/or indicated amounts of purified

proteins (e.g. RbcS, RbcX, RbcL<sub>8</sub> or RbcL<sub>8</sub>S<sub>8</sub>, bead-bound LX<sub>FLAG</sub>-complexes) were mixed with 40-50 µl of the gel bead suspension and the final reaction volume was brought to 1 ml by the addition of lysis buffer lacking lysozyme. Pull-downs were carried out with gentle rotation for indicated times and at indicated temperatures. Afterwards, beads were washed three times for 5 min at 4 °C under gentle rotation with 500 µl TBS and the bound proteins were eluted. For elution under denaturing conditions, beads were boiled with an equal volume of 2 x SDS sample buffer lacking reducing agents (125 mM Tris-HCl, pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 0.004 % (w/v) bromphenolblue) for 5 min at 95 °C. Native elution of proteins was performed by incubating the beads for 30 min on ice with 100 µl of 3x FLAG peptide solution (300 ng/µl). Eluted proteins were separated from the beads by centrifugation and subjected to further analysis as indicated.

#### **3.4.3.11 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 or *in vitro* assembly) was performed using a Superdex 200 (10/30) gel filtration column, equilibrated in 50 mM Tris-HCl, pH 8, 50 mM NaCl, 5 mM MgCl<sub>2</sub>. 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<sub>FLAG</sub> or GroEL by SDS-PAGE and Coomassie staining or immunoblotting.

#### **3.4.3.12 Crosslinking**

Purified proteins (0.5 mg/ml) were crosslinked in presence of 1 mM DSP in 20 mM HEPES-NaOH, pH 7.5, 25 mM NaCl for 30 min at 25 °C. Reactions were stopped by addition of Tris-HCl, pH 7.5, at a final concentration of 20 mM. Samples were taken up in SDS-loading dye lacking reducing agents and analyzed by SDS-PAGE.

#### 3.4.3.13 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.

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 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 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 of the membrane, the latter was stripped after each single use according to the manufacturer's instructions. For this purpose, 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). Afterwards, the membrane was incubated in 10x PBS-buffer (92 mM Na<sub>2</sub>HPO<sub>4</sub> x 12 H<sub>2</sub>O, 16 mM NaH<sub>2</sub>PO<sub>4</sub> x H<sub>2</sub>O, 1500 mM NaCl, pH 7.2 adjusted with NaOH) and in TBST-buffer, both three times for 20 min at RT.

#### 3.4.3.14 Isothermal Titration Calorimetry (ITC)

Calorimetric measurements were performed using a MicroCal VP-ITC MicroCalorimeter. The reference cell was filled with water and all solutions were degassed by stirring under vacuum before use. Syn7002-RbcX (327 µM dimer) or AnaCA-RbcX (323 µM dimer), previously dialyzed against buffer G, were titrated with 2 mM Syn7002-RbcL-peptides (peptide 1: EIKFEFD, peptide 2: KEIKFEFDTVD)

dissolved in the buffer used for dialysis of RbcX. 36 identical aliquots (8  $\mu$ l) of peptide solution were injected at intervals of 3 min from a syringe into the sample cell containing RbcX (1.345 ml). Throughout the experiment, solutions in the sample cell were stirred by the injection syringe stirrer paddle at 315 rpm at 22 °C and the reference power was set to 30  $\mu$ Cal/sec. Control experiments to determine the heat of peptide dilution were carried out with the same experimental parameters by injecting the peptide solution into the cell containing buffer without RbcX. These heats of dilution were subtracted from the data of corresponding titrations including RbcX. Data analysis by peak integration and curve fitting by least-squares deconvolution was performed using the instrumental MicroCal Origin software to give the binding constant ( $K$ ,  $M^{-1}$ ), the enthalpy change ( $\Delta H$ , cal/mole), the entropy ( $\Delta S$ , cal/mole/deg) of the reaction and the stoichiometry ( $N$ ), which indicates the number of peptide binding sites per RbcX (VP-ITC MicroCal Manual: [www.biochem.emory.edu/labs/acorbe2/protein/ITC\\_users\\_manual.doc](http://www.biochem.emory.edu/labs/acorbe2/protein/ITC_users_manual.doc); Pierce *et al.* 1999, Wiseman *et al.* 1989).

### **3.4.4 Crystallography and structure analysis**

#### **3.4.4.1 Analytical subtilisin digest of Syn7002-RbcX**

Syn7002-RbcX was digested at a concentration of 1 mg/ml by different concentrations of subtilisin (0-1000  $\mu$ g/ml) in 20 mM HEPES-NaOH, pH 7.5, 50 mM NaCl on ice. Samples were withdrawn after 60 min, respectively, and supplemented with PMSF at a final concentration of 10 mM to stop the digest. Samples were analyzed on 16 % Tricine-PAGE. Truncated Syn7002-RbcX was identified as Syn7002-RbcX( $\Delta$ C25) by Edman degradation (N-terminal sequencing) and ESI-MS.

#### **3.4.4.2 Tryptophan-fluorescence spectroscopy**

In order to monitor tryptophan-fluorescence of RbcL in refolding reactions, Rr-RbcL or Syn6301-RbcL<sub>8</sub> was denatured in assay-denaturation-buffer 1 for 60 min at 25 °C. The denatured protein was diluted 100-fold (0.25  $\mu$ M RbcL monomer) into ice-cold assay-buffer 1 containing 0.625  $\mu$ M GroEL. After incubation for 5 min at 25 °C, 1.25  $\mu$ M GroES was added. After 2 min, refolding was started by supplementing the reaction with 2 mM ATP and incubation at 25 °C. If only native or denatured substrate or merely

substrate-binding to GroEL had to be analyzed, reactions were modified accordingly. Samples were withdrawn after indicated times and reactions were stopped with 10 mM CDTA. Tryptophan-fluorescence was measured immediately using Spex Fluorolog 3 with the following parameters: excitation 295 nm, emission scan 310-400 nm, incr. 1 nm, interval time 1 sec, slits 5/5. Background fluorescence of chemically identical reactions lacking RuBisCO was subtracted in each assay (Martin *et al.* 1991).

#### **3.4.3.3 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 5  $\mu$ M ANS. Fluorescence was measured with Spex Fluorolog 3 applying the following parameters: excitation 390 nm, emission scan 420-550 nm, incr. 1 nm, interval time 0.5 sec, slits 5/5. Data were corrected for background fluorescence (Martin *et al.* 1991, Wildner 1976).

#### **3.4.4.4 Protein crystallization**

##### **Syn7002-RbcX**

Prior to crystallization, RbcX proteins were transferred into 10 mM Tris-HCl, pH 7.5 by Superdex 200 gel filtration chromatography. Syn7002-RbcX crystals were grown using the hanging drop vapor diffusion method at 20 °C by mixing 1  $\mu$ l protein sample at ~28 mg/ml and 1  $\mu$ l reservoir solution. Initial rod-shaped crystals were obtained overnight from Crystal Screen I (Hampton Research, USA), condition 7 (0.1 M sodium cacodylate, pH 6.5, 1.4 M sodium acetate trihydrate) and condition 29 (0.1 M HEPES-NaOH, pH 7.5, 0.8 M potassium sodium tartrate tetrahydrate). The final optimized crystallization solution contained 0.1 M HEPES-NaOH, pH 7.5 and 1.3-1.55 M sodium acetate. Crystallization of Syn7002-RbcX(Y17A/Y20L, Q29A, R70A) mutant proteins required higher concentrations of 1.5-2.5 M sodium acetate and was performed by Bharathi Vasudeva Rao. For cryo-protection, the crystals were transferred stepwise into mother liquor containing 30 % (v/v) glycerol.

Heavy-atom derivatives were obtained by incubating Syn7002-RbcX crystals with 0.2 mM  $K_2[PtCl_4]$  in mother liquor for 22 h. In order to remove excess of heavy atoms and to

achieve cryo-protection, crystals were then back-soaked in mother liquor supplemented with 20 % (v/v) glycerol for 10-20 min.

#### **Syn7002-RbcX / Syn7002-RbcL peptide**

Crystals of Syn7002-RbcX were incubated for 1 h in buffer (0.1 M HEPES-NaOH, pH 7.5, 1.4 M sodium acetate) containing 1 mM Syn7002-RbcL peptide, EIKFEFD. For cryo-protection, the crystals were transferred stepwise into mother liquor containing 1 mM peptide and 20 % (v/v) glycerol.

For co-crystallization trials, a mixture of Syn7002-RbcX (in 10 mM Tris-HCl, pH 7.5) and Syn7002-RbcL peptide (EIKFEFD (in H<sub>2</sub>O) at a molar ratio of 1:1.5 (467 μM RbcX-dimer, 700.5 μM peptide) was used and crystallization performed as above. Well-diffracting RbcX crystals of space group *I*222 were obtained from PEG / Ion Screen condition 14 (0.2 M potassium thiocyanate, 20 % (w/v) polyethylene glycol 3350) (Hampton Research, USA). These crystals apparently contained no bound peptide.

#### **(SeMet)-Syn7002-RbcX(ΔC25)**

(SeMet)-Syn7002-RbcX(ΔC25) was used at 2.8 mg/ml in crystallization trials. Small rhombic crystals grew overnight in Crystal Screen I condition 22 (0.1 M Tris-HCl, pH 8.5, 0.2 M sodium acetate trihydrate, 30 % (w/v) PEG 4000) and in Crystal Screen II (Hampton Research, USA), conditions 36 (0.1 M HEPES-NaOH pH 7.5, 4.3 M NaCl) and 41 (0.1 M Tris-HCl, pH 8.5, 0.01 M Nickel (II) chloride hexahydrate, 1 M lithium sulfate monohydrate). For structure analysis only crystals from condition 22 were used. Prior to cryo-cooling, they were incubated in cryoprotectant buffer for 10 min (0.085 M Tris-HCl, pH 8.5, 0.17 M sodium acetate trihydrate, 25.5 % (w/v) PEG 4000, 15 % (v/v) glycerol).

#### **AnaCA-RbcX**

Crystals of AnaCA-RbcX<sub>His6</sub> were produced in crystallization trials of the complex of Syn6301-RbcL/AnaCA-RbcX<sub>His6</sub> using a Syn6301-RbcL/AnaCA-RbcX<sub>His6</sub> protein sample at 8.3 mg/ml. Rod-shaped crystals were obtained together with massive precipitate, presumably containing mainly Syn6301-RbcL, after several days at 20°C

from Crystal Screen I condition 11 (0.1 M sodium citrate tribasic dihydrate, pH 5.6, 1 M ammonium phosphate monobasic). Crystals were cryo-protected by stepwise transfer into mother liquor containing 35 % (v/v) glycerol.

#### 3.4.4.5 Structure determination

Structure determination was primarily performed by Dr. Andreas Bracher as follows.

Diffraction data were collected at beamlines BM14 and ID14-4 at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, and at beamline X10SA-PX-II at the Swiss Synchrotron Light Source (SLS) in Villigen, Switzerland. Diffraction data were integrated with Mosflm (Leslie 1992) and Scala (Evans 1997).

The structure of Syn7002-RbcX was solved by SIRAS using crystals incubated with 0.2 mM  $K_2(PtCl_4)$  in mother liquor for 22 h. Six Platinum sites were found by Patterson methods using Solve (Terwilliger and Berendzen 1999). Density modification was carried out with Resolve (Terwilliger 2000). The resulting map was readily interpretable, and continuous backbone density for all six RbcX chains between residues 4 and 109 could be traced with O (Jones *et al.* 1991). Subsequent iterative model improvement and refinement were performed with O and Refmac5 (Murshudov *et al.* 1997).

The structure of SeMet-labeled Syn7002-RbcX( $\Delta$ C25) was solved initially by molecular replacement (MR) using a dimer of RbcX as a search model employing the program Molrep (Vagin and Isupov 2001). The MR-model phases were used to calculate a map using the anomalous differences from the Se K-edge peak dataset, in which 29 Se-sites were located. SHARP was used for refinement of the Se positions, yielding 8 additional Se-sites, and for phase calculation (de la Fortelle and Bricogne 1997). Density modification was performed with Resolve (Terwilliger 2000). The resulting phases were used as restraints during initial model-building and refinement cycles using O and Refmac.

The structure of AnaCA-RbcX was solved by molecular replacement using Molrep with a monomer of Syn7002-RbcX as search model. Model bias was reduced by automated model building with ArpWarp6.1 (Perrakis *et al.* 1999). The ArpWarp-model was completed manually using O and refined with Refmac.

The high resolution crystal form of Syn7002-RbcX was solved by molecular replacement

using Molrep (Vagin and Isupov 2001). This structure and the RbcX mutant structures were edited with Coot (Emsley and Cowtan 2004) and refined with Refmac.

In all structures, residues facing solvent channels without detectable side chain density were modeled as alanines. All models obey good stereochemistry with no Ramachandran outliers as judged by the program Procheck (Laskowski *et al.* 1993). Coordinates were aligned with Lsqman (Kleywegt and Jones 1994). The similarity score was calculated with ESPript using the Rissler scoring matrix (Gouet *et al.* 1999). Figures were generated with the programs Molscript (Kraulis 1991), Bobscrip (Esnouf 1997), Raster-3D (Merrit and Bacon 1997), and Pymol 11 (DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web (<http://www.pymol.org>)).

## 4 RESULTS

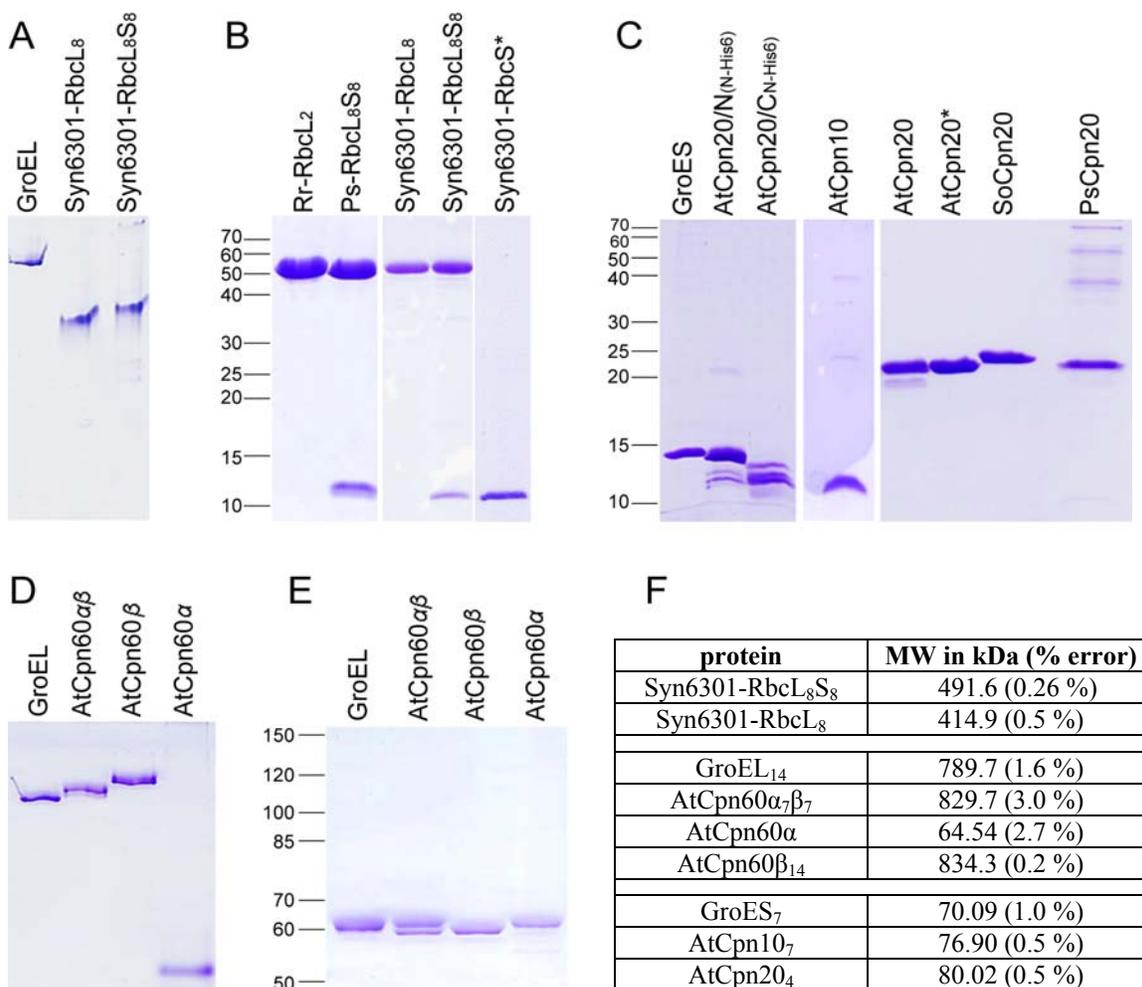
### 4.1 *In vitro* analysis on reconstitution of type I RuBisCO

#### 4.1.1 Characterization of RuBisCO as folding substrate

The underlying interest of the *in vitro* assays described in this chapter (4.1) was focused on the particular nature of refolding and assembly of hexadecameric type I RuBisCO from *Synechococcus* sp. PCC6301 (Syn6301-RbcL<sub>8</sub>S<sub>8</sub>). As mentioned in the introduction, type I RuBisCO has never been successfully reconstituted *in vitro* so far. However, the expression of Syn6301-RuBisCO in *E. coli* results in limited amounts of soluble, active enzyme whose folding has been shown to be assisted by the GroEL/GroES and DnaK/DnaJ/GrpE chaperone systems (Checa and Viale 1997, Goloubinoff *et al.* 1989a). Consequently, both Syn6301-RbcL<sub>8</sub>S<sub>8</sub> holoenzyme and the Syn6301-RbcL<sub>8</sub> cores could be purified as soluble complexes after co-expression with chaperones in *E. coli*. The small subunits (RbcS) could be refolded from inclusion bodies (Fig. 16 A, B and F). The purified RuBisCO complexes or RbcS subunits were used as substrates for the reconstitution assays described below. Since the homodimeric type II enzyme from *Rhodospirillum rubrum* (Rr-RbcL<sub>2</sub>, Fig. 16 B) can be folded and assembled *in vitro* upon interaction with GroEL/GroES (Brinker *et al.* 2001, Goloubinoff *et al.* 1989b), it served as a positive control in all refolding experiments.

The purified Syn6301-RbcL<sub>8</sub>S<sub>8</sub> holoenzyme showed carboxylation activity upon interaction with the sugar substrate RuBP, whereas the Syn6301-RbcL<sub>8</sub> core complexes acquired enzymatic activity only upon supplementation with Syn6301-RbcS subunits. In contrast to the large subunits, the RbcS subunits fold and assemble on RbcL<sub>8</sub> core complexes spontaneously upon dilution from denaturant (Andrews 1988, Andrews and Lorimer 1985, Gatenby and Ellis 1990, Lee and Tabita 1990). Hence, they could be added either as native protein or from denaturant to purified Syn6301-RbcL<sub>8</sub> cores to give the active Syn6301-RbcL<sub>8</sub>S<sub>8</sub> holoenzyme. Thus, the limiting step in *in vitro* reconstitution of type I RuBisCO is not the assembly of small subunits on RbcL<sub>8</sub> cores, but the folding and assembly of the large subunits. For the majority of assays, in which

purified small subunits had to be added, Syn6301-RbcS was used, as it is functionally compatible with RbcL of different phylogenetic origin (see Fig. 48).



### Figure 16. Purified proteins.

All proteins were expressed w/o chloroplast transit peptide or other tags, unless stated otherwise.

(A) GroEL<sub>14</sub> (*Escherichia coli*), RbcL<sub>8</sub> cores and RbcL<sub>8</sub>S<sub>8</sub> holoenzyme (*Synechococcus* sp. PCC6301) purified after overexpression in *E. coli* and analyzed by 6 % Native PAGE.

(B) SDS-PAGE (17.5 %) analysis of purified RbcL<sub>2</sub> RuBisCO from *Rhodospirillum rubrum* (Rr), RbcL<sub>8</sub>S<sub>8</sub> holoenzyme from *Pisum sativum* (pea) as well as RbcL<sub>8</sub>S<sub>8</sub>, RbcL<sub>8</sub> cores and single RbcS small subunits from *Synechococcus* sp. PCC6301 (Syn6301). While pea RuBisCO was purified from chloroplast stroma, the other proteins resulted from recombinant expression in *E. coli* and were extracted as soluble proteins. Only RbcS subunits (+) were purified from inclusion bodies. Molecular weight standards are indicated.

(C) Purified co-chaperones analyzed by 17.5 % SDS-PAGE. GroES<sub>7</sub> (*E. coli*), His-tagged N- and C-terminal domain of AtCpn20 (*A. thaliana*) showing degradation; AtCpn10 (*A. thaliana*), AtCpn20 (*A. thaliana*) with a faster migrating C-terminal degradation product (confirmed by Edman degradation), AtCpn20\* (*A. thaliana*) purified via N-terminal His-tag which was cut off, SoCpn20 (*Spinacia oleracea*) and PsCpn20 (*Pisum sativum*). All co-chaperones were expressed in *E. coli* and purified as soluble proteins; only PsCpn20 was extracted from pea chloroplast stroma.

(D) Purified GroEL<sub>14</sub> (*E. coli*), AtCpn60α<sub>7</sub>β<sub>7</sub>, AtCpn60β<sub>14</sub> and monomeric AtCpn60α (*A. thaliana*), resulting from overexpression in *E. coli* and analyzed by 6 % Native PAGE.

(E) Proteins as in (D) were analyzed by 8 % SDS-PAGE. Molecular weight standards are indicated.

(F) Molecular weight of indicated proteins as determined by light scattering with precedent field flow fractionation or gel filtration (TSK Super 3000 SW), in cooperation with Dr. M. Hayer-Hartl.

#### 4.1.2 Structural characterization of chloroplast chaperonins and co-chaperones used for *in vitro* studies

Syn6301-RbcL<sub>8</sub>S<sub>8</sub> can be produced as soluble and active enzyme in *E. coli*, assisted by GroEL/GroES (Goloubinoff *et al.* 1989a). In view of this fact, *in vitro* refolding experiments for cyanobacterial RuBisCO were performed with the bacterial chaperonin system. Alternatively, the chaperonin homologs from plant chloroplast were used for comparison with the bacterial chaperonin system. Therefore, the chloroplast Cpn60 $\alpha\beta$ , Cpn60 $\alpha$  and Cpn60 $\beta$  from *Arabidopsis thaliana* (At) as well as the co-chaperones Cpn10 or Cpn20 from *A. thaliana*, *Spinacia oleracea* (So) or *Pisum sativum* (Ps) were purified after expression in *E. coli*.

Cpn60 from *A. thaliana* (AtCpn60) was observed to form assembled chaperonin-complexes only when either both  $\alpha$ - and  $\beta$ -subunits or merely  $\beta$ -subunits were expressed. After purification, these complexes showed sizes corresponding to tetradecameric structures (Fig. 16 D and F) and they were composed of equal amounts of  $\alpha$ - and  $\beta$ -subunits (AtCpn60 $\alpha_7\beta_7$ , Fig. 16 E) or only  $\beta$ -subunits (AtCpn60 $\beta_{14}$ ), without detectable incorporation of GroEL subunits, as confirmed by mass spectrometry. On the contrary,  $\alpha$ -subunits were not competent for assembly on their own (Fig. 16 D and F). These observations were similarly described for the homologous proteins from *Pisum sativum* or *Brassica napus* (Cloney *et al.* 1992a, b, Dickson *et al.* 2000, Nishio *et al.* 1999). For this reason, only AtCpn60 $\alpha\beta$  and AtCpn60 $\beta$  (but not AtCpn60 $\alpha$ ) as well as GroEL were used as chaperonins for the refolding analyses described below.

With regard to the co-chaperones, Cpn10 from *A. thaliana* (AtCpn10) could be purified as a heptameric complex similar to *E. coli* GroES (Fig. 16 C and F). The oligomeric structure of Cpn20 is still a matter of discussion (Baneyx *et al.* 1995, Hirohashi *et al.* 1999, Koumoto *et al.* 1999, Sharkia *et al.* 2003), but for purified recombinant AtCpn20 it could be determined by light scattering to be ~80 kDa, which corresponds to a tetrameric composition (Fig. 16 F). This implies interaction of eight GroES-like domains of AtCpn20 with the seven subunits of a Cpn60 ring. The mechanism of such an interaction remains to be determined. The N- and C-terminal domains of AtCpn20 were also purified, but they were susceptible to degradation (Fig. 16 C), as described previously for the spinach Cpn20-domains (Baneyx *et al.* 1995).

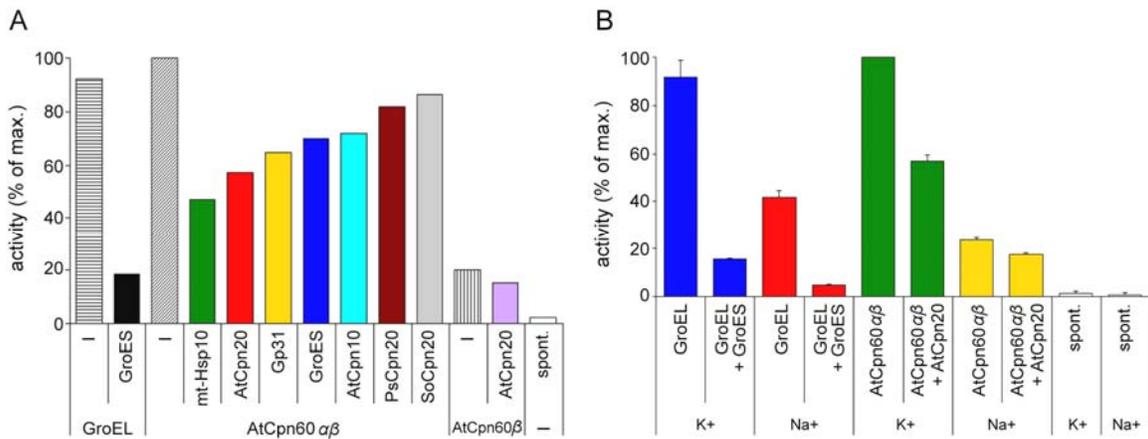
In order to provide further co-chaperones with different phylogenetic origin and functional peculiarities, mitochondrial Hsp10 from yeast (mt-Hsp10) and Gp31 from bacteriophage T4 were also utilized for refolding attempts. Gp31 lacks a roof structure and thus forms a larger folding cage with GroEL, allowing for the encapsulation of larger substrate proteins (Hunt *et al.* 1997). Notably, similar structural and functional features have been suggested for chloroplast Cpn10 (Koumoto *et al.* 2001).

#### **4.1.3 Functional characterization of GroEL, AtCpn60 $\alpha\beta$ , AtCpn60 $\beta$ and their interaction with various co-chaperones**

In order to confirm and compare the activity of the purified chloroplast chaperonins (AtCpn60 $\alpha\beta$  and AtCpn60 $\beta$ ) and *E. coli* GroEL as well as to assess their productive interaction with different co-chaperones, their ATPase and refolding activity were examined *in vitro*. The ATPase activity of Cpn60 is measured in the absence of substrate proteins and is inhibited upon interaction with co-chaperone (Baneyx *et al.* 1995, Martin *et al.* 1991, Viitanen *et al.* 1990).

The recombinant AtCpn60 $\alpha\beta$  showed a similar ATPase activity as GroEL. ATP hydrolysis by GroEL was diminished by 73 % in the presence of GroES. For AtCpn60 $\alpha\beta$ , this inhibition was less pronounced in concert with various co-chaperones and ranged from 53 % to only 14 % in decreasing order of mt-Hsp10, AtCpn20, Gp31, GroES, AtCpn10, PsCpn20 and SoCpn20. The ATPase activity of AtCpn60 $\beta$  amounted to merely 20 % of AtCpn60 $\alpha\beta$  or GroEL and hardly any inhibition could be measured in presence of AtCpn20 (Fig. 17 A). For both GroEL and AtCpn60 $\alpha\beta$ , ATP hydrolysis was higher in presence of ionic potassium than in buffers containing sodium instead (Fig. 17 B).

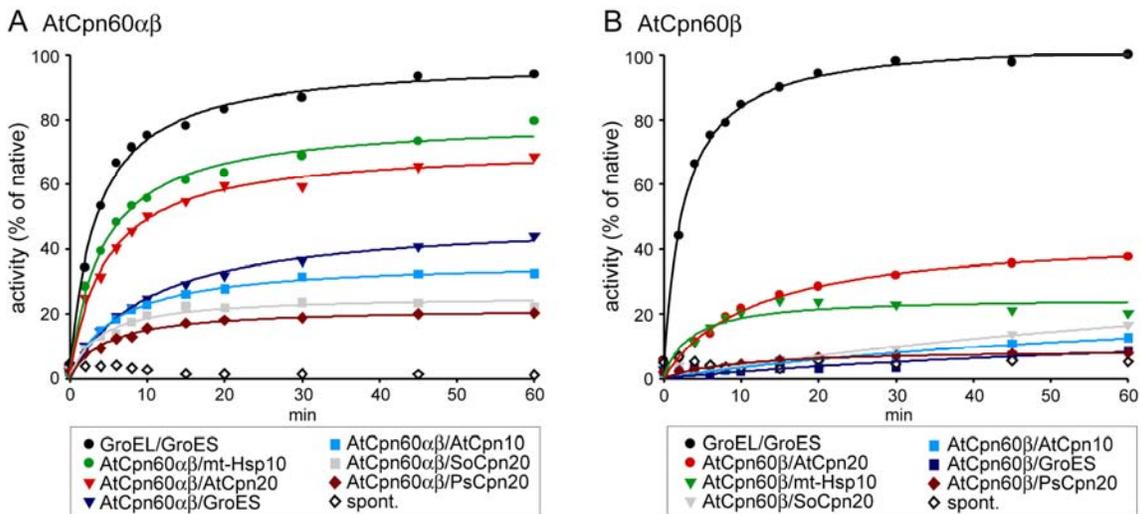
Refolding of malate dehydrogenase (MDH) by both AtCpn60 $\alpha\beta$  and AtCpn60 $\beta$  was most efficient in concert with mt-Hsp10 and AtCpn20, whereas GroES, AtCpn10, SoCpn20 or PsCpn20 accounted for lower refolding yields. However, MDH refolded with considerably higher efficiency with AtCpn60 $\alpha\beta$  than with AtCpn60 $\beta$  (Fig. 18 A and B).



**Figure 17. ATPase activity of GroEL, AtCpn60 $\alpha\beta$  and AtCpn60 $\beta$  and inhibition by co-chaperones.**

(A) Chaperonins (0.5  $\mu$ M) and, if indicated, co-chaperones (1  $\mu$ M) were added to assay-buffer 1 (20 mM MOPS-KOH, pH 7.5, 100 mM KCl, 5 mM MgCl<sub>2</sub>). Reactions were incubated for 5 min at 25 °C, before they were initiated with 2 mM ATP. At indicated time points aliquots of 10  $\mu$ l were withdrawn and stopped with 20 mM CDTA. Phosphate was determined colorimetrically with malachite green reagent. ATPase activity of AtCpn60 $\alpha\beta$  was set to 100 %.

(B) Influence of K<sup>+</sup> and Na<sup>+</sup> on ATPase activity. ATPase activity was determined as in (A) using (K<sup>+</sup>) assay buffer 1 or (Na<sup>+</sup>) buffer (20 mM MOPS-NaOH, pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>).



**Figure 18. Refolding of MDH by AtCpn60 $\alpha\beta$  or AtCpn60 $\beta$  and different co-chaperones.**

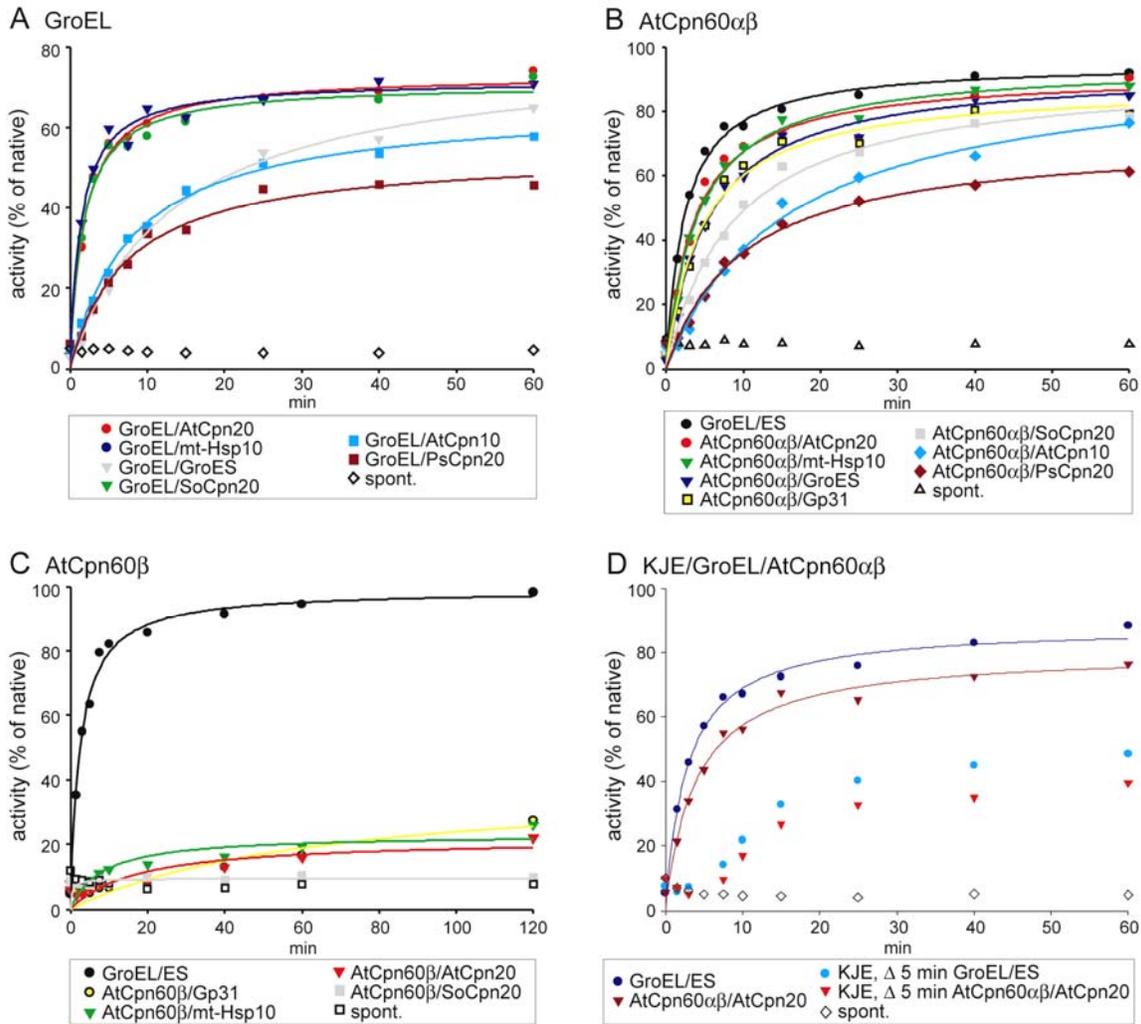
Denatured MDH was diluted 50-fold (0.5  $\mu$ M) at 37 °C into assay-buffer 2 containing (A) AtCpn60 $\alpha\beta$  or (B) AtCpn60 $\beta$  (2  $\mu$ M) and co-chaperones (2  $\mu$ M) as indicated. Refolding was initiated by addition of 2 mM ATP. At indicated time points, samples were taken and immediately stopped with CDTA (65 mM) and the activity of refolded MDH was determined. The enzymatic activity of an equivalent amount of native MDH is set to 100 %.

For Rr-RbcL as substrate, the refolding ability of GroEL and AtCpn60 $\alpha\beta$  were similar, showing highest efficiency in cooperation with AtCpn20, mt-Hsp10, GroES or Gp31 and lower efficiency with SoCpn20, AtCpn10 or PsCpn20 (Fig. 19 A and B). Thus, GroEL

and AtCpn60 $\alpha\beta$  are functionally compatible with co-chaperones from evolutionary diverse sources (bacterial, viral, mitochondrial and chloroplast Cpn10 or Cpn20) (Dickson *et al.* 2000, Rospert *et al.* 1993, Viitanen *et al.* 1995). The slightly reduced effectiveness of their interaction with SoCpn20, AtCpn10 and PsCpn20 is also reflected by the less pronounced inhibition of ATPase activity (Fig. 17 A) or by MDH refolding (Fig. 18 A). In contrast to full-length AtCpn20, neither the N- nor the C-terminal domain alone showed chaperonin-assisted refolding, which is in agreement with previous reports (Baneyx *et al.* 1995, Bertsch and Soll 1995).

Unlike AtCpn60 $\alpha\beta$ , AtCpn60 $\beta$  displayed very poor Rr-RbcL refolding capability. The latter resulted in only 20-30 % of the GroEL/GroES assisted refolding yield and highest efficiency was observed for the interaction of AtCpn60 $\beta$  with Gp31, mt-Hsp10 or AtCpn20 (Fig.19 C). This is in agreement with both the low ATPase activity and MDH-refolding results with AtCpn60 $\beta$ . Dickson *et al.* (2000) have observed a considerably more efficient refolding of Rr-RbcL with Cpn60 $\beta$  from *Pisum sativum* and mouse mt-Cpn10, but no refolding with GroES or SoCpn20 as the co-chaperones. In general, the low activity of AtCpn60 $\beta$  is probably due to the intrinsic properties of this chaperonin. The existence homotetradecameric Cpn60 $\beta$ <sub>14</sub> oligomers in chloroplasts is still an open question (Nishio *et al.* 1999).

The chaperone system of DnaK/DnaJ/GrpE (KJE) was neither alone nor in presence of GroEL/GroES or AtCpn60 $\alpha\beta$ /AtCpn20 capable of supporting the refolding of Rr-RbcL (Fig. 19 D). Upon dilution of denatured Rr-RbcL into buffer containing DnaK/DnaJ/GrpE, refolding did not start before chaperonins and co-chaperones were also added. The refolding yields were observed to be lower in the presence of KJE. The latter effect might be attributable to the interaction between Rr-RbcL and KJE, resulting in a delayed transfer of RbcL to the chaperonin folding machinery, or to an increased aggregation of Rr-RbcL upon dilution from denaturant into buffer lacking GroEL.



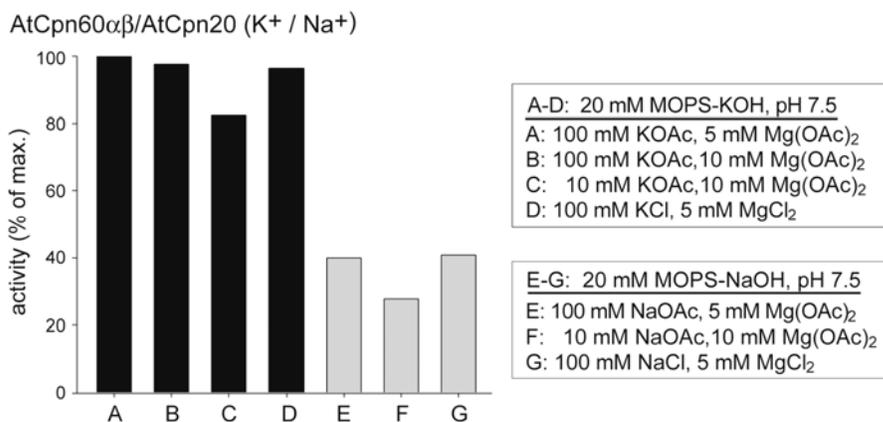
**Figure 19. *In vitro* refolding of Rr-Rbcl.**

(A-C) Denatured Rr-Rbcl was diluted 100-fold (50 nM Rbcl monomer) into ice-cold assay-buffer 1 containing 5 mM DTT, 1 mg/ml BSA and 200 nM of indicated chaperonin. Reactions were incubated for 5 min at 25 °C, before co-chaperone (650 nM) was added as indicated. Refolding was started by addition of 2 mM ATP at 25 °C. Samples were taken at indicated time points, stopped with glucose (10 mM) / hexokinase (2.5 U) and carboxylation activity was determined. The enzymatic activity of an equivalent amount of native Rr-Rbcl is set to 100 %. Note that for C (AtCpn60 $\beta$ ) the curves for GroES, AtCpn10 and PsCpn20 are not shown as they overlap with the control reaction (spont.).

(D) Denatured Rr-Rbcl was diluted into assay buffer 1 containing 5 mM DTT, 1 mg/ml BSA and DnaK/DnaJ/GrpE, followed by addition of GroEL/GroES or AtCpn60 $\alpha\beta$ /AtCpn20 after 5 min. Reactions were otherwise performed as described for A-C. For comparison, reactions lacking DnaK/DnaJ/GrpE were carried out. The stoichiometry of components was 1 Rr-Rbcl (50 nM Rbcl monomer) : 4 GroEL/AtCpn60 : 13 GroES/AtCpn20 : 5 DnaK : 2.5 DnaJ : 5 GrpE. The enzymatic activity of an equivalent amount of native Rr-Rbcl is set to 100 %.

As observed for the ATPase activity, refolding of Rr-RbcL by AtCpn60 $\alpha\beta$ /AtCpn20 was more efficient in the presence of potassium ions compared to sodium ions (Fig. 20). The requirement for K<sup>+</sup> for ATP hydrolysis and thus for refolding by Cpn60 has been shown for *E. coli* GroEL and mitochondrial mt-Cpn60 (Rospert *et al.* 1993, Viitanen *et al.* 1990, 1992b). However, for PsCpn60 similar refolding rates in presence of K<sup>+</sup> and Na<sup>+</sup> were reported, but the authors could not exclude the possibility of presence of bound monovalent cations (Viitanen *et al.* 1995).

In summary, GroEL and AtCpn60 $\alpha\beta$  showed typical chaperonin properties including similar ATPase activity and assisted refolding of the two model substrates MDH and Rr-RbcL. Of all tested co-chaperones, GroES, AtCpn20 and mt-Hsp10 were most effective in the functional interaction with GroEL or AtCpn60 $\alpha\beta$ . For AtCpn60 $\beta$  only comparatively low ATP hydrolysis and refolding of MDH or Rr-RbcL was measured. Consequently, both GroEL/GroES and AtCpn60 $\alpha\beta$ /AtCpn20 were predominantly utilized for further *in vitro* refolding analysis of type I RuBisCO.



**Figure 20. Effect of K<sup>+</sup> and Na<sup>+</sup> on Rr-RbcL refolding by AtCpn60 $\alpha\beta$ /AtCpn20.**

Refolding of Rr-RbcL was analyzed as described in Fig. 19 in the indicated buffers (containing 5 mM DTT, 1 mg/ml BSA) with AtCpn60 $\alpha\beta$ /AtCpn20 as chaperonin system. Indicated activities were measured after 10 min and the highest was set to 100 %.

#### 4.1.4 Type I RuBisCO cannot be refolded *in vitro*

Numerous experiments were performed to attempt to reconstitute Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> *in vitro*. A range of concentrations and combinations of different chaperonins (GroEL, AtCpn60αβ, AtCpn60β) and co-chaperones (GroES, mt-Hsp10, Gp31, AtCpn20, SoCpn20, PsCpn20, AtCpn10) or other chaperone systems (e.g. DnaK/DnaJ/GrpE) together with different species of RbcS and RbcX was tried. Moreover, refolding of RuBisCO was attempted using varying concentrations denaturants, including guanidinium, urea or acid. The effect of different temperatures, buffers (e.g. Tris, MOPS, ethanolamine, RTS-buffer), pH-values (adjusted e.g. with NaOH, KOH, HCl, Acetic acid), salts (e.g. NaCl, KCl, NaOAc, KOAc, MgCl<sub>2</sub>, Mg(OAc)<sub>2</sub>, ammoniumsulfate), the presence of soluble *E. coli* lysate, BSA, casein, chaotropic agent arginine, RuBisCO substrate RuBP, crowding reagents (e.g. dextran, ficoll, PEGs, TMAO), redox reagents or redox pairs (DTT, β-mercaptoethanol, GSH/GSSG) in different concentrations and combinations was examined as well. The folding of Rr-RbcL was used as the standard control. Unfortunately, none of the conditions yielded active Syn6301-RuBisCO enzyme.

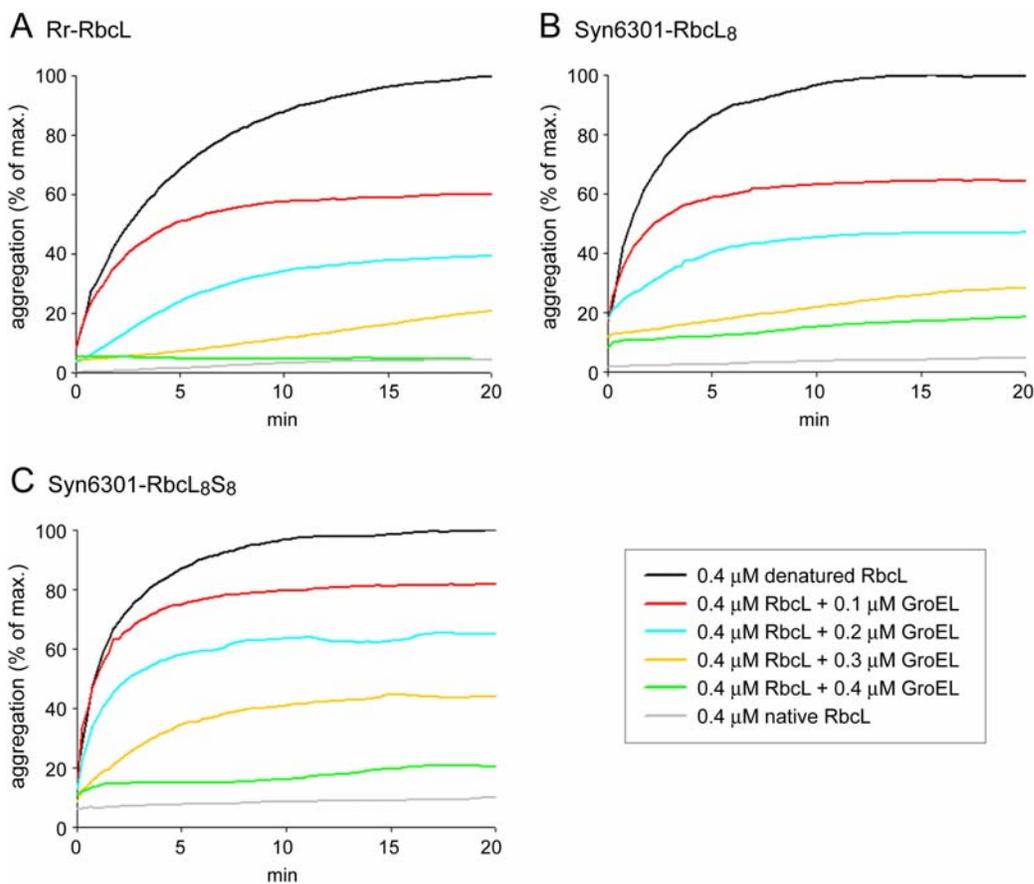
#### 4.1.5 Analysis on the impediment in reconstitution of type I RuBisCO *in vitro*

Since Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> could not be refolded *in vitro*, different aspects of the folding process were analyzed in more detail in order to identify the limiting factor(s). GroEL/GroES was used as chaperonin system for these analyses, since it efficiently supported the production of the active enzyme upon expression in *E. coli* (Goloubinoff *et al.* 1989 a, Fig. 27).

##### 4.1.5.1 GroEL binds unfolded Syn6301-RbcL, thereby preventing aggregation

The chaperonin-assisted folding process is initiated by the binding of substrate protein to the chaperonin. Binding of unfolded proteins can be analyzed under non-permissive conditions, in which these proteins fail to refold spontaneously but instead aggregate. This aggregation is prevented upon binding to chaperonins. To confirm that GroEL can bind unfolded Rr-RbcL and Syn6301-RbcL, aggregation prevention was analyzed

photometrically for denatured Rr-RbcL<sub>2</sub>, Syn6301-RbcL<sub>8</sub> and Syn6301-RbcL<sub>8</sub>S<sub>8</sub> under non-permissive conditions. Dilution of these proteins from denaturant into buffer alone resulted in aggregation, whereas GroEL prevented this process for all three substrates by complex formation in a concentration dependent manner showing fully efficient prevention of aggregation at a 1:1 ratio (Fig. 21). Therefore, it can be concluded that unfolded Rr-RbcL and Syn6301-RbcL are efficiently captured by GroEL.



**Figure 21. Aggregation prevention by GroEL.**

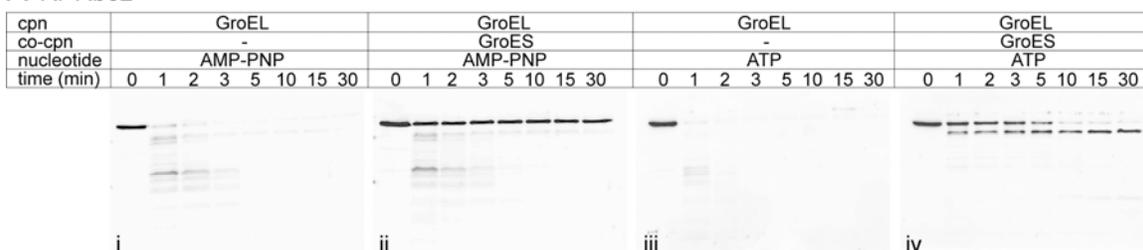
Denatured Rr-RbcL<sub>2</sub> (A), Syn6301-RbcL<sub>8</sub> (B) or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> (C) were diluted 100-fold (0.4  $\mu\text{M}$  RbcL monomers) into ice-cold assay-buffer 1 containing the indicated amounts of GroEL. Aggregation was measured at 320 nm and 25 °C. Dilution of an equivalent amount of native protein served as a control to demonstrate that the readily folded and assembled proteins do not aggregate or cause an increase in turbidity. Aggregation in the absence of GroEL is set to 100 %.

#### 4.1.5.2 Syn6301-RbcL is encapsulated in the GroEL cage with properly cycling GroES

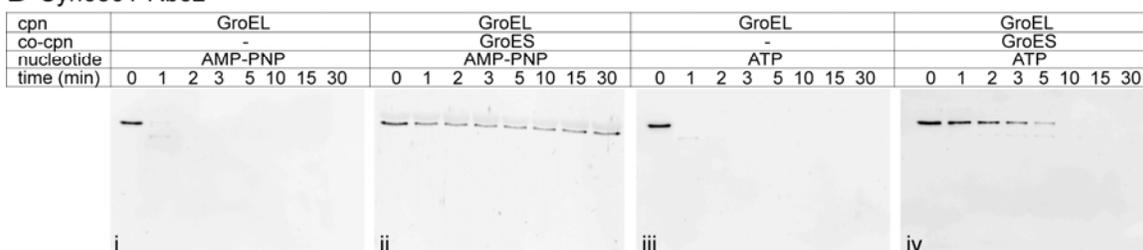
Furthermore it was examined, if Syn6301-RbcL is properly encapsulated in the cavity formed by GroEL/GroES. In general, the chaperonin cavity has capacity for proteins up to 60 kDa (Hartl and Hayer-Hartl 2002). Since Syn6301-RbcL has a size of approximately 52 kDa, it should become completely encapsulated and free to fold in the GroEL/GroES cage. In order to estimate the efficiency of encapsulation, Proteinase K degradation of GroEL-bound RbcL was analyzed in absence or presence of GroES and ATP or AMP-PNP (non-hydrolysable analogue of ATP).

Both Rr-RbcL and Syn6301-RbcL are completely digested by Proteinase K when bound to GroEL in absence of GroES (Fig. 22 A and B, i and iii). Addition of GroES together with AMP-PNP to the GroEL-RbcL-complex resulted in protection of the substrate from Proteinase K digestion (Fig. 22 A and B, ii). However, in the presence of ATP, the RbcL was slowly digested (Fig. 22 A and B, iv) due to cycling of GroES. In summary, both Rr-RbcL and Syn6301-RbcL are efficiently encapsulated in the GroEL/GroES cage.

##### A Rr-RbcL



##### B Syn6301-RbcL



#### Figure 22. Encapsulation of RbcL and cycling of GroES (Proteinase K protection).

Denatured Rr-RbcL(A) or Syn6301-RbcL<sub>8</sub>(B) were diluted 100-fold (300 nM RbcL monomer) into ice-cold assay-buffer 1 containing 300 nM GroEL. Reactions were immediately incubated for 5 min at 25 °C, before aggregated material was removed by centrifugation. Where indicated, reactions were then supplemented with 1 μM GroES and incubated for further 5 min, followed by addition of ATP or AMP-PNP (4 mM). Digestion was initiated by addition of Proteinase K (2.5 μg/ml) at 25 °C. At indicated times samples were taken and digestion was stopped with 1 mM PMSF. Samples were subjected to SDS-PAGE, followed by immunoblotting against RbcL.

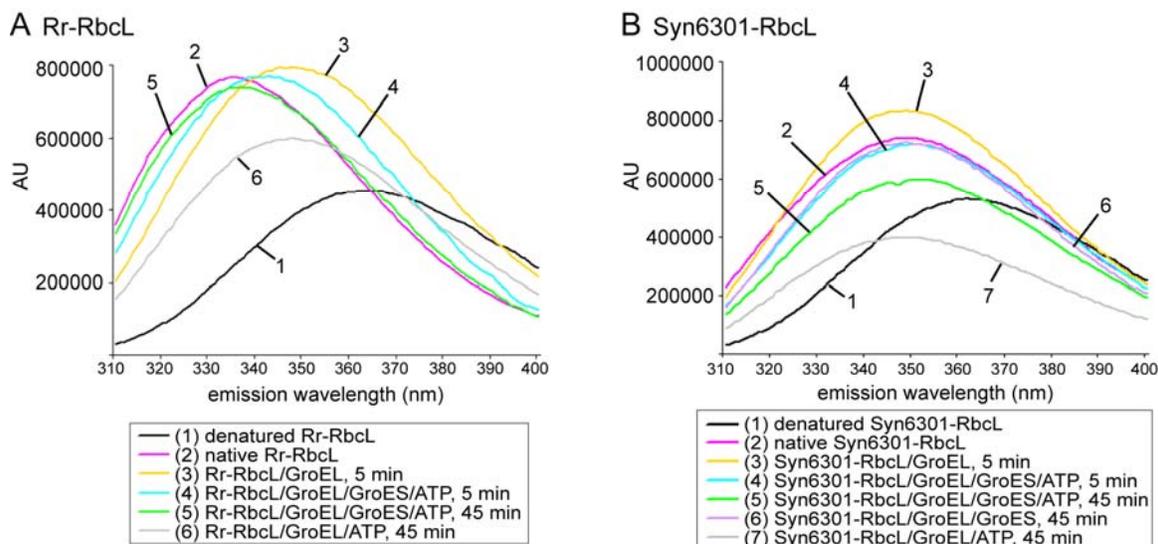
#### 4.1.5.3 Conformational changes of Syn6301-RbcL upon interaction with the chaperonin system

Considering that Syn6301-RbcL can be bound and encapsulated by GroEL, the question arises, if the substrate protein undergoes any structural conversions upon interaction with the chaperonin system. To address this issue, intrinsic tryptophan fluorescence was measured during the refolding reaction of Rr-RbcL<sub>2</sub> and Syn6301-RbcL<sub>8</sub>. In Rr-RbcL and Syn6301-RbcL the numbers of tryptophan residues are six and nine, respectively, while GroEL and GroES lack tryptophans. Changes in tryptophan fluorescence can indicate 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 situation, which is mostly brought about by a conformational change typical for protein folding (Schmid 1990). Measurement of tryptophan fluorescence was complemented with analysis of ANS(1-anilino-8-naphthalene-sulphonate)-binding. ANS can be used as a probe for exposed apolar binding sites and is thus used as indicator for alterations in protein tertiary or quaternary structure (Martin *et al.* 1991, Wildner 1976).

For Rr-RbcL tryptophan fluorescence changed markedly upon denaturing in 6 M guanidinium: the emission maximum shifted from 335 nm of the native protein (Fig. 23 A, curve 2) to 365 nm for the unfolded enzyme (Fig. 23 A, curve 1), accompanied by a considerable decrease of fluorescence intensity. The emission maximum of GroEL-bound Rr-RbcL was at 348 nm (Fig. 23 A, curve 3), i.e. it shifted ~ 57 % from the denatured to the folded state. The fluorescence intensity of GroEL-bound Rr-RbcL was similar to the native enzyme, but decreased upon ATP-hydrolysis in absence of GroES, probably due to loss of some protein as a consequence of aggregation (Fig. 23 A, curve 6). Apparently, unfolded Rr-RbcL is stabilized by GroEL in a conformation lacking ordered tertiary structure (molten globule state). This was confirmed by high ANS-fluorescence (Fig. 24, column 1), which reflects massive binding of ANS to the hydrophobic surfaces of early folding intermediates. Upon folding of Rr-RbcL in presence of GroEL/GroES and Mg-ATP, the tryptophan fluorescence emission maximum was shifted back to 335 nm (Fig. 23 A, curve 5), identical with the native protein, indicating completed reconstitution of the dimeric Rr-RbcL<sub>2</sub> enzyme. 50 % of this blueshift was reached already after only 5

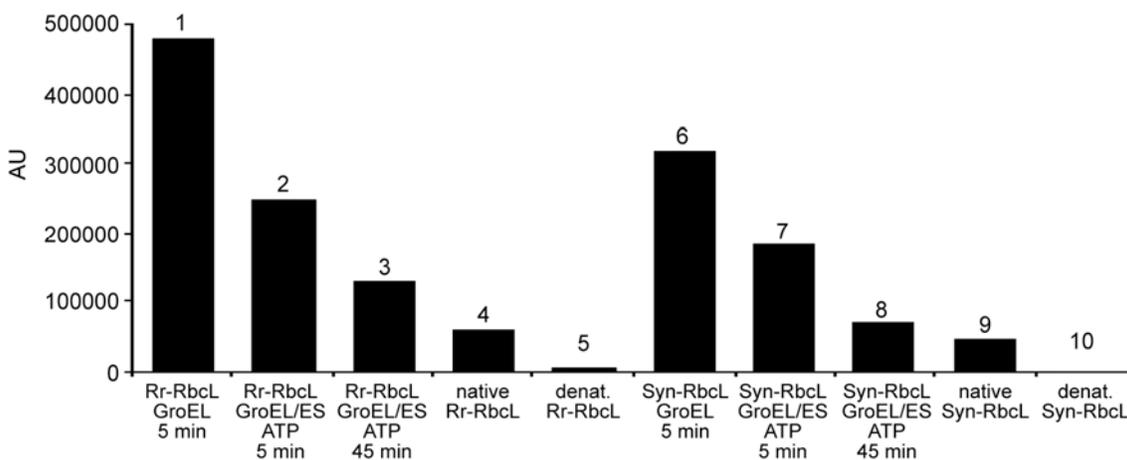
min (Fig. 23 A, curve 4), indicating high efficiency of refolding and assembly. Concomitantly, the intensity of ANS-fluorescence decreased, confirming changes in the tertiary or quaternary structure of the enzyme molecule during the refolding reaction (Fig. 24, columns 2-3).

Unfolding of Syn6301-RbcL<sub>8</sub> resulted in a redshift of the fluorescence maximum from 350 nm (native state, Fig. 23 B, curve 2) to 361 nm (unfolded state, Fig. 23 B, curve 1), accompanied by a decrease of emission intensity of more than 25 %. Binding of unfolded Syn6301-RbcL to GroEL shifted the fluorescence maximum back to 350 nm, identical to the native protein, but fluorescence intensity was higher (Fig. 23 B, curve 3). Both increased tryptophan fluorescence intensity and high ANS-fluorescence (Fig. 24, column 6) of GroEL-bound Syn6301-RbcL might reflect its unordered structural state. Upon incubation of the GroEL/Syn6301-RbcL-complex with only GroES or both GroES and Mg-ATP, the tryptophan fluorescence maximum remained at 350 nm and the intensity decreased to a level similar to the native enzyme (Fig. 23 B, curves 4 and 6). But over time, the emission intensity in the refolding reaction went below that of the native state (Fig. 23 B, curve 5), which probably reflects the monomeric folded state of RbcL. The same effect could be observed for both the GroEL/Syn6301-RbcL-complex (Fig. 23 B, curve 7) and the GroEL/Rr-RbcL-complex (Fig. 23 A, curve 6) upon incubation with only Mg-ATP in absence of GroES. Accordingly, this scenario could argue against progressive folding of Syn6301-RbcL. However, decrease of ANS-fluorescence throughout the refolding reaction indicated conformational changes typical for folding (Fig. 24, columns 7-8).



**Figure 23. Tryptophan fluorescence in RbcL refolding.**

Denatured Rr-RbcL (A) or Syn6301-RbcL<sub>8</sub> (B) were diluted 100-fold (0.25  $\mu$ M RbcL monomer) into ice-cold assay-buffer 1 containing 0.625  $\mu$ M GroEL. After incubation for 5 min at 25  $^{\circ}$ C, 1.25  $\mu$ M GroES was added. After 2 min, refolding was started by supplementing the reactions with 2 mM ATP and incubation at 25  $^{\circ}$ C. If only native or denatured substrate or merely substrate-binding to GroEL had to be analyzed, reactions were modified accordingly. Samples were withdrawn at indicated times and reactions were stopped with 10 mM CDTA. Immediately, tryptophan fluorescence was measured (excitation 295 nm, emission scan 310-400 nm) and background fluorescence of chemically identical reactions lacking RbcL was subtracted.

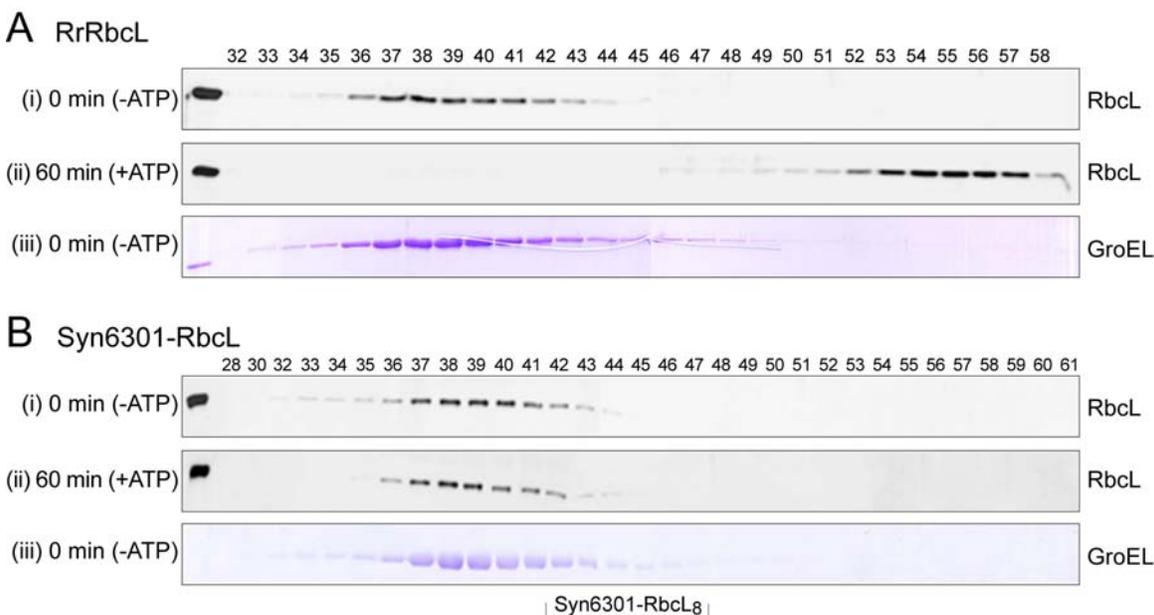


**Figure 24. ANS-fluorescence in RbcL refolding.**

As indicated, native or denatured Rr-RbcL<sub>2</sub> or Syn6301-RbcL<sub>8</sub> were diluted 100-fold (0.25  $\mu$ M RbcL monomer) into ice-cold assay-buffer 1 containing 5  $\mu$ M ANS and for reactions 1-3 and 6-8 also 0.625  $\mu$ M GroEL. After incubation for 5 min at 25  $^{\circ}$ C, fluorescence was measured directly for reactions 1, 4-6 and 9-10, whereas reactions 2-3 and 7-8 were supplemented with 1.25  $\mu$ M GroES and (after further 2 min) 2 mM ATP to start refolding. Refolding reactions 2-3 and 7-8 were incubated at 25  $^{\circ}$ C, samples were withdrawn and stopped with 10 mM CDTA after 5 min or 45 min, respectively. A fluorescence scan was immediately performed for all reactions (excitation 390 nm, emission scan 420-550 nm), data were corrected for background fluorescence of chemically identical reactions lacking RbcL and emission at 470 nm was depicted.

#### 4.1.5.4 Syn6301-RbcL is not released from GroEL as assembly-competent protein

In the experiments described thus far, no carboxylation activity was measurable at the end of refolding reactions for Syn6301-RuBisCO. Carboxylation activity would indicate the formation of properly folded and assembled holoenzyme. To address if folded and soluble Syn6301-RbcL is released from GroEL during folding, the refolding reactions were analyzed by gel filtration.



**Figure 25. Gel filtration at the beginning and end of RuBisCO refolding reactions.**

(A) Denatured Rr-RbcL was diluted 100-fold (0.25  $\mu$ M RbcL monomers) into ice-cold assay-buffer 1 containing 5 mM DTT and 1  $\mu$ M GroEL. After incubation for 5 min at 25  $^{\circ}$ C, aggregates were removed by centrifugation, 3.25  $\mu$ M GroES was added and the reaction incubated for 5 min at 25  $^{\circ}$ C. Refolding was initiated by addition of 2 mM ATP at 25  $^{\circ}$ C. Samples of 200  $\mu$ l were withdrawn before (i, iii) and 60 min after (ii) addition of ATP, and immediately applied to a Superdex 200 (10/30) gel filtration column in assay-buffer 1. Fractions were analyzed by SDS-PAGE, followed by immunoblotting against RbcL (i,ii) or Coomassie staining to visualize GroEL (iii).

(B) Syn6301-RbcL<sub>8</sub> was subjected to a refolding reaction as in (A) with concentrations of 0.8  $\mu$ M RbcL monomers, 1.5  $\mu$ M GroEL and 4.5  $\mu$ M GroES under otherwise same conditions.

Samples which were withdrawn prior to initiation of folding by the addition of ATP showed elution of both RbcL and GroEL in the same fraction (Fig. 25 A and B, i and iii). This indicates the presence of GroEL-bound RbcL after dilution of unfolded Rr-RbcL or Syn6301-RbcL into GroEL-containing buffer. After 60 min refolding, Rr-RbcL eluted considerably later than GroEL, indicating that folding and assembly was accomplished and yielded the dimeric enzyme Rr-RbcL<sub>2</sub> (Fig. 25 A, ii and Fig. 19 A). On the contrary, Syn6301-RbcL still co-eluted with GroEL after 60 min of refolding (Fig. 25 B, ii). This

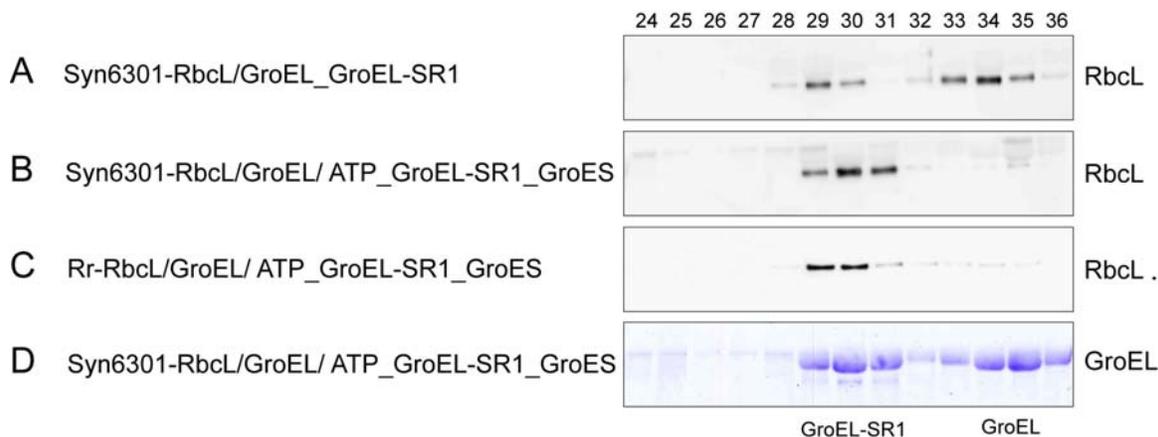
result shows that Syn6301-RbcL stayed bound to the chaperonin and thus failed to assemble. Furthermore, no carboxylation activity could be measured at the end of refolding reactions upon addition of RbcS, confirming the failure to obtain assembled Syn6301-RbcL<sub>8</sub> core complexes.

#### 4.1.5.5 Syn6301-RbcL cycles on GroEL

Considering the above-described results, one can assume that all requirements for folding of Syn6301-RbcL are being fulfilled inside the chaperonin cage. However, the obstacle in producing active Syn6301-RbcL<sub>8</sub>S<sub>8</sub> is most likely caused by the inefficient release of the folded substrate from the chaperonin into the environment in which assembly can proceed. Many proteins have to pass through several folding cycles until they reach their native state, i.e. incompletely folded polypeptide chains are released from Cpn60 and are rapidly recaptured for further folding attempts (Hartl and Hayer-Hartl 2002, Mayhew *et al.* 1996). In order to find out if Syn6301-RbcL subunits are released from GroEL, a refolding experiment was carried out to see if GroEL-bound RbcL was transferred to a non-cycling single-ring variant of GroEL (GroEL-SR1). GroEL-SR1 binds and encapsulates unfolded proteins, but under low salt conditions it does not release GroES, resulting in the accumulation of substrate protein in the closed folding cage (Weissman *et al.* 1995).

For this experimental approach, GroEL-bound RbcL was incubated with GroEL-SR1, either in absence or presence of ATP and GroES. The products of these reactions were subjected to field flow fractionation and the resulting fractions were analyzed for GroEL, GroEL-SR1 and RbcL by SDS-PAGE followed by gel staining or immunoblotting, respectively. GroEL and GroEL-SR1 are well separated upon field flow fractionation (Fig. 26 D). Thus the association of RbcL with either of them is clearly discernable. In the absence of ATP and GroES, the majority of Syn6301-RbcL stayed bound to GroEL and a smaller amount was transferred to GroEL-SR1 (Fig. 26 A). When ATP and GroES were added to the reaction, virtually all formerly GroEL-bound Syn6301-RbcL was transferred to and captured by GroEL-SR1 (Fig. 26 B). The same result was achieved when the experiment was performed with type II Rr-RbcL, which folds and assembles to the dimeric enzyme assisted by the chaperonin system (Fig. 26 C). Consequently, it can

be concluded that Syn6301-RbcL, just as Rr-RbcL, can be released and re-bound by GroEL and that improper cycling of Syn6301-RbcL on GroEL is not the reason for the failure of the *in vitro* refolding of type I RuBisCO.



**Figure 26. Cycling of RbcL on GroEL.**

Syn6301-RbcL<sub>8</sub> (A, B, D) or Rr-RbcL (C) was denatured in low salt assay-denaturation-buffer and diluted 100-fold (0.5 μM RbcL monomer) into ice-cold low salt assay-buffer containing 0.5 μM GroEL. After incubation for 10 min at 25 °C, aggregated material was removed by centrifugation. To the resulting solution of GroEL-bound RbcL either (A) GroEL-SR1 (2 μM) was added alone or (B-D) 2 mM ATP, 2 μM GroEL-SR1 and 3 μM GroES were added with a time lag of 15 sec and 60 sec, respectively. 50 μl of the each reaction were subjected to field flow fractionation. Collected fractions were TCA-precipitated and analyzed by SDS-PAGE with consecutive immunoblotting against RbcL (A-C) or Coomassie staining (D). Field flow fractionation was performed in cooperation with Dr. Manajit Hayer-Hartl.

## 4.2 Requirement of chaperonin for efficient production of assembled Syn6301-RuBisCO in *E. coli*

As described in the introduction, higher plant RuBisCO has never been produced as an active enzyme in *E. coli*, even in presence of co-expressed chloroplast chaperonins (Cloney *et al.* 1993, Gatenby 1984, Gatenby *et al.* 1987, Somerville *et al.* 1986). However, recombinant expression of type I RuBisCO from some cyanobacterial species has been shown to yield limited amounts of soluble, assembled and functional enzyme (Goloubinoff *et al.* 1989a, Gurevitz *et al.* 1985, Larimer and Soper 1993, Tabita and Small 1985).

The effect of different *E. coli* chaperones on *de novo* folding and assembly of RuBisCO from *Synechococcus* sp. PCC6301 (Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub>) was next investigated. Soluble cell lysates resulting from Syn6301-RuBisCO expression were

analyzed by Native PAGE. Without overproduction of the chaperonin- or Hsp70-system, the majority of newly synthesized RbcL formed insoluble aggregates and only minute amounts ( $\leq 1\%$ ) of assembled enzyme could be detected in the soluble fraction of *E. coli* cell lysate (Fig. 27 A and B, lanes 1). However, transient overproduction of the chaperonin system GroEL/GroES prior to induction of RbcL/RbcS increased the synthesis of assembled Syn6301-RuBisCO complexes to a considerable extent (Fig. 27 A and B, lanes 2). This effect has also been previously shown by Goloubinoff *et al.* (1989a) and is most likely caused by GroEL-mediated binding and folding of Syn6301-RbcL subunits, resulting in elevated levels of soluble and assembly-competent Syn6301-RbcL (Fig. 31, lane 3).

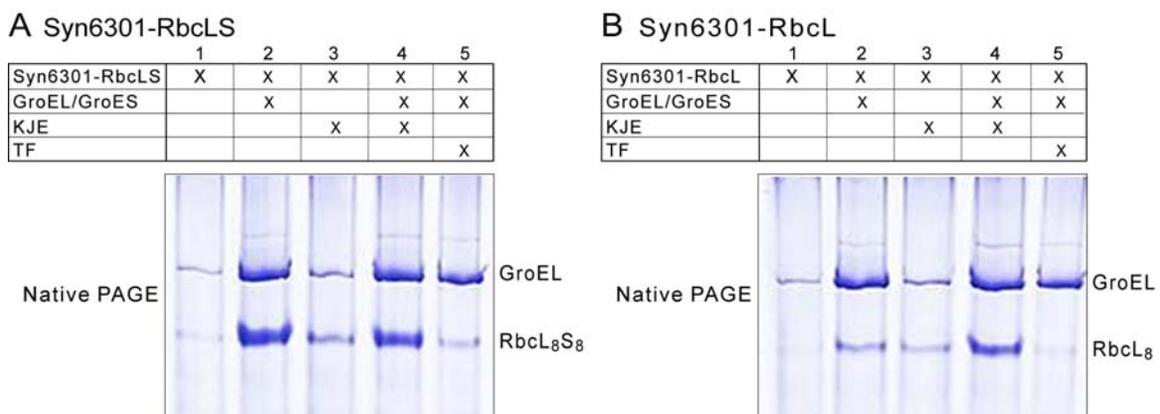
However, Syn6301-RbcL might also interact with other chaperones prior to or after folding by GroEL. Increased levels of trigger factor (TF), which interacts with translating polypeptides emerging from the ribosome exit tunnel, quenched the positive effect of GroEL/GroES almost completely (Fig. 27 A and B, lanes 5). This indicates a possible co-translational interaction between TF and Syn6301-RbcL, presumably leading to both delayed GroEL-mediated folding and RbcL subunit stabilization (Agashe *et al.* 2004, Nishihara *et al.* 2000). The observed negative effect of overproduced TF was absent in presence of only endogenous wildtype TF levels (Fig. 27 A and B, lanes 2)

The overproduction of the Hsp70 system, DnaK/DnaJ/GrpE, resulted in increased amounts of assembled RbcL<sub>8</sub> and RbcL<sub>8</sub>S<sub>8</sub> (Fig. 27 A and B, lanes 3) compared to the wildtype situation (Fig. 27 A and B, lanes 1). Hence, the effect is similar to GroEL, albeit less prominent. Likewise, the *in vivo* requirement of the DnaK/DnaJ/GrpE system for productive folding of type I RuBisCO was reported by Checa and Viale (1997). They observed extensive aggregation of RuBisCO in *dnaK* null mutants and enzyme activity was restored upon co-expression of *dnaK*.

Like TF, the Hsp70 system can bind to, shield and stabilize *de novo* synthesized nascent polypeptides of RbcL, thus preventing them from aggregation and keeping them in a folding competent state (Hartl and Hayer-Hartl 2002). In this respect, co- and post-translational interaction with TF and/or DnaK/DnaJ/GrpE appears to be important to prevent Syn6301-RbcL from aggregation and to keep it stabilized in a folding competent state for subsequent interaction with the GroEL/GroES system. Therefore, it is not

surprising that highest levels of assembled RuBisCO were found in the presence of wildtype TF levels combined with elevated amounts of both DnaK/DnaJ/GrpE and GroEL/GroES (Fig. 27 A and B, lanes 4).

Similar results were obtained from expression of Syn6301-RbcLS (Fig. 27 A) and Syn6301-RbcL (Fig. 27 B), confirming that formation of RbcL<sub>8</sub> cores is the limiting step in the RuBisCO folding and assembly pathway (see chapter 4.1) and that RbcL<sub>8</sub> production occurs in an RbcS-independent manner prior to the final assembly of RbcL<sub>8</sub>S<sub>8</sub> holoenzyme (Andrews 1988, Gatenby and Ellis 1990, Lee and Tabita 1990).



**Figure 27. Influence of chaperones on assembly of Syn6301-RbcL<sub>8</sub>S<sub>8</sub> or Syn6301-RbcL<sub>8</sub>.** Syn6301-RbcLS (A) or Syn6301-RbcL (B) was expressed in *E. coli* with or without prior transient overproduction of GroEL/GroES, DnaK/DnaJ/GrpE and TF as indicated. Soluble cell lysates were subjected to 6 % Native PAGE analysis followed by Coomassie blue staining.

### 4.3 Functional and structural characterization of RbcX

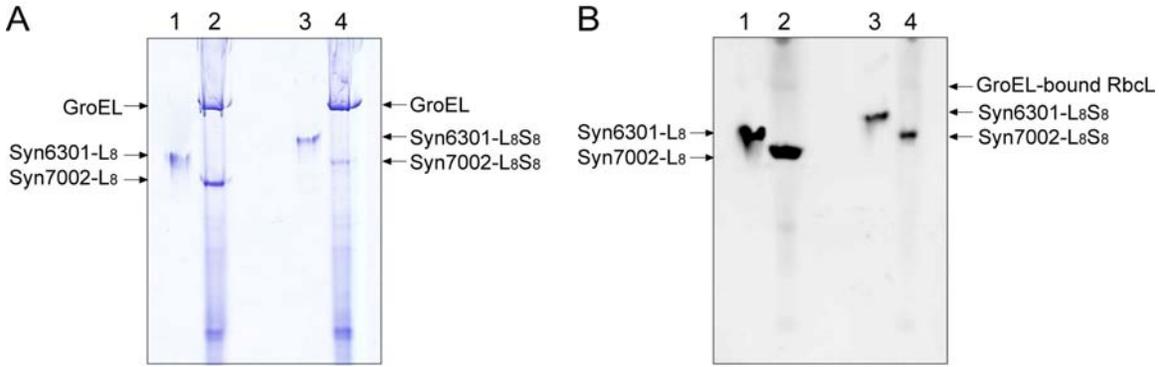
In addition to the chaperonin and the Hsp70-systems, a factor termed RbcX was reported to be either supportive or essential for the *in vivo* production of active RuBisCO from several cyanobacterial species (see introduction; Emlyn-Jones *et al.* 2006, Li and Tabita 1997, Onizuka *et al.* 2004). Since RbcX had been characterized only poorly and since analysis of its functional mechanism could be beneficial for a better understanding of the assembly of type I RuBisCO, both the function and the structure of RbcX were examined in detail for this study.

### 4.3.1 Sequential operation of GroEL/GroES and RbcX in folding and assembly of Syn7002-RuBisCO

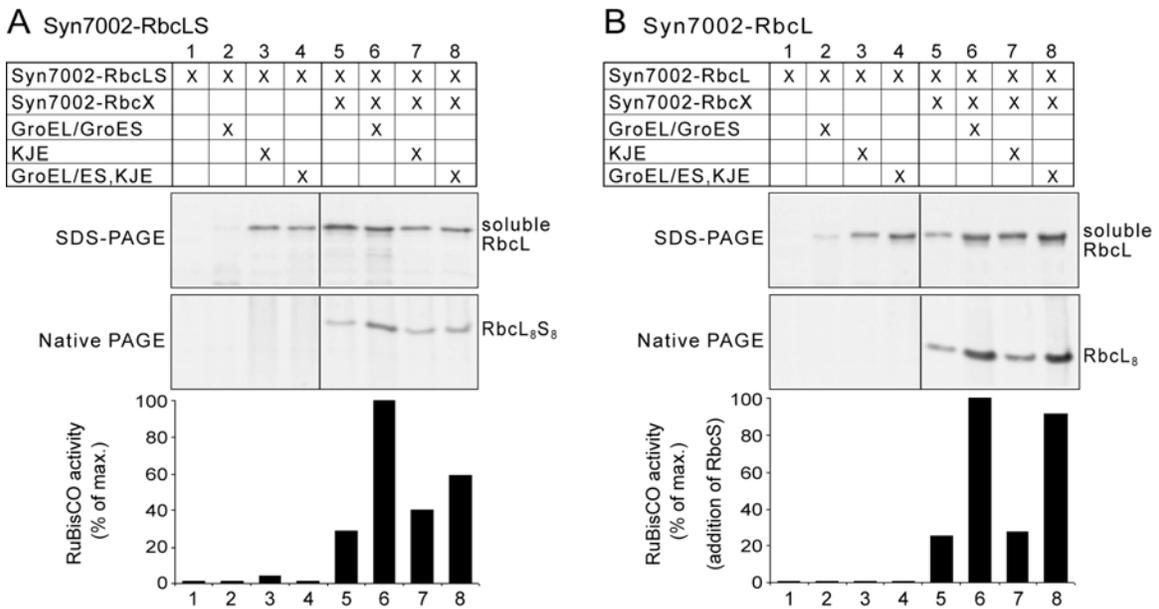
The requirement and effect of RbcX and molecular chaperones was analyzed for the folding and assembly of type I RuBisCO from the cyanobacterium *Synechococcus* sp. PCC7002 (Syn7002), which encodes *rbcX* in the same operon as the large and small RuBisCO subunit genes (*rbcLXS* operon; Onizuka *et al.* 2002, 2004).

For this purpose, Syn7002-RuBisCO was expressed alone or in concert with Syn7002-RbcX either with or without prior transient overproduction of GroEL/GroES and/or DnaK/DnaJ/GrpE in *E. coli* at 30 °C. In order to assess folding and assembly of RuBisCO, soluble cell lysate protein was subjected to SDS- and Native PAGE analysis with subsequent immunoblotting. Both Native PAGE and RuBisCO activity assay revealed if formation of assembled and enzymatically functional RbcL<sub>8</sub> or RbcL<sub>8</sub>S<sub>8</sub> complexes occurred. Both Syn7002-RbcL<sub>8</sub> and Syn7002-RbcL<sub>8</sub>S<sub>8</sub> migrated as high molecular weight complexes by Native PAGE (Fig. 28 A and B).

Expression of merely Syn7002-RbcLS or Syn7002-RbcL resulted predominantly in insoluble protein, and virtually no soluble RbcL or carboxylation activity could be detected (Fig. 29 A and B, lanes 1). Transient overproduction of GroEL/GroES, DnaK/DnaJ/GrpE or of both systems yielded increased amounts of soluble RbcL, but hardly any complex formation or RuBisCO activity was observed (Fig. 29 A and B, lanes 2-4). In contrast, co-expression of Syn7002-RbcLS or Syn7002-RbcL with Syn7002-RbcX resulted not only in elevated levels of soluble Syn7002-RbcL, but also in the assembly of RbcL<sub>8</sub>S<sub>8</sub> holoenzymes or RbcL<sub>8</sub> cores and concomitant occurrence of carboxylation activity (Fig. 29 A and B, lanes 5). Additional overproduction of GroEL/GroES (but not overproduction of DnaK/DnaJ/GrpE) could enhance this effect 3- to 4-fold (Fig. 29, lanes 6-8).



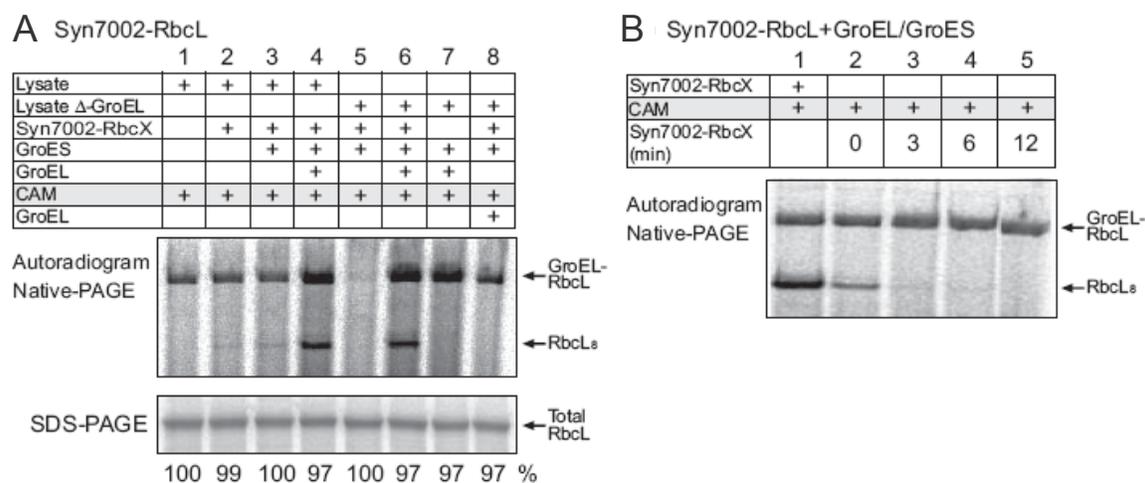
**Figure 28. Syn6301- and Syn7002-RbcL<sub>8</sub> and RbcL<sub>8</sub>S<sub>8</sub> in Native PAGE.**  
 Purified Syn6301-RbcL<sub>8</sub> (1), purified Syn6301-RbcL<sub>8</sub>S<sub>8</sub> (3) and soluble *E. coli* lysates, resulting from co-expression of Syn7002-RbcL/Syn7002-RbcX/GroEL/GroES (2) or co-expression of Syn7002-RbcL/Syn7002-RbcX/GroEL/GroES (4), were analyzed by 6 % Native PAGE, followed by Coomassie blue staining (A) or immunoblotting against RbcL (B).



**Figure 29. Influence of chaperones on solubility and assembly of Syn7002-RuBisCO.**  
 Syn7002-RbcLS (A) or Syn7002-RbcL (B) was expressed in *E. coli* with or without co-expression of Syn7002-RbcX. Where indicated, GroEL/GroES and/or DnaK/DnaJ/GrpE were transiently overproduced prior to expression of the *rbc*-genes. Soluble cell lysate was analyzed for soluble RbcL by 16 % SDS-PAGE and for complex assembly by 6 % Native PAGE, followed by immunoblotting against RbcL. Carboxylation activity in the soluble cell lysates was measured either directly (A) or upon addition of purified Syn6301-RbcS (B).

In summary, these results indicate that the solubility of Syn7002-RbcL benefits from elevated levels of DnaK/DnaJ/GrpE and/or GroEL/GroES, suggesting post-translational interaction of both chaperone systems with the large subunits. Moreover, the assembly of Syn7002-RbcL<sub>8</sub> is stringently dependent on RbcX, both in absence or presence of RbcS. This finding could be verified by *in vitro* experiments, in which Syn7002-RbcL was expressed in the RTS100 HY *E. coli in vitro* transcription/translation system. Newly synthesized Syn7002-RbcL was found bound to GroEL and only if the system was supplemented with GroEL/GroES and Syn7002-RbcX, substantial amounts of assembled RbcL<sub>8</sub> cores could be detected (Fig. 30 A, lanes 1-4). Translation of Syn7002-RbcL in a GroEL-immunodepleted lysate confirmed that RbcL<sub>8</sub> assembly only occurred upon supplementation with both GroEL/GroES and RbcX (Fig.30 A, lanes 5-7). When translation was performed in absence of GroEL, RbcX could not capture the newly synthesized RbcL for subsequent interaction with post-translationally added GroEL. Instead non-productive GroEL-RbcL complexes formed (Fig. 30 A, lane 8). Hence, proper folding and assembly of Syn7002-RbcL requires the presence of GroEL/GroES during or directly after translation of RbcL. At the same time RbcX should be present (Fig. 30 B, lanes 1 and 2), since delayed post-translational addition of RbcX resulted in progressively reduced amounts of RbcL<sub>8</sub> complexes accompanied by increased aggregation of RbcL (Fig. 30 B, lanes 3-5).

It can be concluded that interaction with DnaK/DnaJ/GrpE is necessary to keep newly-synthesized Syn7002-RbcL in a soluble state and to prevent it from aggregation, while the GroEL/GroES system is required for proper folding of Syn7002-RbcL. After release from the chaperonin system, RbcX is essential for the productive assembly of RbcL<sub>8</sub> complexes. The sequential operation mode of GroEL/GroES and RbcX seems to be optimized such that fast and efficient RbcX-mediated assembly might elevate the turnover-rate of GroEL/GroES by shifting the equilibrium from GroEL-bound RbcL to the assembly products, thereby increasing the efficiency of the whole system to supply maximal amounts of assembled, active enzyme.



**Figure 30. Effect of GroEL/ES and RbcX on production of assembled Syn7002-RuBisCO.**

(A) Syn7002-RbcL was translated in the RTS100 HY *E. coli in vitro* transcription/translation system in the presence of  $^{35}\text{S}$ -Met for 1.5 h at 30 °C. When indicated, GroEL/GroES (0.5  $\mu\text{M}$ /1.0  $\mu\text{M}$ ) and Syn7002-RbcX (40  $\mu\text{M}$ ) were added to normal lysate (lanes 1-4) or to GroEL-depleted lysate ( $\Delta$ -GroEL, lanes 5-8). In lane 8, GroEL was added after stopping translation with CAM. Assembled RbcL<sub>8</sub> and total RbcL protein were analyzed by 6 % Native-PAGE and 12.5 % SDS-PAGE, respectively, followed by autoradiography.

(B) Syn7002-RbcL was translated in the presence of GroEL/GroES as in (A). RbcX was either present during translation (lane 1), was added together with CAM (lane 2) or after CAM addition (lanes 3-5) and assembled RbcL<sub>8</sub> was analyzed by 6 % Native-PAGE followed by autoradiography. Experiments (A) and (B) were performed by Karnam Vasudeva Rao.

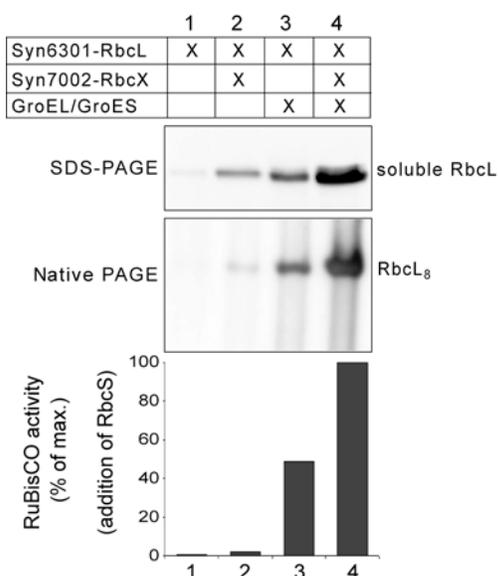
### 4.3.2 Different requirement for RbcX in the assembly of RuBisCO from *Synechococcus* sp. PCC6301 and *Synechococcus* sp. PCC7002

Interestingly, the requirement of RbcX for assembly of Syn7002-RuBisCO is reflected in the presence of the *rbcX* gene in the *rbcLXS* operon of *Synechococcus* sp. PCC7002 (Onizuka *et al.* 2002, 2004). Thus, coordinated expression of RbcX with the RbcL and RbcS subunits might be necessary to provide adequate amounts of RbcX for RuBisCO assembly in this cyanobacterial species.

In contrast, *rbcX* of *Synechococcus* sp. PCC6301 is encoded outside the *rbcLS* operon. Although expression of Syn6301-RbcL in *E. coli* resulted in a very low yield of assembled RbcL<sub>8</sub> complexes, coexpression of Syn7002-RbcX failed to significantly increase the amount of soluble and assembled RbcL (Fig. 31, lane 2). However, overexpression of GroEL/GroES resulted in the formation of considerable amounts of RbcL<sub>8</sub> cores which assembled with RbcS to give active holoenzymes (Fig. 31, lane 3). RbcX together with the chaperonin system was nevertheless beneficial for the solubility

of Syn6301-RbcL and enhanced the GroEL-mediated folding and assembly process more than twofold (Fig. 31, lane 4). This stresses again the probable function of RbcX as an assembly chaperone and defines the assembly of Syn6301-RuBisCO as an RbcX-supported but -independent process.

On the one hand it might be that the absence of *rbcX* in the *rbc* operon of *Synechococcus* sp. PCC6301 is an evolutionary consequence of RbcX-independent RuBisCO assembly. On the other hand it is plausible that RuBisCO assembly in any photosynthetic organism might benefit of RbcX irrespective of its absolute necessity for assembly and/or of the genomic localization of the *rbcX* gene.



**Figure 31. Effect of GroEL/GroES and Syn7002-RbcX on solubility and assembly of Syn6301-RuBisCO.**

Syn6301-RbcL was expressed in *E. coli* with or without co-expression of Syn7002-RbcX. Where indicated, transient overproduction of the GroEL/GroES-system preceded expression of the *rbc*-genes. Soluble cell lysate was analyzed for soluble RbcL by 16 % SDS-PAGE and for complex assembly by 6 % Native PAGE, followed by immunoblotting against RbcL. Carboxylation activity in the soluble cell lysate was measured upon addition of Syn6301-RbcS.

### 4.3.3 Further chaperone activities of RbcX

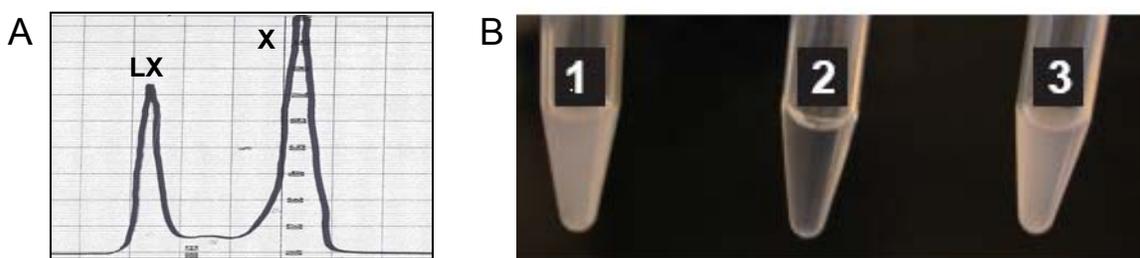
Based on our observations that RbcX probably functions as a RuBisCO assembly chaperone in the cellular environment of *E. coli* or in *E. coli*-derived translation systems, it was relevant to examine if it exhibits any other chaperone activity *in vitro*.

Unlike various other molecular chaperones, RbcX did not prevent aggregation of unfolded Rr-RbcL or Syn6301-RbcL under non-permissive conditions. It has neither ATPase activity of its own nor inhibits the ATPase activity of GroEL. In *in vitro* refolding of MDH or Rr-RbcL, RbcX could not replace GroEL or GroES and it did not alter the refolding rates of GroEL/GroES, even though different concentrations of RbcX were tested at temperatures ranging from 18 to 37 °C. The attempt to reconstitute *in vitro*

Syn6301-RbcL<sub>8</sub> or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> by GroEL/GroES-mediated folding in presence of various concentrations of RbcX also did not result in active enzyme.

However, when Syn7002-RbcL/RbcX-complexes were purified, it became evident that under certain conditions (e.g. pH <9) excess of RbcX is helpful for stabilization of the complexes and thus for the prevention of RbcL aggregation. This was observed when Syn7002-RbcL/Syn7002-RbcX<sub>N-His6</sub>-complexes were purified. For this purpose, the cell lysate from co-expression of RbcL and RbcX<sub>N-His6</sub> was subjected to affinity chromatography with Ni-NTA-agarose, resulting in a mixture of RbcL/RbcX<sub>N-His6</sub>-complexes and free RbcX<sub>N-His6</sub>-dimers. When subsequently the excess of free RbcX-dimers was removed by gel filtration at pH 7.5-8, a white precipitate was formed in the initially eluting fractions containing the RbcLX-complexes. Re-addition of RbcX could dissolve these aggregates whereas addition of an equal volume of buffer did not have this effect, indicating chaperone character of RbcX (Fig. 32 B). Notably, the stability of RbcL/RbcX-complexes could also be increased by performing the critical gel filtration at pH ≥ 9.

Based on these results, a detailed functional analysis of RbcX was not possible under mere *in vitro* conditions. Hence, further studies, focusing on the RbcX-mediated support of Syn7002-RuBisCO assembly, were performed in *E. coli* or in *E. coli* lysate.

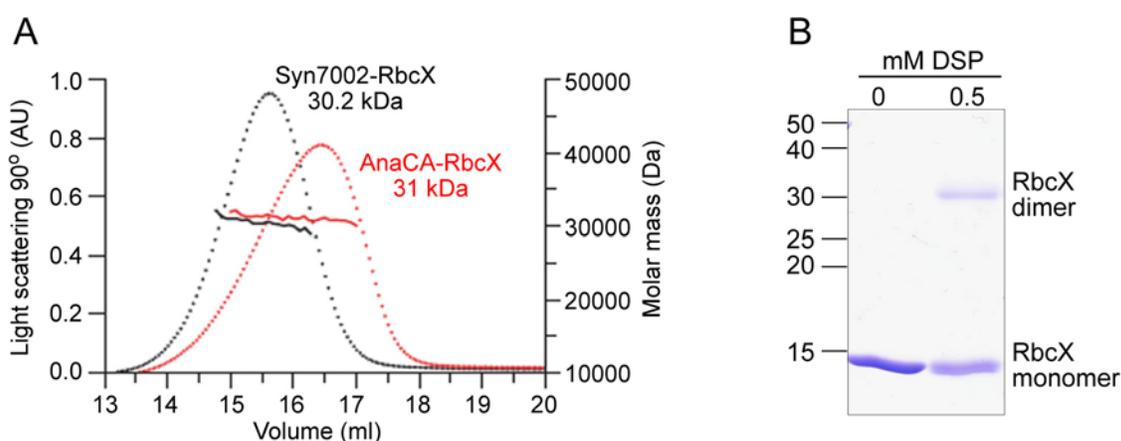


**Figure 32. Dissolution of RbcL aggregates by RbcX.**

(A) The final gel filtration step in the purification of Syn7002-RbcLX<sub>N-His6</sub>-complexes was performed in 20 mM Tris-HCl, pH 7.5 (instead of pH ≥ 9), resulting in elution of a first peak containing RbcLX<sub>N-His6</sub>-complexes and a second peak of free RbcX<sub>N-His6</sub>-dimers. At pH 7.5, a white precipitate appeared in the fractions of the first peak (B1), due to removal of excess RbcX. Equal amounts of this precipitate fraction were mixed with identical volumes of RbcX<sub>N-His6</sub> (B2), which was re-added from the second peak, or buffer (B3).

#### 4.3.4 Structural characterization of RbcX

Prior to further functional analyses of RbcX, the structure of this assembly chaperone needed to be examined in more detail. To this end, Syn7002-RbcX was expressed in *E. coli* and purified as soluble, homogeneous protein, running at an apparent molecular weight of ca. 15 kDa in SDS-PAGE. In order to determine its oligomeric state, the protein was subjected to field flow fractionation in conjunction with multiangle light scattering (FFF-MALS) and to crosslinking. Both methods revealed that Syn7002-RbcX forms dimers of ~30 kDa (Fig. 33 A and B).

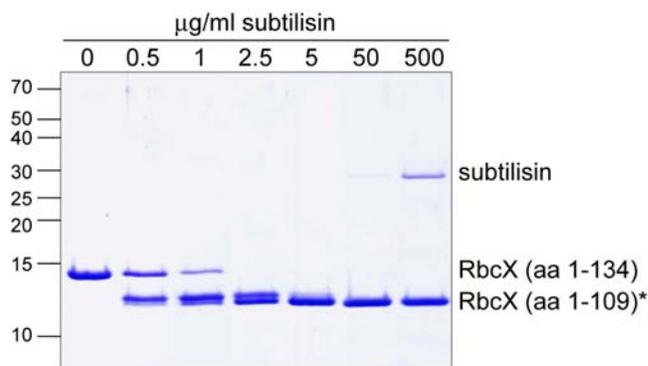


**Figure 33. Size determination of RbcX.**

(A) Molecular mass determination of Syn7002-RbcX (30170 Da) and AnaCA-RbcX (21040 Da) by FFF-MALS. Horizontal lines across the peaks indicate molar mass and homogeneity of the sample. FFF-MALS was performed in cooperation with Dr. Manajit Hayer-Hartl.

(B) Crosslinking and SDS-PAGE analysis. Syn7002-RbcX (0.5 mg/ml) was incubated in absence or presence of 1 mM DSP in 20 mM HEPES-NaOH, pH 7.5, 25 mM NaCl for 30 min at 25 °C. Crosslinking was stopped by addition of 20 mM Tris-HCl, pH 7.5. Samples were analyzed in absence of reducing agents by SDS-PAGE, followed by Coomassie blue staining.

With the intention to recognize potentially unstructured regions in Syn7002-RbcX, limited proteolysis by subtilisin was carried out. It resulted in the deletion of the C-terminus of RbcX comprising residues 110-134 (Fig. 34), as confirmed by Edman degradation and mass spectrometry. The C-terminal amino acids represent a sequence element that is only poorly conserved among ~150 cyanobacterial RbcX homologs (Fig. 36). In contrast, the protease resistant core domain, Syn7002-RbcX( $\Delta$ C25), is highly conserved, well structured and functional in supporting the formation of RbcL<sub>8</sub> complexes (Fig. 40, lane 30).

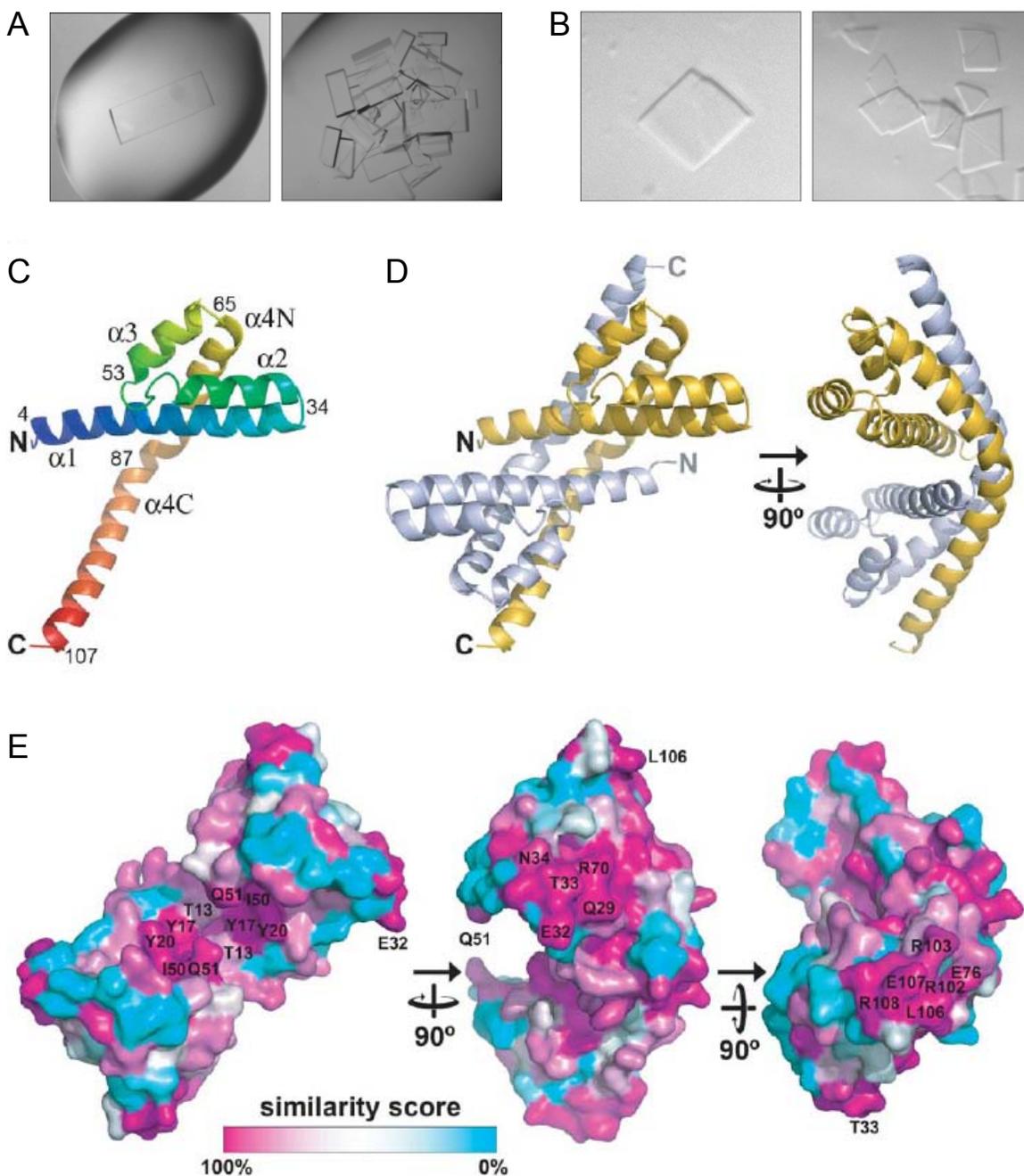


**Figure 34. Limited proteolysis of Syn7002-RbcX.**

Syn7002-RbcX was digested on ice at a concentration of 1 mg/ml at indicated concentrations of subtilisin in 20 mM HEPES-NaOH pH 7.5, 50 mM NaCl. Samples were withdrawn after 60 min, stopped with 10 mM PMSF and analyzed by 16 % Tricine-PAGE. \*Digest of the flexible C-terminus ( $\Delta$  aa 110-134) was confirmed by Edman degradation and mass spectrometry.

Crystal of both Syn7002-RbcX and (SeMet)-Syn7002-RbcX( $\Delta$ C25) were obtained using the hanging drop vapor diffusion method at 20 °C (Fig. 35 A and B). The crystals were subjected to X-ray structural analysis in cooperation with Dr. Andreas Bracher (Table A6). Unmodified Syn7002-RbcX crystallized in space group  $P4_12_12$ , containing three dimers per asymmetric unit. Its structure was solved by Pt-SIRAS at a resolution of 2.8 Å. The structure of (SeMet)-Syn7002-RbcX( $\Delta$ C25), with space group  $P1$  and six dimers per unit cell, was solved at 2.7 Å. Whereas continuous electron density was observed throughout the core domains of all protein chains, their termini were progressively disordered, presumably due to mobility as indicated by proteolytic susceptibility (Fig. 34).

The structure of the Syn7002-RbcX core domain, comprising residues 4-107, shows an unusual helix bundle consisting of four  $\alpha$ -helices ( $\alpha$ 1-4) per monomer (Fig. 35 C). Helix  $\alpha$ 2 (residues 35-48) turns backwards relative to  $\alpha$ 1 (residues 4-33) with only one residue (N34) forming a turn of 151°. At the junction of  $\alpha$ 3 (52-63) and  $\alpha$ 4 (residues 65-107) a similar arrangement with an inter-helix angle of 161° can be found. Comprising only 3 and 2 turns, respectively, the central helices  $\alpha$ 2 and  $\alpha$ 3 are comparatively short. They are connected by a linker of 5 residues. Typical coiled-coil interactions between the side chains cannot be observed in the helix bundle, but its core contains a high density of conserved hydrophobic residues (Fig. 36). Helix  $\alpha$ 4 kinks by ca. 60° in the vicinity of residue 84 and forms a 35 Å long extension ( $\alpha$ 4C) pointing away from the helix bundle.



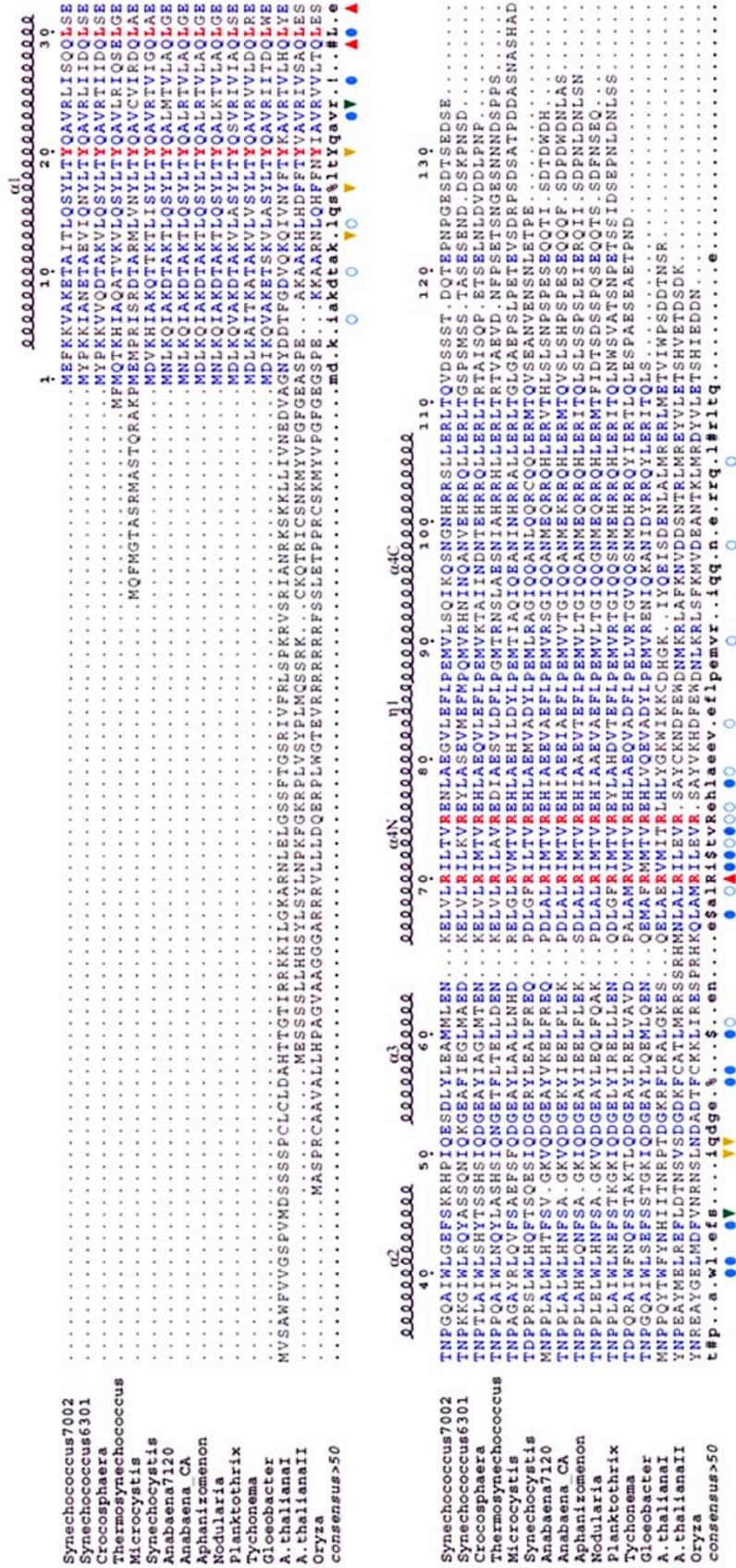
**Figure 35. Crystal structure of Syn7002-RbcX.**

(A-B) Crystals of Syn7002-RbcX (A) were obtained in 0.1 M HEPES-NaOH, pH 7.5, 1.4 M NaOAc. (SeMet)-Syn7002-RbcX( $\Delta$ C25) (B) crystallized in 0.1 M Tris-HCl, pH 8.5, 0.2 M NaOAc, 30 % PEG4000.

(C) Ribbon representation of the RbcX monomer. The peptide backbone is depicted from N- to C-terminus using a color gradient from blue to red. Secondary structure elements, selected residue numbers and chain termini are indicated. Figures C-E were prepared by Dr. Andreas Bracher.

(D) Structure of the RbcX dimer. Protomers are shown in yellow and blue.

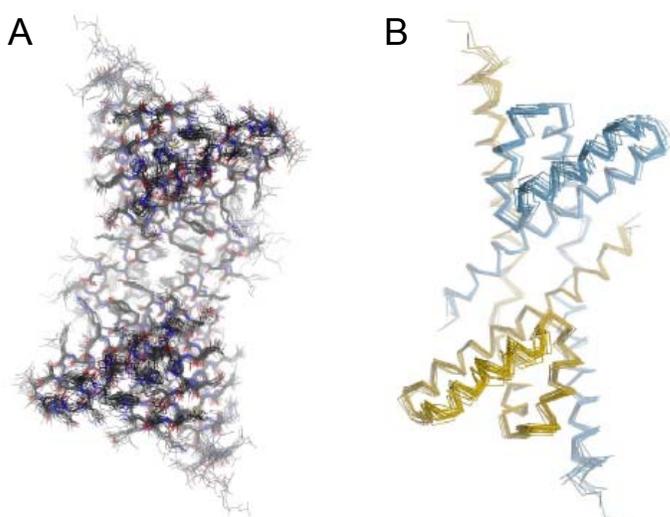
(E) Surface conservation in RbcX. The similarity score from an alignment of 151 sequences of cyanobacterial RbcX in the PFAM database was plotted onto the accessible surface of the RbcX dimer. Sequence conservation is indicated by a color gradient showing highly conserved residues in magenta and variable regions in cyan. Positions of conserved surface residues are indicated.



**Figure 36. Alignment of amino acid sequences of selected RbcX homologs.**

Amino acid sequences of selected RbcX homologs from various cyanobacterial species and from two higher plants were aligned using the MultAlin analysis program (Corpet 1988). Residues indicated in red are invariable and residues indicated in blue are conserved in more than 50 % of the sequences. A consensus sequence according to these criteria using the standard symbols of MultAlin 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 yellow and green inverted triangles, respectively. Red triangles denote residues at the corner surfaces of RbcX implicated by mutagenesis in binding to an uncharacterized surface of RbcL. Open blue circles denote structural residues involved in dimerization and filled blue circles indicate residues forming the hydrophobic core of the four-helix bundle. Structure and alignment analysis was performed in cooperation with Dr. Andreas Bracher. SwissProt accession numbers of the sequences are listed in table A4.

The dimer of Syn7002-RbcX has overall dimensions of approximately 64 x 33 x 31 Å (length x height x width). As depicted in Fig. 35 D, the two protomers align in an almost anti-parallel fashion along the  $\alpha$ 4-helices, such that the helical bundles are located at opposed ends, thus allowing the central  $\alpha$ 1 helices to form additional symmetrical contacts. On the whole, the RbcX dimer can be described as an arc-shaped complex with a narrow diagonal groove which is primarily lined by helices  $\alpha$ 1 and  $\alpha$ 2 (Fig. 35 D and E). The central part of the complex, containing the bottom of that groove, is the most rigid region, whereas helices  $\alpha$ 2 and  $\alpha$ 3 represent the most mobile elements (Fig. 37).



**Figure 37. Crystal structure of RbcX.**

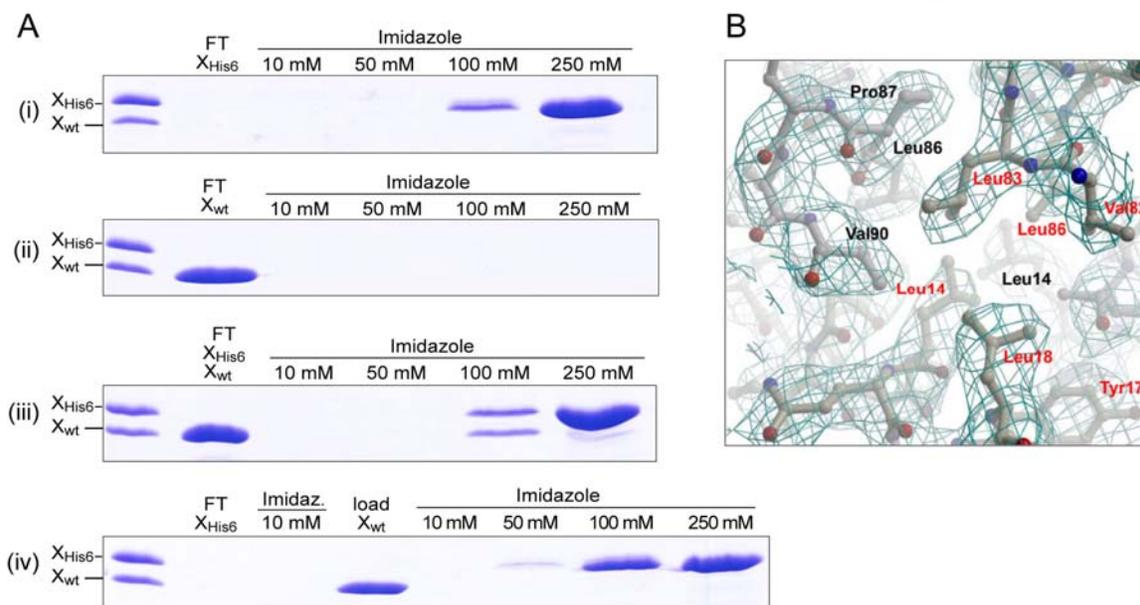
(A) Conformational ensemble generated by superposition of 18 individual conformations observed in the crystal lattices. The average r.m.s. deviation for  $C\alpha$  positions was determined to be 0.623 Å. Covalent bonds are shown as sticks. Carbon, nitrogen, oxygen and sulfur atoms are depicted in black, blue, red and yellow, respectively.

(B) Superposed peptide backbones of RbcX dimers. Helix  $\alpha$ 2 in the foreground appears to be the most mobile element in the structure. Figures were prepared by Dr. A. Bracher.

The dimer interface of Syn7002-RbcX is predominantly uncharged and hydrophobic and important intermolecular contact residues appear to be V6, T10,L, Y17, M61, L69, L72, T73, A79, L83, V90, and L105 (Fig. 36). Additionally, dimer stability appears to be supported by a polar network around the conserved residue N98 (contacting the backbone residue L72) and R75/E76 in the opposing protomer. Interestingly, the RbcX dimer could be a candidate for a domain-swapping structure with helix  $\alpha$ 4C as the swapped domain

(Liu and Eisenberg 2002). However, the double mutant R75L/E76A resulted in a considerable decrease of RbcX solubility (Fig. 40 B and C), presumably due to destabilization of the dimeric structure. At the central junction between the helix bundles, tight packing of the conserved hydrophobic residues L83, L86, P87, M89, and V90 between opposing strands and against the hydrophobic residues L14 and L18 from helices  $\alpha 1$  of both protomers was observed (Fig. 38 B). Disruption of these hydrophobic interactions by introducing charged residues at the dimer interface (L14E or L14E/V90E) resulted in the formation of insoluble RbcX. Furthermore, the approach to generate a monomeric form of Syn7002-RbcX by truncation of  $\alpha 4C$  produced insoluble Syn7002-RbcX(aa 1-83) (Fig. 40 B and C). These results clearly indicate the importance of the polar and hydrophobic interactions for the integrity of the dimeric protein fold and suggest that the existence of RbcX monomers is unlikely.

In order to confirm experimentally the stability of RbcX dimers or the absence of monomer-dimer equilibrium, the formation of mixed dimers between His-tagged and wildtype Syn7002-RbcX was examined as follows. Syn7002-RbcX<sub>N-His6</sub> bound to Ni-NTA-agarose and eluted upon application of high concentrations of imidazole, whereas untagged Syn7002-RbcX<sub>wt</sub> showed no affinity to the resin (Fig. 38 A, i and ii). When both types of RbcX were incubated together for 1 h and were subsequently applied to the Ni-NTA-agarose affinity column, the majority of untagged RbcX was in the flowthrough and His-tagged RbcX eluted only at high imidazole. However, little amounts of both RbcX types eluted in 100 mM imidazole in equal ratio. Thus, it can be assumed, that they formed mixed RbcX<sub>N-His6</sub>/RbcX<sub>wt</sub> dimers in solution, although the efficiency of this monomer exchange reaction appears to be very low (Fig. 38 A, iii). When Syn7002-RbcX<sub>N-His</sub> was firstly bound to Ni-NTA-agarose and afterwards incubated with Syn7002-RbcX<sub>wt</sub>, only His-tagged RbcX and no mixed complexes were observed, indicating that RbcX loses its ability to “cycle” as soon as it is bound to an affinity-matrix. Based on these results and considering the above described crystallographic and mutational analyses, it can be concluded that Syn7002-RbcX dimers are highly stable.

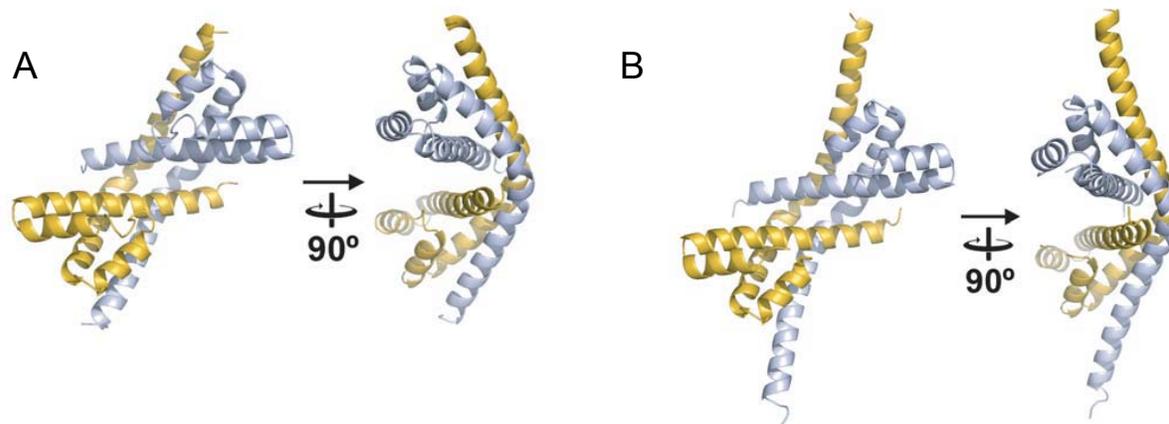


**Figure 38. Syn7002-RbcX dimer stability.**

(A) Minute monomer exchange in Syn7002-RbcX dimers. Syn7002-RbcX<sub>N-His6</sub> (i) and Syn7002-RbcX<sub>wt</sub> (ii) were applied to a Ni-NTA-agarose column, which was eluted with increasing concentrations of imidazole as indicated. (iii) A mixture of both proteins was incubated for 1 h at RT and subsequently loaded onto the column. (iv) Syn7002-RbcX<sub>N-His6</sub> was bound to Ni-NTA-agarose, before the matrix was washed with low imidazole and incubated with Syn7002-RbcX<sub>wt</sub> for 1 h at RT. Afterwards, the resin was washed and bound protein eluted with increasing concentrations of imidazole. FT, flowthrough.

(B) Model showing residues of hydrophobic interaction in the central junction of the Syn7002-RbcX dimer. Contact residues on opposing monomers are indicated in black and red. Covalent bonds are shown as sticks. Carbon, nitrogen and oxygen atoms are indicated by black, blue and red color, respectively. Electron density is depicted as meshwork in cyan. The model was prepared by Dr. Andreas Bracher.

Since RbcX homologs from different cyanobacterial species and orders share a high degree of sequence conservation, it can be supposed that they also have the same overall protein architecture. This assumption was supported by the observation that RbcX from *Synechococcus* sp. PCC7002 and from another cyanobacterium, *Anabaena* sp. CA, revealed nearly identical crystallographic structures (r.m.s.d. 1.14 Å over 103 Cα atoms, Fig. 39, Table A6).



**Figure 39. Structure comparison of RbcX.**

Structures of Syn7002-RbcX (A) and AnaCA-RbcX (B). A residual mean square deviation of 1.09 Å for 190 C $\alpha$  atoms was calculated, indicating high structure identity in the common region.

#### 4.3.5 Protein interaction surfaces on RbcX

Based on the RbcX sequence conservation (Fig. 35 E and Fig. 36) several residues of Syn7002-RbcX were subjected to a mutational analysis in order to find out if they are of functional or structural importance. RbcX mutants were co-expressed with Syn7002-RbcL in *E. coli* and the resulting soluble lysates were analyzed for the presence of soluble RbcX and RbcL as well as for RbcL<sub>8</sub> complex formation and potential activity (Fig. 40).

As described above, some mutations (L14E, L14E/V90E, R75L/E76A and  $\Delta$ 84-134) severely impaired dimer formation and caused RbcX to become insoluble or only poorly soluble (Fig. 40 B and C, triangles). Consequently, they could not support the production of soluble RbcL (Fig. 40 A, triangles), resulting in a situation resembling the expression of RbcL alone (Fig. 40 A, lane 1). All other RbcX mutants were expressed as soluble proteins and since their migration pattern in Native PAGE was similar to wildtype RbcX (except for the smaller deletion mutants), the above described dimer structure can be assumed to be valid for all of them (Fig. 40 B and C). Several RbcX mutations did not considerably affect the solubility or assembly of Syn7002-RbcL (Fig. 40 A, circles). However, the mutational analyses confirmed the importance of the conserved regions of RbcX as potential interaction sites between RbcX and RbcL: the central groove of the dimer and the peripheral polar regions around the corners of the molecule.

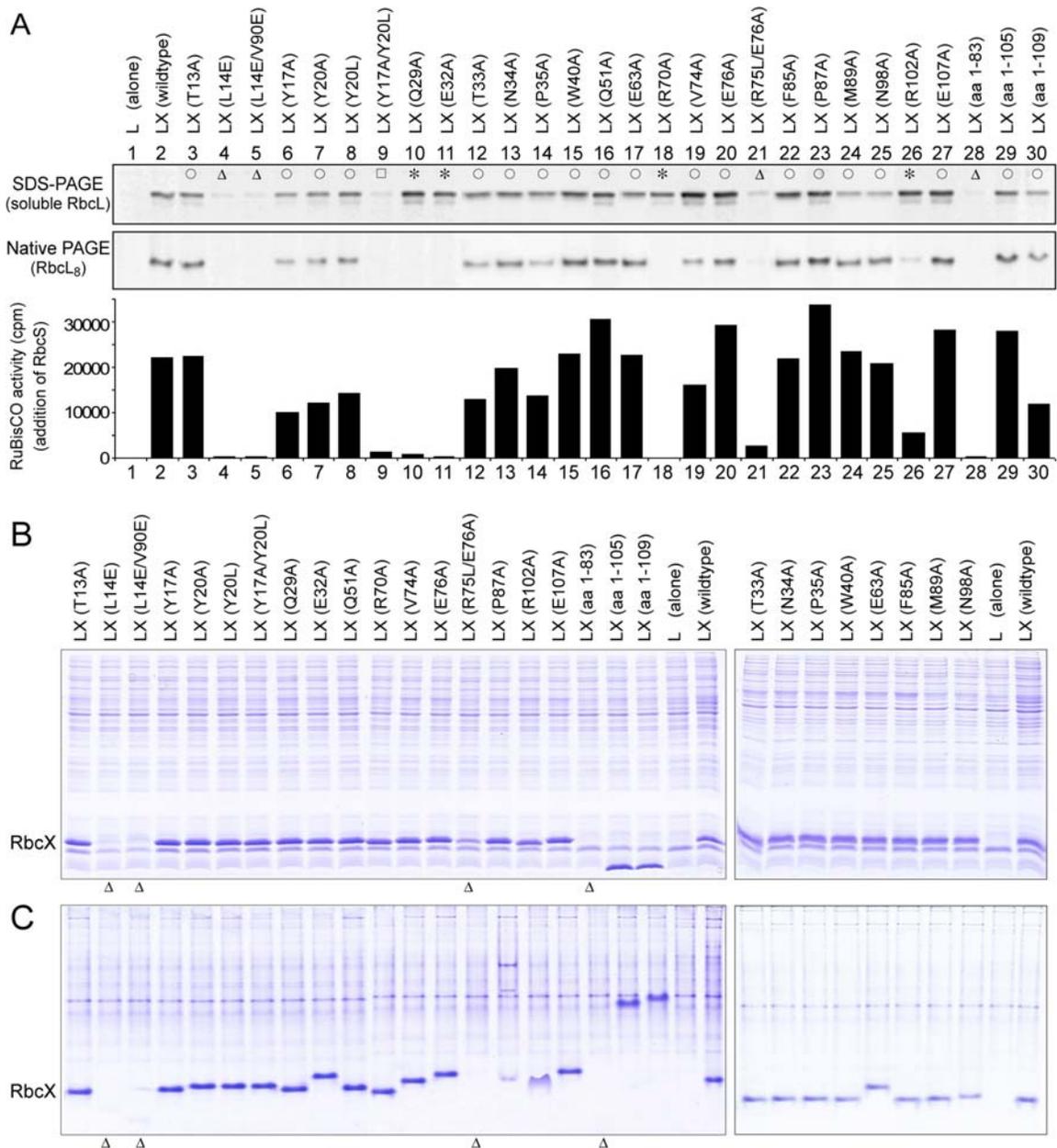
On the one hand, the double mutant Y17A/Y20L prevented the formation of soluble RbcL almost completely (Fig. 40 A, square). When only the single tyrosines were

mutated, solubility and assembly of RbcL were reduced, but not abolished (Fig. 40 A, lanes 6-8). Hence, the functional cooperation of both Y17 and Y20 is required for proper operation of RbcX. Y17 and Y20 form in concert with I50 a highly conserved hydrophobic area which is located on the surface of each protomer in the central groove of the RbcX dimer (Fig. 35 E). With a central access of 5.4 Å, this groove is wide enough to accommodate a polypeptide chain in an extended conformation and it thus represents a potential site for the interaction of RbcX with an unstructured peptide of RbcL.

On the other hand, the mutations Q29A, E32A, R70A and R102A allowed the accumulation of soluble RbcL but did not support its proper assembly to RbcL<sub>8</sub> cores, as indicated by Native PAGE and by the absence of RuBisCO activity upon supplementation with RbcS (Fig. 40 A, asterisks). All these critical residues are located in conserved polar surface regions around the edges of the RbcX dimer (Fig. 35 E). These areas are formed by helices  $\alpha_4$  and by the turn region between helices  $\alpha_1$  and  $\alpha_2$ . They might be suitable for interaction with the corresponding polar surface of the RbcL subunit.

Mutation of other residues which are located within or in close proximity to these two regions (e.g. T13, T33, N34, E76, E107 or  $\Delta$ 106-134) were without measurable effect, presumably due to minor or different functional relevance.

Since the crystal structure of three Syn7002-RbcX mutants (Y17A/Y20L, Q29A and R70A) was nearly identical to that of wildtype RbcX (Table A6), the functional significance of both potential protein-protein interaction areas (central groove and peripheral corner surfaces) is assignable to both mutant and wildtype RbcX.



**Figure 40. Functional analysis of Syn7002-RbcX mutants.**

(A) Syn7002-RbcL was co-expressed in *E. coli* with wildtype or mutant Syn7002-RbcX as indicated. Soluble cell lysate was analyzed for soluble RbcL by 16 % SDS-PAGE and for RbcL<sub>8</sub> complex assembly by 6 % Native PAGE, followed by immunoblotting against RbcL. Carboxylation activity in the soluble cell lysate was measured upon addition of purified Syn6301-RbcS. Syn7002-RbcL expressed without RbcX (1) or with wildtype RbcX (2) served as negative or positive control, respectively. (o) RbcX mutants similar to wt RbcX, supporting synthesis of soluble RbcL and properly assembled RbcL<sub>8</sub> with carboxylation activity in presence of RbcS. (\*) RbcX mutants supporting production of soluble RbcL but not or hardly their assembly to RbcL<sub>8</sub> complexes. (Δ) RbcX mutants that are not or very poorly soluble themselves and therefore not or only poorly supporting production of soluble RbcL. (□) RbcX is soluble but not competent to allow production of soluble RbcL.

(B-C) Soluble cell lysate from expressions in (A) was analyzed by 17.5 % SDS-PAGE (B) and by 6 % Native PAGE (C), followed by Coomassie blue staining. Indicated RbcX mutants (Δ) were insoluble or poorly soluble and thus mainly found in the pellet fraction (not shown).

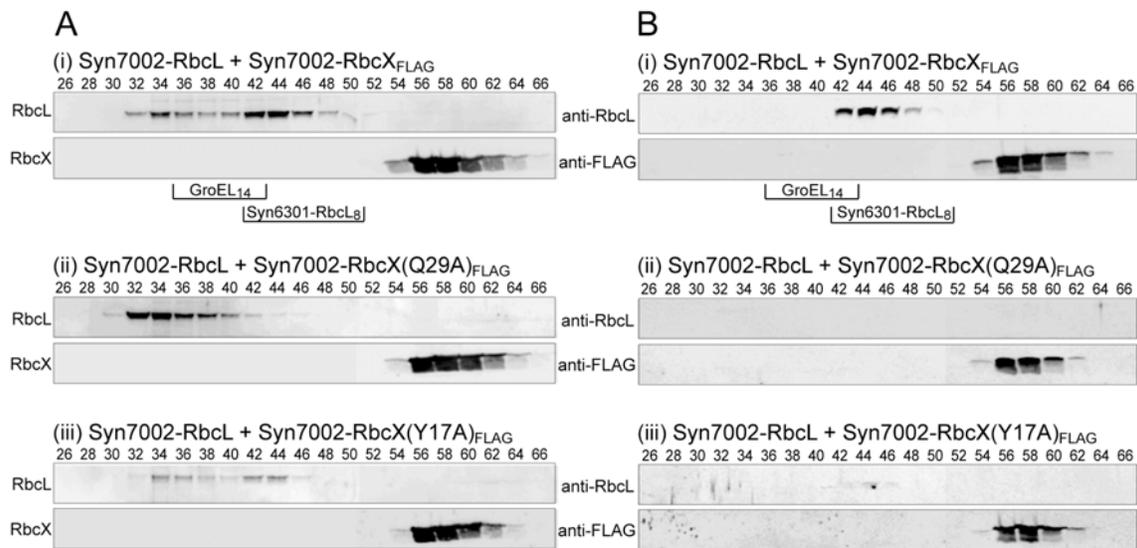
In order to characterize the interaction of Syn7002-RbcL with wildtype or mutant Syn7002-RbcX, RbcL was co-expressed in *E. coli* with N-terminally FLAG-tagged RbcX and soluble cell lysates were subjected to gel filtration chromatography.

When Syn7002-RbcL was co-expressed with functionally unimpaired wildtype Syn7002-RbcX<sub>FLAG</sub>, the majority of soluble RbcL fractionated as complex of approximately 400 kDa, consistent with the elution pattern of purified Syn6301-RbcL<sub>8</sub> cores. Assembly intermediates, such as RbcL monomers or dimers, could not be detected, suggesting that assembly of RbcL subunits into L<sub>8</sub> cores is a very efficient process. Some RbcL eluted in very early fractions and was most likely misassembled, aggregated RbcL (> 800 kDa), migrating with a molecular weight behind GroEL (Fig. 41 A, i). RbcX did not co-fractionate with RbcL by gel filtration, suggesting that the interaction between Syn7002-RbcL and Syn7002-RbcX<sub>FLAG</sub> must be dynamic. However, the interaction between Syn7002-RbcL and Syn7002-RbcX<sub>FLAG</sub> was confirmed by co-immunoprecipitation from the soluble cell lysates with ANTI-FLAG affinity beads. Precipitated proteins were eluted from the immunobeads under native conditions with excess of 3x FLAG-peptide and subsequently subjected to gel filtration. RbcL fractionated as RbcL<sub>8</sub> complex with no misassembled, aggregated or monomeric RbcL observed, indicating that wildtype RbcX interacts with Syn7002-RbcL<sub>8</sub> cores (Fig. 41 B, i). However, despite their interaction during the immunoprecipitation, RbcX again failed to co-fractionate with RbcL. This observation suggests that the interaction between RbcL and RbcX is rather labile and susceptible to disruption by gel filtration.

When Syn7002-RbcL was co-expressed with Syn7002-RbcX(Q29A)<sub>FLAG</sub>, only misassembled or aggregated RbcL was observed migrating in high molecular weight fractions upon gel filtration of soluble cell lysate (Fig. 41 A, ii). This is in agreement with the observation that mutations in the peripheral polar surface areas of RbcX dimers do not support the proper assembly of soluble RbcL into RbcL<sub>8</sub> complexes, as shown by Native PAGE (Fig. 40 A, asterisks). No association between RbcL and Syn7002-RbcX(Q29A)<sub>FLAG</sub> could be detected by immunoprecipitation (Fig. 41 B, ii), confirming that the assembly failure by Syn7002-RbcX(Q29A)<sub>FLAG</sub> is probably due to unstable interaction between RbcL and RbcX, leading to misassembly and/or aggregation of RbcL.

Syn7002-RbcX(Y17A)<sub>FLAG</sub>, a mutant with a moderately impaired function of the central groove, resulted (besides some aggregated RbcL) in a reduced formation of assembled RbcL<sub>8</sub> and inefficient co-immunoprecipitation (Fig. 41 A and B, iii).

Based on the results depicted in Fig. 40 and 41, two distinct regions on RbcX are necessary for productive interaction with RbcL. The functional integrity of the central binding groove is a prerequisite for the production of soluble and thus assembly-competent RbcL. On the other hand, the polar peripheral areas at the edges of RbcX dimers are crucial for the proper assembly of soluble RbcL into RbcL<sub>8</sub> cores.



**Figure 41. Functional analysis of Syn7002-RbcX-mutants.**

Syn7002-RbcL was co-expressed with Syn7002-RbcX<sub>FLAG</sub> (i), Syn7002-RbcX(Q29A)<sub>FLAG</sub> (ii) or Syn7002-RbcX(Y17A)<sub>FLAG</sub> (iii) in *E. coli*.

(A) Soluble cell lysates were subjected to gel filtration (Superdex 200, 10/30) in 50 mM Tris-HCl, pH 8, 50 mM NaCl, 5 mM MgCl<sub>2</sub>. Fractions were analyzed for RbcL and RbcX<sub>FLAG</sub> by SDS-PAGE, followed by immunoblotting. The elution pattern for GroEL (ca. 800 kDa) and Syn6301-RbcL<sub>8</sub> (ca. 400 kDa) is indicated.

(B) Soluble cell lysates were subjected to co-immunoprecipitation of RbcL with anti-FLAG affinity beads. Pulled down proteins were eluted under native conditions with excess of 3x FLAG peptide and were analyzed by gel filtration as described in (A).

### 4.3.6 Interaction of RbcX with the C-terminus of RbcL

#### 4.3.6.1 The C-terminal RbcL motif EIKFEFD is recognized and bound by RbcX

As demonstrated above, Syn7002-RbcX and Syn7002-RbcL are interacting *via* two potential sites, the central hydrophobic groove as well as the peripheral polar corner surfaces. Since the crevice of RbcX is only large enough to bind a peptide in an extended conformation, the next step was to look for a sequence in RbcL that is capable to bind in this groove.

For this purpose, a peptide array of the Syn7002-RbcL sequence was screened for interaction with wildtype Syn7002-RbcX, the peripheral polar surface mutant Syn7002-RbcX(Q29A), the groove mutant Syn7002-RbcX(Y17A/Y20L) and a heterologous RbcX from the cyanobacterium *Anabaena* sp. CA (AnaCA-RbcX). The peptide array consisted of 231 peptides which covered the entire sequence of Syn7002-RbcL. Peptides of 12 amino acids in length, overlapping with 10 amino acids, were N-terminally acetylated and C-terminally covalently attached to a cellulose-PEG-membrane. This peptide array was incubated with the different types of RbcX, unbound protein was removed by washing and bound RbcX was detected with a specific antiserum against RbcX.

The most intense interaction of wildtype Syn7002-RbcX was found with four adjacent peptides, that represented the C-terminus of Syn7002-RbcL and that shared the sequence motif EIKFEFD (Fig. 42 A and B). Interestingly, a number of other peptides showing weaker binding contained one or two phenylalanines (F), consistent with the consensus sequence EIKFEFD. The peripheral polar surface mutant Syn7002-RbcX(Q29A) bound to the same C-terminal peptide as wildtype RbcX. However, since Syn7002-RbcL could not be co-immunoprecipitated with Syn7002-RbcX(Q29A)<sub>FLAG</sub> (Fig. 41 B, ii), it can be assumed that binding of RbcL in the groove of RbcX is not sufficient to maintain a stable interaction between RbcX and RbcL. Furthermore, the peptide binding pattern also confirmed that the central groove of Syn7002-RbcX(Q29A) is functional and that its impotence to support proper RbcL assembly is indeed ascribable to alterations in the polar peripheral surface areas of RbcX. The groove mutant Syn7002-RbcX(Y17A/Y20L), on the other hand, did not bind to any of the Syn7002-RbcL peptides, thereby verifying that the functional integrity of the groove is essential for the initial binding to RbcL. Interestingly, the heterologous AnaCA-RbcX also bound to the

same C-terminal peptides of Syn7002-RbcL. The common binding motif (EIKFEFD) is fairly conserved among RbcL of cyanobacteria, green algae and higher plants (Fig. 42 C) and thus it can be assumed that the interaction of RbcX with the C-terminus of type I RbcL is a universal mechanism. In comparison, the dimeric type II RuBisCO of the proteobacterium *Rhodospirillum rubrum* does not require such a mechanism, which is in agreement with the absence of both RbcX and the C-terminal RbcL binding motif in this bacterium (Fig. 42 D).

Unambiguous specific or common binding motifs for the interaction of Syn7002-RbcL with other molecular chaperones (GroEL, trigger factor, DnaK or DnaJ) or with Syn7002-RbcS<sub>FLAG</sub> could not be found by this method. Notably, Syn7002-RbcS<sub>FLAG</sub> bound in a very unspecific manner to numerous RbcL peptide spots, suggesting that folded and/or assembled RbcL is required to provide a definite binding site for RbcS (Data not shown).

#### 4.3.6.2 Structure of the RbcX/RbcL-peptide complex

In order to verify and visualize the binding of the Syn7002-RbcL C-terminal peptide, the consensus peptide binding motif (EIKFEFD) was used soaked native crystals of Syn7002-RbcX and the structure of the co-crystals was determined.

Indeed, the RbcL peptide was found to bind in the central hydrophobic groove of RbcX (Fig. 42 E, Table A6). There, it adopts an extended conformation with the side chains of F462 and F464 extending into cavities lined by the sidechains of T13, Y17, Y20, I50, and Q51 of each monomer of RbcX. The polar sidechains of the peptide face outwards and are mostly disordered. I460 of the peptide contacts Y20, R24, and S45 of one RbcX chain. In addition, three hydrogen bonds are formed between the peptide backbone and the side chains of the residues Y20, R24, and Q51 of RbcX. 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 (Fig. 42 F). The opening of the RbcX crevice is also constricted to  $\sim 5.4$  Å by the side-chains of the Q51 residues.

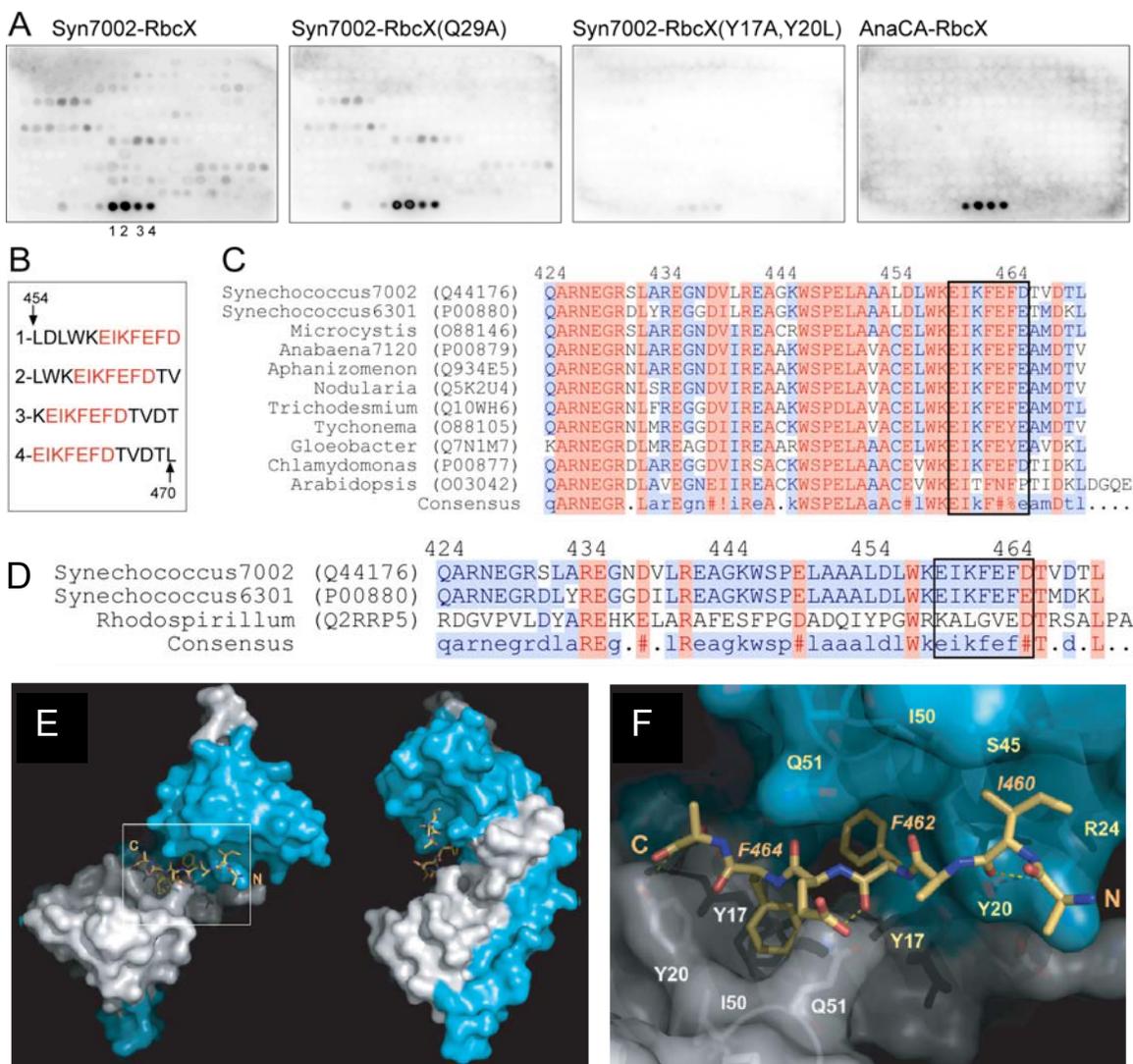
#### 4.3.6.3 Low affinity of Syn7002-RbcX for C-terminal RbcL-peptides

The strength of the interaction between RbcX and the C-terminus of RbcL was analyzed by isothermal titration calorimetry. Two peptides covering different lengths of the Syn7002-RbcL C-terminus were used for this analysis. The first peptide EIKFEFD represented the consensus binding motif for RbcX, observed with the peptide array. The second peptide is a longer peptide KEIKFEFDTVD, which also included the highly conserved residue D468. The latter might be involved in stabilizing the tight position of the C-terminal RbcL strand when the active site of RuBisCO is in its closed state (Duff *et al.* 2000).

Syn7002-RbcX had a rather low affinity to both peptides, as seen by the high dissociation constants ( $K_D$ ) of approximately 231  $\mu\text{M}$  for EIKFEFD and 158  $\mu\text{M}$  for KEIKFEFDTVD (Fig. 43 A and B). However, the high dissociation constants are an indication that the interaction between Syn7002-RbcX and Syn7002-RbcL peptides is unstable or dynamic in nature. These dissociation rates could also explain the instability of the Syn7002-RbcLX-complexes observed during gel filtration (Fig. 41 A and B, i) or Native PAGE.

Peptide titration was also performed with the heterologous AnaCA-RbcX, which was shown to bind to the C-terminal peptides of Syn7002-RbcL more specifically (Fig. 42 A). With dissociation constants of only 60  $\mu\text{M}$  for EIKFEFD and 5  $\mu\text{M}$  for KEIKFEFDTVD, the affinity of AnaCA-RbcX for the peptides was significantly higher (Fig. 43 C and D). This might be caused by structural differences of AnaCA-RbcX, such as a shorter loop region between helices  $\alpha_2$  and  $\alpha_3$  or an additional positive charge (K49) close to the central groove (Fig. 36).

For both Syn7002-RbcX and AnaCA-RbcX the titration analysis confirmed that the stoichiometry of interaction is one RbcL-peptide bound to one dimer of RbcX, which is also shown by structural analysis (Fig. 42 E).



**Figure 42. Interaction of RbcX with the C-terminus of RbcL.**

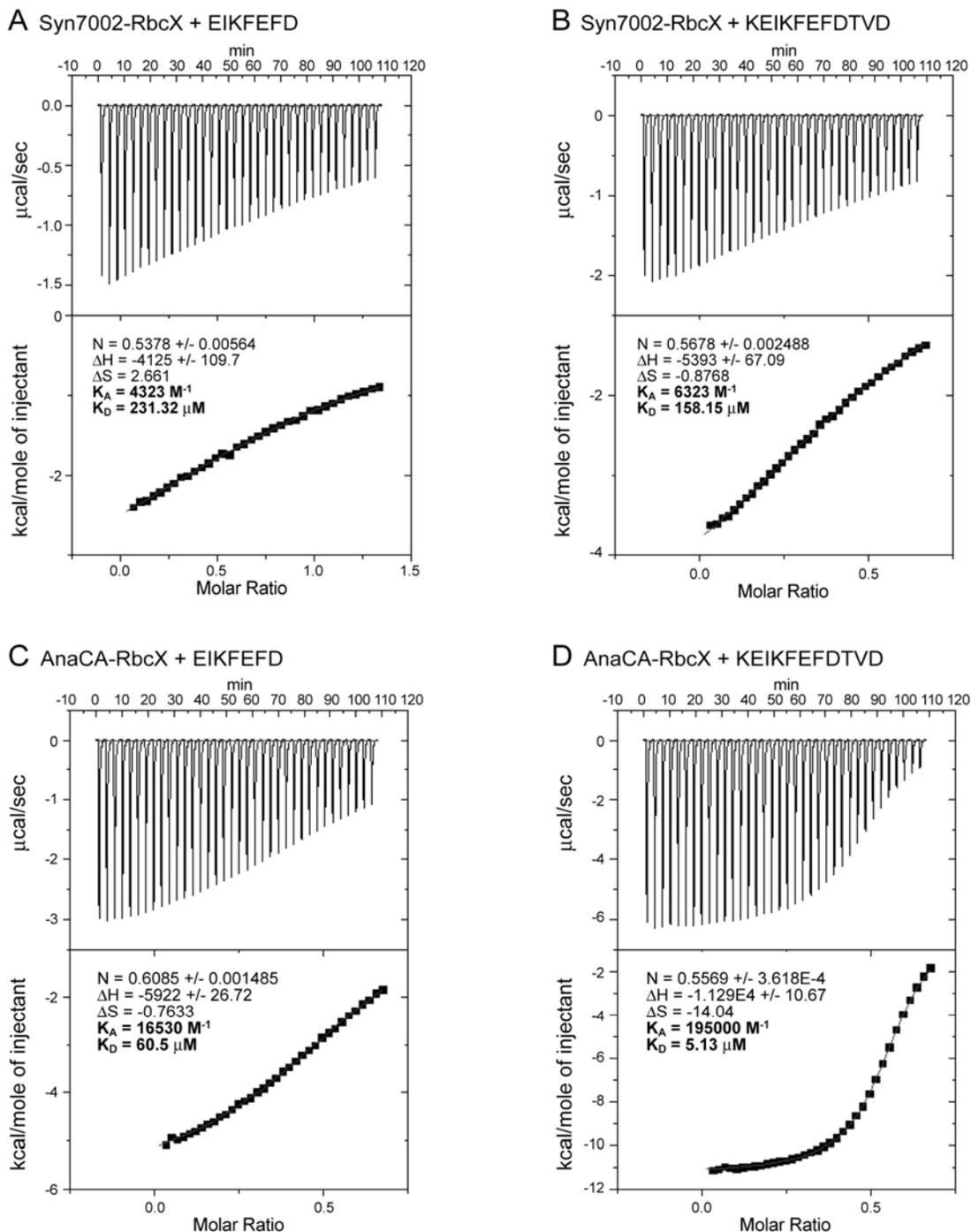
(A-B) Overlapping dodecameric peptides covering the sequence of Syn7002-RbcL were synthesized on a cellulose membrane. The membrane was probed with the indicated RbcX proteins (10  $\mu$ g/ml). Peptide-bound RbcX was visualized by immunodetection with RbcX specific antiserum. Sequences of the indicated binding spots with the consensus (red) are listed in (B).

(C) Alignment of C-terminal RbcL amino acid sequences of the indicated cyanobacterial, green algal and higher plant species, performed with MultAlin (Corpet 1988). High consensus level ( $\geq 90\%$ ) is depicted in red and low consensus level ( $\geq 50\%$ ) in blue. Protein accession numbers for Swiss-Prot are given in brackets or in Table A5. Residue numbers are based on Syn7002-RbcL.

(D) Alignment of RbcL C-termini of *Synechococcus* sp. PCC7002 and PCC6301 (cyanobacteria, type I RuBisCO) and proteobacterium *Rhodospirillum rubrum* (Rhodospirillales, type II RuBisCO).

(E) Structure of the complex of RbcL-peptide EIKFEFD bound to a dimer of Syn7002-RbcX. The peptide is shown in ball-and-stick-presentation with indicated N- and C-termini. RbcX is depicted as a molecular surface with protomers colored in white and blue, respectively. Structural analyses were performed by Dr. Andreas Bracher.

(F) Magnified boxed area from (E) showing the molecular interactions between the RbcL-peptide and RbcX. Hydrogen bonds are indicated as dashed lines. Residues (white or yellow) of RbcX monomers participating in peptide binding are shown in ball-and-stick presentation below the transparent surface of the molecule. Important hydrophobic peptide residues are labeled as well.

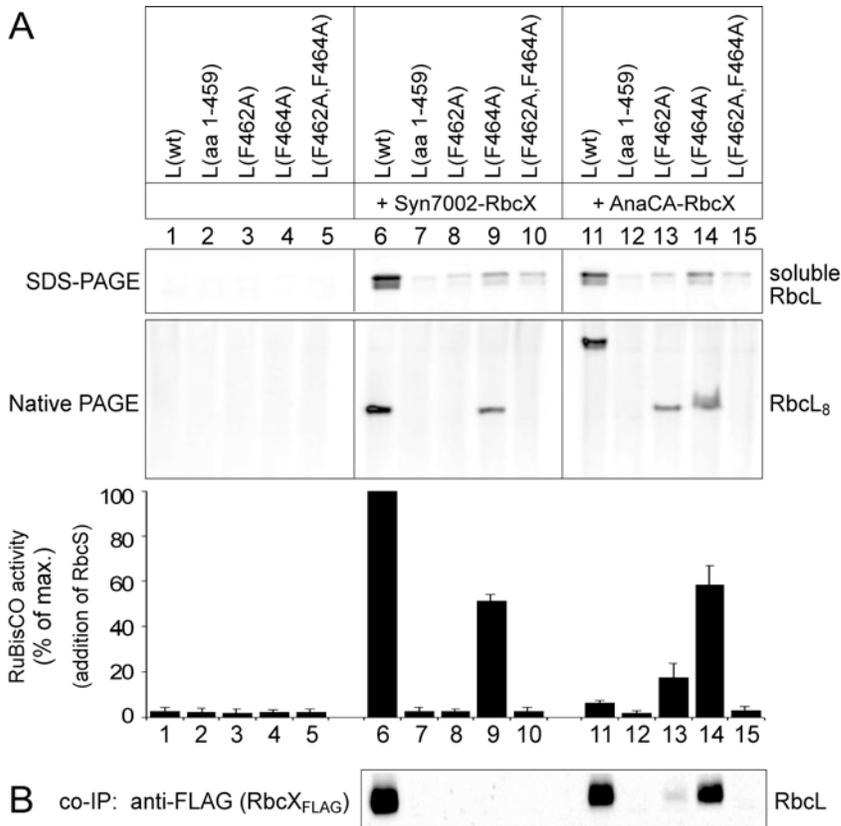


**Figure 43. Analysis of RbcX–RbcL(peptide) interaction by isothermal titration calorimetry.** Titration of Syn7002-RbcX with Syn7002-RbcL peptide EIKFEFD (A) or KEIKFEFDTVD (B). Titration of AnaCA-RbcX with peptide EIKFEFD (C) or KEIKFEFDTVD (D). Data were corrected for titration in absence of RbcX. Raw data and integrated data of net heat exchange are depicted on top or bottom, respectively. For the latter, the best fit curve was obtained by least-squares deconvolution. Presented are the reaction stoichiometry ( $N$ , peptide: $X_{\text{mon}}$ ), enthalpy change ( $\Delta H$ , cal/mole), entropy ( $\Delta S$ , cal/mole/deg), binding constant ( $K_A$ ,  $\text{M}^{-1}$ ), dissociation constant ( $K_D$ ,  $\mu\text{M}$ ).

#### **4.3.6.4 The RbcL C-terminus (in particular its phenylalanines) is essential for interaction with RbcX and thus for proper RbcL<sub>8</sub> assembly in *E. coli***

So far, analysis of the interaction between RbcX and the C-terminus of RbcL was restricted to *in vitro* experiments using synthetic RbcL peptides. Consequently, the functional relevance of the RbcL C-terminus for interaction with RbcX and for proper assembly had also to be analyzed in the cellular environment of *E. coli*.

For this purpose, C-terminal mutants of Syn7002-RbcL were co-expressed with RbcX in *E. coli* and compared to wildtype RbcL with regard to the formation of soluble RbcL and assembly of RbcL<sub>8</sub> complexes. Neither wildtype nor C-terminal mutants of Syn7002-RbcL resulted in the production of soluble RbcL when RbcX was absent (Fig. 44 A, lanes 1-5). Upon coexpression of Syn7002-RbcX, wildtype Syn7002-RbcL formed soluble RbcL<sub>8</sub> cores which assembled with RbcS to active holoenzyme, and RbcL could efficiently be co-immunoprecipitated with FLAG-tagged RbcX (Fig. 44 A and B, lane 6). However, after deletion of the C-terminal recognition motif, Syn7002-RbcL(aa 1-459) was mainly expressed as insoluble protein and was not found to associate with co-expressed RbcX (Fig. 44 A and B, lane 7). This observation confirms the functional relevance of the C-terminal RbcL strand. Within the C-terminus, the two phenylalanines seem to be of special importance. Mutation of the second phenylalanine (Syn7002-RbcL(F464A)) resulted in the production of reduced amounts of soluble and assembled RbcL and absence of RbcL-RbcX interaction upon co-immunoprecipitation (Fig. 44 A and B, lane 9). However, mutation of the first phenylalanine (Syn7002-RbcL(F462A)) or the double phenylalanine mutant Syn7002-RbcL(F462A/F464A) produced insignificant amounts of soluble RbcL and assembly of RbcL<sub>8</sub> or RbcL-RbcX interaction were basically absent (Fig. 44 A and B, lanes 8 and 10), suggesting that the first phenylalanine may be more important in anchoring RbcL to RbcX.

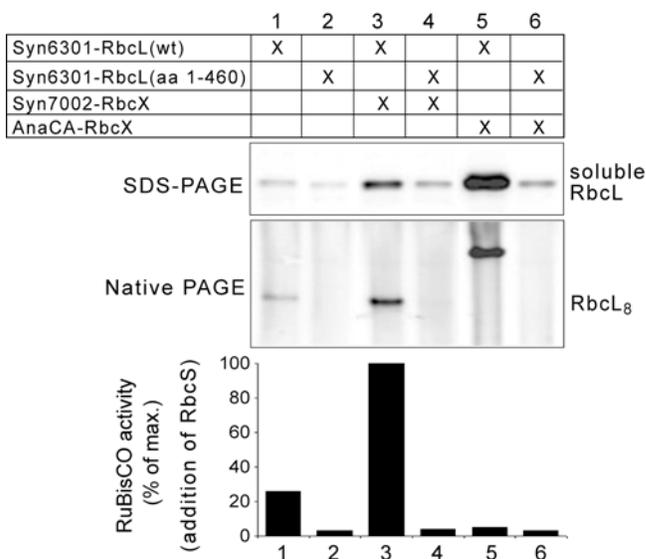


**Figure 44. Effect of C-terminal mutations in Syn7002-RbcL on RbcX binding strength and RbcL solubility and assembly.**

(A) Syn7002-RbcL wildtype or indicated mutants were expressed in *E. coli* either alone or with Syn7002-RbcX or AnaCA-RbcX. Soluble cell lysate was analyzed for RbcL by 16 % SDS-PAGE and for RbcL<sub>8</sub> complex formation by 6 % Native PAGE, followed by immunoblotting. Carboxylation activity was measured upon addition of Syn6301-RbcS and data present averages of three independent assays.

(B) Soluble lysates from coexpression of Syn7002-RbcL wildtype or mutants with Syn7002-RbcX<sub>FLAG</sub> or AnaCA-RbcX<sub>FLAG</sub> were subjected to immunoprecipitation with ANTI-FLAG affinity beads overnight at 4 °C. Co-immunoprecipitated proteins were analyzed by SDS-PAGE and RbcL was detected by immunoblotting.

Similar observations could be made for Syn6301-RbcL, whose assembly is supported by RbcX but not dependent on it (Fig. 45, lanes 1 and 3, Fig. 31). Deletion of the C-terminus completely eliminated the assembly of this protein, both in absence or presence of RbcX (Fig. 45, lanes 2 and 4). This allows the conclusion that the C-terminus is not only necessary for the interaction with RbcX, but that it is also important for the RbcL<sub>8</sub> assembly process itself. Otherwise the RbcX-independent Syn6301-RbcL could have assembled in absence of RbcX despite deletion of the C-terminus.



**Figure 45. Effect of the RbcL C-terminus and RbcX on assembly of Syn6301-RbcL<sub>8</sub>.**

Syn6301-RbcL wildtype or  $\Delta$ C(aa1-460) was expressed in *E. coli* either alone, with Syn7002-RbcX or AnaCA-RbcX as indicated. Soluble cell lysates were analyzed for RbcL by SDS-PAGE or Native PAGE, followed by immunoblotting against RbcL. RubisCO carboxylation activity was determined upon addition of Syn6301-RbcS.

#### 4.3.7 Dynamic interaction between Syn7002-RbcL and Syn7002-RbcX and its physiological importance

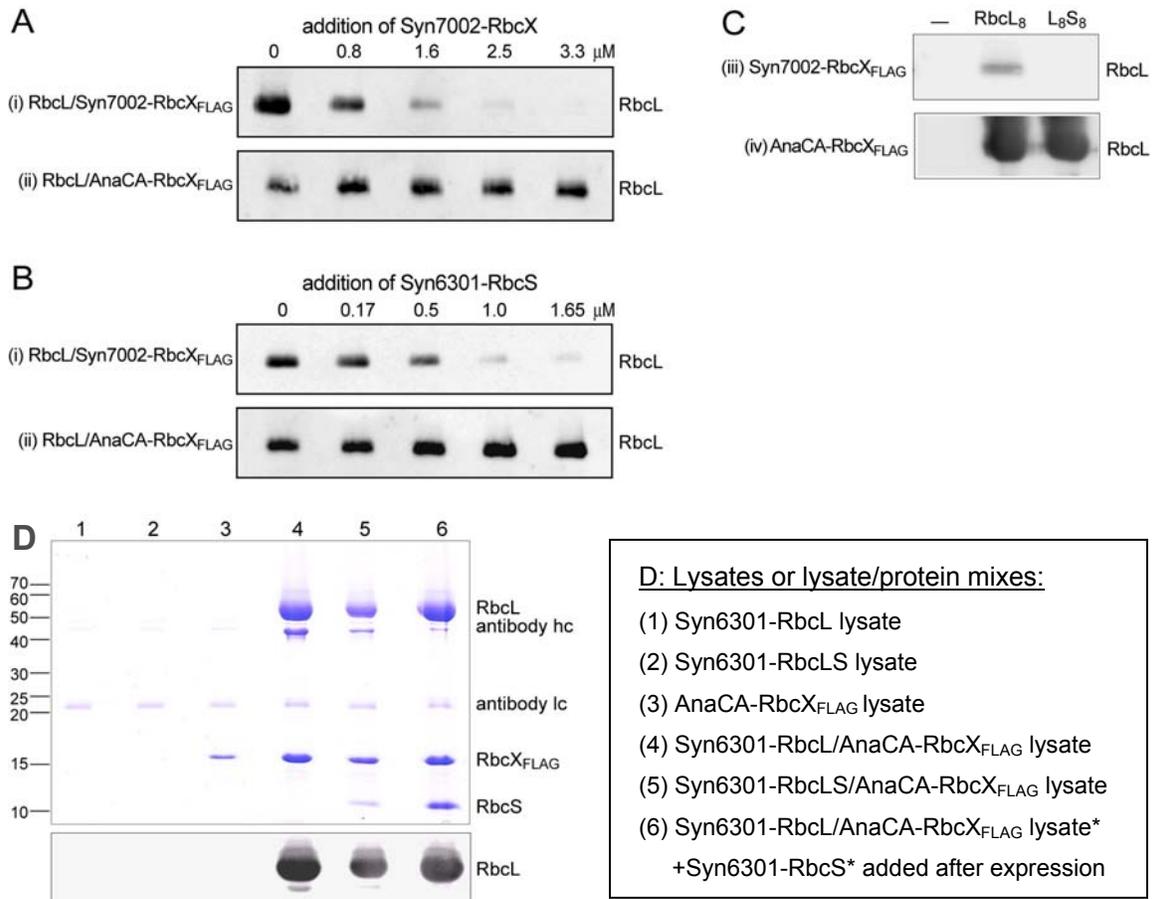
In several of the above described experiments, indications for a dynamic interaction between Syn7002-RbcL and Syn7002-RbcX were observed. Firstly, the dissociation constant for Syn7002-RbcX/RbcL-peptide was relatively high. Secondly, gel filtration or Native PAGE usually disrupted Syn7002-RbcLX-complexes which were isolated by co-immunoprecipitation. In addition, FFF-MALS of purified Syn7002-RbcL/Syn7002-RbcX<sub>N-His6</sub>-complexes revealed a complex size of ~425 kDa, consistent with the size of RbcL<sub>8</sub>, indicating a possible removal of RbcX from the complex during field flow fractionation (Table 7). Since the dynamic nature of the interaction between Syn7002-RbcL and Syn7002-RbcX might be of functional importance, it was analyzed in more detail.

In order to examine the exchangeability of RbcL-bound RbcX, complexes of Syn7002-RbcL/Syn7002-RbcX<sub>FLAG</sub> were immunoprecipitated from *E. coli* lysates co-expressing both proteins. These complexes were incubated with increasing concentrations of non-tagged Syn7002-RbcX, before immunoprecipitation was repeated and bound RbcL was detected by SDS-PAGE and immunoblotting. With increasing concentrations of non-tagged RbcX, progressively less RbcL was pulled-down, indicating displacement of tagged RbcX<sub>FLAG</sub> by non-tagged RbcX, thus demonstrating the dynamic interaction or

exchangeability (Fig. 46 A, i). The half maximal displacement of tagged RbcX was reached at  $\sim 1 \mu\text{M}$  non-tagged RbcX, reflecting a considerable higher affinity as compared to the affinity seen with only the C-terminal RbcL peptide by ITC ( $K_D \sim 160 \mu\text{M}$ , Fig. 43 B). This suggests that besides the central binding cleft of RbcX (binding the C-terminal RbcL strand) the polar peripheral surfaces of RbcX (interacting with regions of folded RbcL which remain to be determined) also contribute to the overall binding strength and dynamics.

In an analogous experiment, the displacement reaction was performed with RbcS (Fig. 46 B, i). Increasing concentrations of RbcS resulted in decreasing amounts of precipitated Syn7002-RbcL/Syn7002-RbcX<sub>FLAG</sub>-complexes, suggesting that RbcX<sub>FLAG</sub> was displaced by RbcS upon formation of RbcL<sub>8</sub>S<sub>8</sub> complexes. Since RbcX was present in molar excess over RbcS and since the displacement occurred at submicromolar concentrations of RbcS, the affinity of RbcL for RbcS exceeds that for RbcX by far. Binding of RbcS and Syn7002-RbcX to Syn7002-RbcL appears to be mutually exclusive, because displacement was complete at higher RbcS concentrations and immunoprecipitation never resulted in complexes containing all three proteins. Moreover, Syn7002-RbcX<sub>FLAG</sub> was found to interact with readily assembled RbcL<sub>8</sub> cores (purified Syn6301-RbcL<sub>8</sub>), whereas no such interaction could be observed with the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme (purified Syn6301-RbcL<sub>8</sub>S<sub>8</sub>, Fig. 46 C, i).

These results lead to the conclusion that the interaction between Syn7002-RbcL and Syn7002-RbcX must be dynamic to support the final formation of RbcL<sub>8</sub>S<sub>8</sub> holoenzyme in the following way: RbcX is required for stabilization and assembly of RbcL until formation of RbcX-associated RbcL<sub>8</sub> cores is completed. RbcS can then assemble on RbcL<sub>8</sub>, thereby enforcing the release of the loosely bound RbcX, resulting in active RbcL<sub>8</sub>S<sub>8</sub> enzyme. Re-binding of RbcX is then prevented by RbcS, presumably due to steric hinderance of attachment or recognition sites for RbcX or as a consequence of RbcS-induced conformational changes in RbcL that make the C-terminus of RbcL inaccessible for RbcX.



### Figure 46. Interaction between RbcL and RbcX.

(A) Complexes of (i) Syn7002-RbcL/Syn7002-RbcX<sub>FLAG</sub> or (ii) Syn7002-RbcL/AnaCA-RbcX<sub>FLAG</sub> were produced by co-expression in *E. coli* and immunoprecipitated from soluble cell lysate with ANTI-FLAG affinity beads overnight at 4 °C. Bead-bound complexes were incubated with the indicated concentrations of non-tagged Syn7002-RbcX for 15 min at RT. 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<sub>FLAG</sub> or (ii) Syn7002-RbcL/AnaCA-RbcX<sub>FLAG</sub>, 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 RT. Bead-bound proteins were analyzed for RbcL by SDS-PAGE and immunoblotting.

(C) Syn7002-RbcX<sub>FLAG</sub> (iii) or AnaCA-RbcX<sub>FLAG</sub> (iv) were expressed in *E. coli*. The soluble lysate was subjected directly or after addition of purified Syn6301-RbcL<sub>8</sub> (100  $\mu\text{g}/\text{ml}$ ) or Syn6301-RbcL<sub>8</sub>S<sub>8</sub> (120  $\mu\text{g}/\text{ml}$ ) to immunoprecipitation with ANTI-FLAG affinity beads overnight at 4 °C. Bead-bound proteins were analyzed for RbcL by SDS-PAGE and immunoblotting.

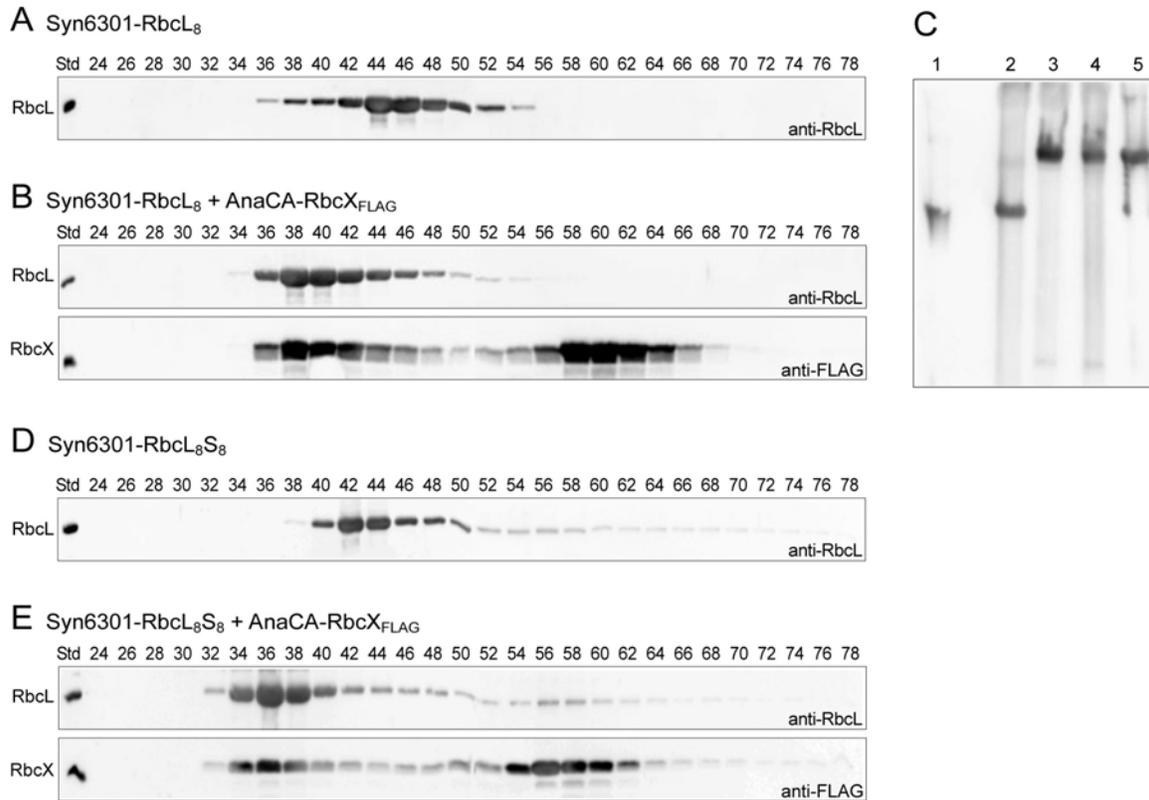
(D) Soluble *E. coli* lysates resulting from the indicated (co-)expressions were subjected directly (1-5) or after supplementation with RbcS (6) to co-immunoprecipitation with ANTI-FLAG affinity beads overnight at 4 °C. Bead-bound proteins were analyzed by SDS-PAGE, followed by Coomassie staining or immunoblotting against RbcL.

The functional significance of the dynamic RbcLX-interaction with regard to production of active holoenzyme became evident in experiments which were performed with AnaCA-RbcX instead of Syn7002-RbcX. As described above, these two cyanobacterial species of RbcX share similar size, structure and the Syn7002-RbcL binding motif (Fig. 33 A, 37, 39, 42 A-C and 43). However, although AnaCA-RbcX supports the expression of soluble RbcL, it prevents the production of RbcL<sub>8</sub> cores capable to form active RbcL<sub>8</sub>S<sub>8</sub> complexes in presence of RbcS. Instead, a high molecular weight complex between RbcL and AnaCA-RbcX is formed, which can be precipitated with immunobeads (Fig. 44 A and B, lane 11, Fig. 45, lane 5). Such a complex could also be purified upon co-expression of Syn6301-RbcL and AnaCA-RbcX<sub>N-His6</sub> in *E. coli*. The molecular weight of this complex was determined to be approximately 730 kDa, corresponding to L<sub>8</sub>(X<sub>2</sub>)<sub>8</sub> or L<sub>8</sub>(X<sub>2</sub>)<sub>9</sub> (Table 7) and suggesting that one RbcX dimer is bound per RbcL protomer. Additionally, the complex could be produced *in vitro* by incubation of purified AnaCA-RbcX with purified Syn6301-RbcL<sub>8</sub> cores (Fig. 47 C, lane 5), confirming that it is composed of merely RbcL and RbcX without involvement of GroEL or another protein. Gel filtration verified the high molecular weight of these *in vitro* complexes, as they fractionated earlier than Syn6301-RbcL<sub>8</sub> cores (~400 kDa, Fig. 47 A and B).

sample	theoretical MW (Da)	measured MW (Da)	hydrodynamic radius (nm)
Syn6301-RbcL <sub>8</sub>	419581.6	414900	5.0
Syn6301-RbcL <sub>8</sub> S <sub>8</sub>	526247.2	491600	5.2
Syn6301-RbcL/AnaCA-RbcX <sub>N-His6</sub>		730000	6.1
Syn7002-RbcL/Syn7002-RbcX <sub>N-His6</sub>		425300	4.6
AnaCA-RbcX	15506.6 (monomer) 31013.2 (dimer)	31040	1.8
Syn7002-RbcX	15269.1 (monomer) 30538.2 (dimer)	30170	1.8

**Table 7. Mass analysis of recombinant proteins.**

The theoretical MW was calculated using the ExpASY ProtParam server. The molar mass and the hydrodynamic radius of the purified proteins were analyzed by FFF-MALS, performed by Dr. Manajit Hayer-Hartl.



**Figure 47. Complex formation between Syn6301-RbcL<sub>8</sub> or -RbcL<sub>8</sub>S<sub>8</sub> and AnaCA-RbcX<sub>FLAG</sub>.** (A and D) Gel filtration chromatography (Superdex 200, 10/30, in 50 mM Tris-HCl, pH 8.5, 50 mM NaCl, 5 mM MgCl<sub>2</sub>) of 2  $\mu$ M Syn6301-RbcL<sub>8</sub> (A) or 1.7  $\mu$ M Syn6301-RbcL<sub>8</sub>S<sub>8</sub> (D). Fractions were analyzed by 17.5 % SDS-PAGE, followed by immunoblotting against RbcL. (B and E) 2  $\mu$ M Syn6301-RbcL<sub>8</sub> and 40  $\mu$ M AnaCA-RbcX<sub>FLAG</sub> (B) or 1.7  $\mu$ M Syn6301-RbcL<sub>8</sub>S<sub>8</sub> and 34  $\mu$ M AnaCA-RbcX<sub>FLAG</sub> (D) were incubated for 60 min at 4 °C and subjected to gel filtration as above, followed by SDS-PAGE and immunoblotting to detect RbcL and RbcX<sub>FLAG</sub>. (C) Native PAGE (6 %) analysis of purified Syn6301-RbcL<sub>8</sub> (1), soluble *E. coli* lysate of Syn6301-RbcL expression (2), soluble lysate of Syn6301-RbcL/AnaCA-RbcX co-expression (3), soluble lysate of Syn6301-RbcL expression after addition of purified AnaCA-RbcX (4), mixture of purified Syn6301-RbcL<sub>8</sub> and purified AnaCA-RbcX (5).

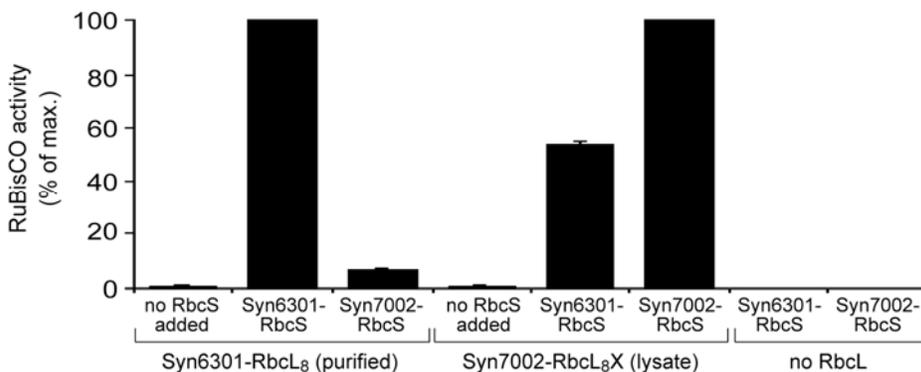
In contrast to Syn7002-RbcX, AnaCA-RbcX co-fractionated with RbcL, indicating a higher stability or a less dynamic nature of RbcL/AnaCA-RbcX-complexes compared to complexes formed with Syn7002-RbcX (Fig. 47 B compared to Fig. 41 A and B, i). This is in agreement with the high affinity of AnaCA-RbcX for the C-terminal RbcL-peptides, indicated by the low dissociation constants resulting from isothermal titration calorimetry (Fig. 43 C and D). Considering these observations, it is not surprising, that AnaCA-RbcX<sub>FLAG</sub> could not be displaced from the complex with RbcL in presence of excess non-tagged Syn7002-RbcX (Fig. 46 A, ii). Even RbcS could not disrupt this complex (Fig. 46

B, ii). Instead, complexes of AnaCA-RbcX<sub>FLAG</sub>/RbcL/RbcS were formed and could be immunoprecipitated with ANTI-FLAG affinity beads (Fig. 46 D, lanes 5 and 6). Consequently, AnaCA-RbcX was not only found to interact with readily assembled RbcL<sub>8</sub> cores (purified Syn6301-RbcL<sub>8</sub>), but also with the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme (purified Syn6301-RbcL<sub>8</sub>S<sub>8</sub>) (Fig. 46 C, ii, Fig. 47 D and E). This suggests that the affinity of AnaCA-RbcX for RbcL is strong enough to occur even in the presence of bound RbcS. It can also not be excluded that AnaCA-RbcX might interact with RbcL in a spatially or mechanistically different manner compared to Syn7002-RbcX. Considering that RbcX binds to the C-terminus of RbcL, which is involved in structural regulation of catalysis at the enzyme's active site (Duff *et al.* 2000), the tight binding of AnaCA-RbcX to the RuBisCO complex might result in a steric hinderance of the active site. As a consequence, proper binding of the substrate RuBP would be prevented, leading to the absence of carboxylation activity (Fig. 44 A, lane 11, Fig. 45, lane 5).

Only when the strong interaction between RbcL and AnaCA-RbcX is disturbed, production of active holoenzyme upon assembly with RbcS is possible. Interestingly, this could be achieved by mutations in the C-terminus of Syn7002-RbcL. When the whole RbcX binding motif was deleted in Syn7002-RbcL(aa 1-459) or when both phenylalanines were mutated in Syn7002-RbcL(F462A/F464A), RbcL became insoluble and interaction with RbcX was abolished as indicated by the absence of co-immunoprecipitation (Fig. 44 A and B, lanes 12 and 15). However, the single mutations in Syn7002-RbcL(F462A) or Syn7002-RbcL(F464A) weakened the interaction or affinity of AnaCA-RbcX for these RbcL mutants. Consequently, RbcL<sub>8</sub> complexes could form which were capable to assemble with RbcS to active holoenzyme (Fig. 44 A and B, lanes 13 and 14). As revealed by co-immunoprecipitation, Syn7002-RbcL(F464A) interacted more strongly with AnaCA-RbcX than Syn7002-RbcL(F462A), stressing the more crucial role of F462, as already described for the interaction with Syn7002-RbcX.

In summary, the high affinity of AnaCA-RbcX for heterologous RbcL results in the formation of a high molecular weight dead end complex (most likely L<sub>8</sub>(X<sub>2</sub>)<sub>8</sub>). The extremely strong contacts between RbcL and RbcX in this complex cannot be disrupted by RbcS and therefore the assembly of active RbcL<sub>8</sub>S<sub>8</sub> holoenzyme is not possible. In light of these results, the necessity of the dynamic nature of RbcL-RbcX interaction is

evident. The observed limitation of interspecies RbcX exchangeability to produce active RuBisCO enzyme also suggests the co-evolution of RbcX, RbcL and RbcS. Similarly, the co-evolution of RbcL and RbcS was reflected by limited functional compatibility between these subunits from different cyanobacterial species (Fig. 48).



**Figure 48. Interspecies compatibility of RbcS for production of active RuBisCO.**

As indicated, purified Syn6301-RbcL<sub>8</sub> cores (40 nM) and Syn7002-RbcL/Syn7002-RbcX-complexes (~40 nM, in *E. coli* lysate) were incubated alone, with Syn6301-RbcS (7 μM) or with Syn7002-RbcS<sub>FLAG</sub> (7 μM) in assay-buffer 3 for 5 min at RT, before carboxylation activity was determined. As control, absence of carboxylation activity was confirmed for RbcS alone. Maximal activity was set to 100 % and data present averages from three assays.

### 4.3.8 RbcX-mediated assembly of RbcL

So far, the necessity and/or benefit of RbcX in the assembly of cyanobacterial type I RuBisCO and the optimal dynamic character of RbcL-RbcX interaction for the productive assembly of RbcL<sub>8</sub> cores and RbcL<sub>8</sub>S<sub>8</sub> holoenzyme has been described. However, there is still need to analyze, at which stage of the assembly RbcX does interact with RbcL. In the current view of the RbcL<sub>8</sub>S<sub>8</sub> holoenzyme maturation pathway, large subunits are folded by and released from the chaperonin system in a manner dependent on Mg<sup>2+</sup> and ATP, followed by the formation of RbcL<sub>2</sub> dimers. The latter subsequently tetramerize around a fourfold axis to RbcL<sub>8</sub> cores, before eight small subunits spontaneously associate with the RbcL<sub>8</sub> core complex on top and bottom, resulting in RbcL<sub>8</sub>S<sub>8</sub> (Fig. 15) (Gatenby and Ellis 1990, Goloubinoff *et al.* 1989a, Hubbs and Roy 1993b, Roy *et al.* 1988).

Purification of LX-complexes and their immunoprecipitation from *E. coli* lysates or after *in vitro* assembly have shown that RbcX can interact with RbcL<sub>8</sub> cores (Fig. 41 A and B,

i, 46 C, 47 B and Table 7). It was also observed that RbcX interacts with RbcL subsequent to release from GroEL (Fig. 30 B), that functional RbcX is required to prevent RbcL misassembly or aggregation and to support proper RbcL<sub>8</sub> assembly (Fig. 40 and 41). The latter observations suggest that association of RbcX and RbcL is likely to occur already prior to the RbcL<sub>8</sub> assembly. In order to identify possible assembly intermediates, expression of RbcL was performed using the RTS100 HY *E. coli in vitro* transcription/translation system, which is sensitive enough to detect minute amounts of RbcL assembly intermediates.

For this approach, Syn7002-RbcL was translated in presence of externally added GroEL/GroES and RbcX<sub>FLAG</sub>. In order to synthesize radioactively labeled RbcL, the translation reaction was supplemented with <sup>35</sup>S-methionine for 6 min, followed by a chase with excess of unlabeled methionine. Samples of the reactions were withdrawn up to 160 min post chase and translation was stopped with chloramphenicol on ice. Soluble protein was analyzed by discontinuous Bis-Tris Native PAGE, allowing separation of native protein complexes according to their mass. Labeled RbcL could be detected by autoradiography. Proteins visible as prominent radioactive bands in Native PAGE, were confirmed to contain full-length RbcL by excision and re-analysis by SDS-PAGE.

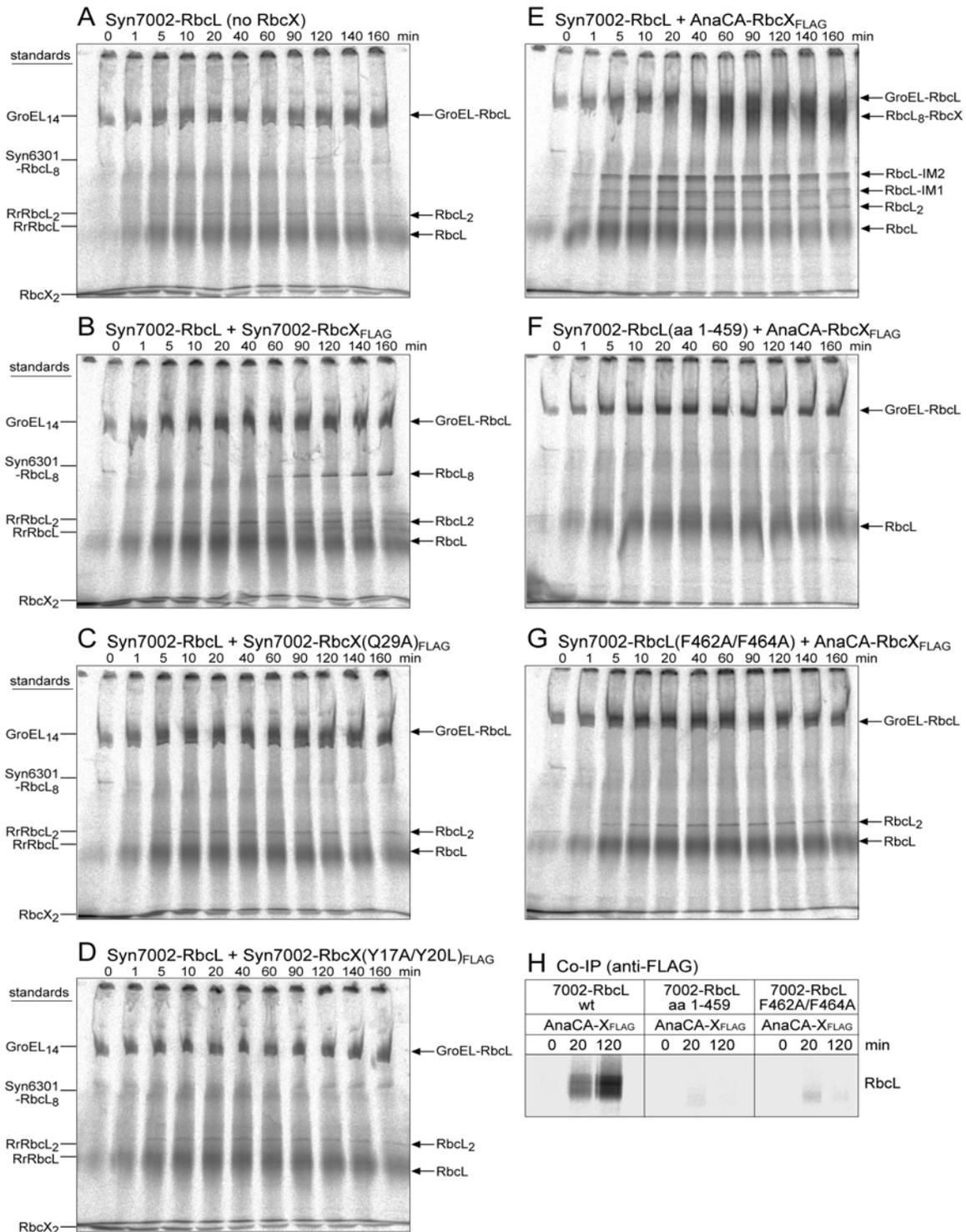
When Syn7002-RbcL was translated in absence of RbcX, soluble RbcL was mainly observed bound to GroEL (~800 kDa) at the beginning of the chase. Some minutes later, a fuzzy and a sharp band of RbcL appeared, which migrated similar to monomeric mutant Rr-RbcL (~50 kDa) or dimeric Rr-RbcL<sub>2</sub> (~100 kDa), respectively (Fig. 49 A). When Syn7002-RbcX was present in excess, intensity of the latter two bands started to decrease after 60 min, concomitant with the appearance of progressively more RbcL<sub>8</sub> core complexes (~400 kDa). Intermediary complexes of RbcL and RbcX could not be detected unambiguously, presumably due to their lability in the gel system (Fig. 49 B). The addition of the peripheral surface mutant Syn7002-RbcX(Q29A) or of the groove mutant Syn7002-RbcX(Y17A/Y20L) resulted in virtually the same pattern as observed in absence of RbcX, confirming the inability of these mutants to support RbcL<sub>8</sub> assembly (Fig. 49 C, D).

AnaCA-RbcX, which has a manifold higher affinity for RbcL than Syn7002-RbcX, resulted in the formation of stable assembly intermediates (IM1, IM2) migrating at a size

of approximately 200-350 kDa (Fig. 49 E). It is likely that they are composed of both RbcL and RbcX, because they are not seen in the assembly with Syn7002-RbcX (Fig. 49 A), reflecting the higher affinity of AnaCA-RbcX for RbcL. Appearance of these intermediates was followed by the production of a high molecular weight complex, observed as a fuzzy high molecular weight band of about 730 kDa (Fig. 44, lane 11, Fig. 45, lane 5, Fig. 47 B and Table 7), migrating just below the GroEL/RbcL complex. The assembly intermediates are also in equilibrium with the monomeric and dimeric forms of RbcL.

When the C-terminal mutants Syn7002-RbcL(aa 1-459) or Syn7002-RbcL(F462A/F464A) were translated in presence of AnaCA-RbcX, the intermediates IM1 and IM2 could not be observed and the RbcL mutant could hardly be immunoprecipitated with AnaCA-RbcX<sub>FLAG</sub> (Fig. 49 F-H). This confirms the importance of the C-terminus for RbcL-RbcX interaction to initiate RbcL assembly, as described previously (Fig. 44 and 45). Also, deletion of the C-terminus abolishes the ability of RbcL to dimerize (Fig. 49 F).

In summary, these results suggest that release of folded Syn7002-RbcL from the chaperonin-machinery is followed by the presence of RbcL as monomers or dimers. In the absence of RbcX, the monomers and dimers of RbcL decrease with time probably due to aggregation. In the presence of RbcX, they are consumed and form RbcX-bound intermediates, which might include monomers, dimers, tetramers and/or hexamers, which finally assemble to RbcX-bound RbcL<sub>8</sub> core complexes.



**Figure 49. Pulse-chase analysis of Syn7002-RbcL assembly upon *in vitro* translation.** Syn7002-RbcL (A-E wildtype, F-G mutants) was translated in the RTS100 HY *E. coli in vitro* transcription/translation system in absence of RbcX (A) or in presence of Syn7002-RbcX<sub>FLAG</sub> (B), Syn7002-RbcX<sub>FLAG</sub> mutants (C, D) or AnaCA-RbcX<sub>FLAG</sub> (E-H). Samples were withdrawn and stopped at indicated times after chase. They were analyzed by discontinuous Native PAGE (6 % / 13 %) and autoradiography (A-G) or subjected to immunoprecipitation (H).

## 5 DISCUSSION

The folding of numerous newly synthesized proteins is dependent on the assistance of molecular chaperones (Hartl and Hayer-Hartl 2002). Moreover, some oligomeric proteins have been found to require the assistance of specific chaperones for complex assembly (Ellis 2006). The mechanism and requirements for the folding and assembly of RuBisCO have been intensively studied for a long time as a consequence of the high abundance, functional importance and economic relevance of this enzyme. In this regard, special focus is put on the folding and assembly of the hexadecameric type I RuBisCO (RbcL<sub>8</sub>S<sub>8</sub>), which can be found in eukaryotic and most prokaryotic photosynthetic organisms, i.e. plants, algae and most cyanobacteria. In contrast to the dimeric type II RuBisCO, which can be folded and assembled as active enzyme both *in vitro* and *in vivo* with the help of the chaperonin system GroEL/GroES (Goloubinoff *et al.* 1989b, Larimer *et al.* 1986, Viitanen *et al.* 1990), the folding and assembly of type I RuBisCO has been largely unsuccessful so far. It has been established that the Cpn60 chaperonin machinery is essential for the production of active type I enzyme (Gatenby and Ellis 1990, Goloubinoff *et al.* 1989a, Gutteridge and Gatenby 1995, Larimer and Soper 1993) and involvement of the Hsp70 system has been suggested as well (Checa and Viale 1997). The production of active type I RuBisCO in *E. coli* is only possible, albeit inefficient, for some cyanobacterial enzymes, but not for higher plants' RuBisCO, which is prone to aggregate even upon co-expression of the chloroplast chaperonin system (Cloney *et al.* 1993, Gatenby 1984, Gatenby *et al.* 1987, Somerville *et al.* 1986). Moreover, functional type I enzyme has never been reconstituted from the unfolded state *in vitro* in presence of the chaperonin system. This represented an obstacle to any practical attempt to engineer the higher plant enzyme. It also suggests that, besides the chaperonins, further factors or special conditions are required for the productive folding and assembly of type I RuBisCO. In particular, one or several factor(s) may be necessary for the assembly of the structurally more complex type I RuBisCO. It is unclear, why RuBisCO of some cyanobacteria can assemble in *E. coli*, whereas that of higher plants cannot - despite the high degree of sequence and structure similarity (Knight *et al.* 1990, Newman *et al.* 1993). The large RuBisCO subunits (RbcL) have been suggested to require some post-translational modifications for proper folding and assembly, which might be implemented

in chloroplasts but perhaps not in *E. coli*. Indeed, several post-translational modifications of RbcL (e.g. deformylations, acetylations, hydroxylations, methylations or carbamoylations) have been reported. However, some of them affect the enzyme's activity or stability more than its assembly, others occur only in selected but not all plant or algal species, few appear only after assembly and several modifications can also be found in prokaryotes (Houtz and Portis 2003). Also, *in vitro* refolding of readily modified enzyme, purified from its native sources, has never been possible. Consequently, it is less likely that missing post-translational modifications cause the failure to produce or reconstitute functional plant RuBisCO in *E. coli* or *in vitro*. Furthermore, it has been reported that the small RuBisCO subunits (RbcS) can assemble spontaneously on preformed RbcL<sub>8</sub> cores (Andrews 1988, Lee and Tabita 1990), indicating that docking of RbcS cannot be the limiting step in production of active enzyme *in vitro*. Rather, the problem seems to occur in the ensemble of processes that result in formation of RbcL<sub>8</sub> cores. In this regard, the recent discovery of RbcX has attracted attention, since this factor was suggested to be required for the efficient formation of cyanobacterial type I RuBisCO (Li and Tabita 1997, Onizuka *et al.* 2004, Emlyn-Jones *et al.* 2006).

Consequently, the objective of the work presented here was to analyze in detail the mechanisms of folding and assembly of cyanobacterial type I RuBisCO of the species *Synechococcus* sp. PCC7002. To this end, the requirement of the chaperonin system and RbcX was investigated, including the structural and functional characterization of RbcX.

### **5.1 RbcL is a class III chaperonin substrate**

In agreement with previous reports (Goloubinoff *et al.* 1989a, Larimer and Soper 1993), the production of soluble, assembled and active cyanobacterial type I RuBisCO (of *Synechococcus* sp. PCC6301 and *Synechococcus* sp. PCC7002) in *E. coli* or in *E. coli*-derived *in vitro* transcription/translation lysate was supported by the presence of increased amounts of GroEL/GroES, whereas absence or wildtype levels of the chaperonin system resulted in mainly aggregated RbcL (Fig. 27, 29 and 30). This observation confirms the ability and necessity of GroEL to bind and fold RbcL, yielding assembly-competent subunits.

Moreover, the co- and/or post-translational interaction of type I RbcL with trigger factor (TF) and the Hsp70 system (DnaK/DnaJ/GrpE) was indicated (Fig. 27), which is consistent with previous reports where extensive aggregation of RuBisCO was observed for *E. coli dnaK* null mutants (Checa and Viale 1997). This interaction might be essential during the ribosomal translation process in order to prevent newly-synthesized RbcL chains from aggregation and to keep them stabilized in a folding-competent state throughout the period between completion of translation and beginning of chaperonin-mediated folding. The Hsp70 system appeared to be beneficial to secure increased levels of soluble RbcL, however, it did not accomplish folding and assembly of RbcL on its own – neither for type I RuBisCO (Fig. 29) nor for the type II enzyme (Fig. 19 D). For the productive folding and assembly processes the chaperonin system (GroEL/GroES) was required.

These observations classify type I RbcL as a class III substrate of the chaperonin system. *Per definitionem*, class III substrates are absolutely dependent on the assistance of chaperonin for folding to their native state. As unfolded proteins, they are highly aggregation prone and are unable to fold spontaneously. The Hsp70 system is able to bind unfolded class III substrate proteins effectively, thereby suppressing their aggregation until they are transferred to and folded by the chaperonin system (Kerner *et al.* 2005).

## 5.2 Unproductive reconstitution of type I RuBisCO *in vitro*

Despite the possibility to support the production of active cyanobacterial type I RuBisCO in *E. coli* by increased levels of GroEL/GroES, all *in vitro* attempts to reconstitute this enzyme from *Synechococcus* sp. PCC6301 (Syn6301-RbcL<sub>8S8</sub>) in presence of chaperonins and under conditions that successfully reconstitute dimeric type II RuBisCO from *Rhodospirillum rubrum* (Rr-RbcL<sub>2</sub>) failed. Numerous experimental approaches were made to optimize the folding conditions. For example, the chloroplast homologs of GroEL (AtCpn60 $\alpha\beta$  and AtCpn60 $\beta$ ) and a variety of bacterial, mitochondrial, chloroplast or viral co-chaperones (GroES, mt-Hsp10, Gp31, AtCpn20, SoCpn20, PsCpn20 and AtCpn10) were applied in different concentrations and combinations. The Gp31 and AtCpn10 co-factors were used to allow folding of larger substrate proteins, since they

both lack roof structures, resulting in an enlarged chaperonin cage compared to the cage formed by GroEL/GroES (Hunt *et al.* 1997, Koumoto *et al.* 2001). The activity of all applied chaperonins as well as their functional interaction with the co-chaperones was first confirmed by examination of the chaperonin ATPase-activity or by *in vitro* refolding of the stringent chaperonin substrates MDH and Rr-RbcL<sub>2</sub>.

Moreover, refolding attempts for type I RuBisCO in presence of DnaK/DnaJ/GrpE or RbcX, performed under different conditions, also failed to yield active enzyme. The effect of different denaturants (guanidinium, urea or acid), temperatures, buffers (e.g. Tris, MOPS, ethanolamine, RTS-buffer), pH-values (adjusted e.g. with NaOH, KOH, HCl, Acetic acid), salts (e.g. NaCl, KCl, NaOAc, KOAc MgCl<sub>2</sub>, Mg(OAc)<sub>2</sub>, ammoniumsulfate), the presence of soluble *E. coli* lysate, BSA, casein, chaotropic agent arginine, RuBisCO substrate RuBP, crowding reagents (e.g. dextran, ficoll, PEGs, TMAO), redox reagents or redox pairs (DTT,  $\beta$ -mercaptoethanol, GSH/GSSG) in different concentrations and combinations were thoroughly tested as well.

Consequently, there must be one or more steps in the folding and assembly pathway of type I RuBisCO, which work effectively in the natural host or in the cytosol of *E. coli*, respectively, but not under the *in vitro* conditions tested. Notably, this problem is not existent for the dimeric type II enzyme, which can be produced by GroEL/GroES both in *E. coli* and *in vitro* (Fig. 19).

In order to specify the obstacle for *in vitro* reconstitution of type I RuBisCO, the refolding and assembly process was analyzed stepwise in more detail. It could be shown that GroEL effectively binds unfolded Syn6301-RbcL and thus prevents its aggregation upon dilution from denaturant under non-permissive conditions (Fig. 21 B and C). Next to aggregation prevention, another essential feature of chaperonin-mediated substrate folding is encapsulation of the substrate inside the GroEL/GroES cage. The resulting confinement can accelerate folding of some proteins by smoothing the folding energy landscape and by destabilization of intermediates resulting in their flux to the native state (Brinker *et al.* 2001, Tang *et al.* 2006). The size of RbcL (50-55 kDa) does not exceed the capacity of the GroEL cavity, which can accommodate proteins up to 60 kDa (Hartl and Hayer-Hartl 2002). Indeed, Syn6301-RbcL (~52 kDa) was found to be encapsulated in the GroEL/GroES cage. Additionally, ATP-hydrolysis by GroEL induced cycling of

GroES which is required for coordinated substrate folding and release (Fig. 22 B, Hartl 1996).

However, it appeared that the efficient release of Syn6301-RbcL from GroEL is severely impaired, since throughout the whole refolding reaction Syn6301-RbcL was found associated with GroEL (Fig. 25 B). On the other hand, type II Rr-RbcL was efficiently folded by GroEL/GroES and spontaneously assembled to give the active dimeric enzyme (Fig. 25 A). Since cycles of binding and release of Syn6301-RbcL could be verified by observing the transfer of GroEL-bound RbcL to the non-cycling single ring variant of GroEL, GroEL-SR1 (Fig. 26), it is possible that incompletely folded Syn6301-RbcL is released from GroEL and becomes immediately recaptured for further folding steps. However, it appears that it is never released from GroEL as an assembly-competent soluble protein. Rather, it seems to re-bind the chaperonin before it can assemble. Perhaps the equilibrium of this reaction is located on the side of GroEL-bound RbcL and requires a specific (assembly) factor or condition in order to shift towards assembly competence, i.e., properly folded RbcL (upon release from the chaperonin) enters the RbcL<sub>8</sub> assembly pathway thereby preventing re-binding to GroEL. Maybe such factor or condition would support the successful *in vitro* reconstitution of type I RuBisCO.

This possibility is reasonable in light of the fact that the *in vitro* system used for the above described studies, is still profoundly different to those systems that allow production of active type I enzyme, such as *E. coli* (Goloubinoff *et al.* 1989a), *E. coli* lysate (RTS system, Fig. 30), chloroplasts (Roy *et al.* 1982) or chloroplast extracts (Hubbs and Roy 1992, 1993a, b). All these latter systems contain numerous additional factors (e.g. further chaperones or enzymes) that might be essential for specific interactions with type I RbcL either prior or after its interaction with GroEL. Moreover, they provide unique reaction conditions which might be advantageous for the production of active enzyme. Additionally, in all these productive systems folding and assembly of RuBisCO is directly following transcription and translation. In contrast, *in vitro* refolding follows artificial unfolding of the RbcL with denaturants. Possibly and for unknown reasons, the denaturant-unfolded Syn6301-RbcL misfolds and is hence not competent for assembly compared to an elongating RbcL polypeptide chain that emerges from the ribosome and already begins to interact with chaperones before completion of translation.

The identification of the critical factor(s) or condition(s) necessary to accomplish *in vitro* reconstitution of type I RuBisCO was not further investigated in this study.

### 5.3 The RuBisCO assembly chaperone RbcX

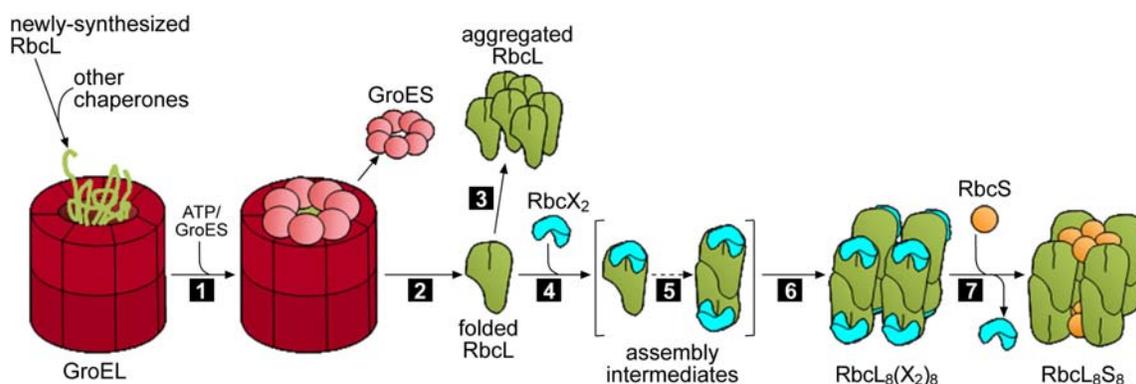
Interestingly, several cyanobacterial species encode *rbcX* in the intergenic space between the genes of the large (*rbcL*) and the small (*rbcS*) RuBisCO subunits. Assuming that the resulting *rbcLXS* operon and the concomitant co-expression of RbcX and the RuBisCO subunits are not fortuitous, a functional relation or interaction between RbcX and RbcL is an obvious presumption. Indeed, initial studies have suggested a potential role of RbcX in the assembly of some cyanobacterial hexadecameric type I RuBisCO, however, without providing any insight into its mechanistic function (Larimer and Soper 1993, Li and Tabita 1994, 1997, Onizuka *et al.* 2002, Onizuka *et al.* 2004, Emlyn-Jones *et al.* 2006). In the work presented here, the role of RbcX as a specific assembly chaperone for cyanobacterial RuBisCO was confirmed by structural and functional analyses.

#### 5.3.1 RbcX-mediated RuBisCO assembly

The current knowledge about type I RuBisCO assembly (Goloubinoff *et al.* 1989a, Gatenby and Ellis 1990, Gutteridge and Gatenby 1995) and the here presented characteristics of RbcX and its interaction with RbcL allow to propose a possible pathway for RbcX-dependent RuBisCO assembly e.g. in the cyanobacterium *Synechococcus* sp. PCC7002 (Fig. 50).

As described above, newly synthesized RbcL subunits might be stabilized by co- and/or post-translational interaction with the ribosome associated trigger factor and/or with the Hsp70 system. Folding of RbcL is accomplished by the chaperonin system GroEL/GroES in an ATP-dependent manner. Upon release from the GroEL cavity, folded RbcL is prone to misassemble or aggregate. Therefore, monomers or perhaps spontaneously formed dimers of RbcL associate immediately with RbcX, resulting in assembly intermediates competent for efficient assembly progression and protected from uncontrolled misassembly. As described below (chapter 5.3.2), the large subunits interact with RbcX *via* their C-terminus, which binds into the central hydrophobic groove of RbcX. Based on this initial attachment, polar region(s) of RbcL probably additionally interact with the

corresponding polar peripheral surfaces of RbcX. These polar contacts may be important for the proper positioning between RbcL subunits for assembly. The absolute composition of the assembly intermediates is not yet defined, but they might include monomers, dimers, tetramers and/or hexamers of RbcL in association with RbcX (Fig. 49 E). The final assembly product is the RbcL<sub>8</sub> core complex which has most likely one RbcX dimer bound to every RbcL subunit (L<sub>8</sub>(X<sub>2</sub>)<sub>8</sub>). The dynamic nature of the interaction between RbcL and RbcX allows the final assembly of eight small RuBisCO subunits on the RbcL<sub>8</sub> core, thereby enforcing the release of RbcX and formation of the stable and active holoenzyme RbcL<sub>8</sub>S<sub>8</sub>.



**Figure 50. Model of folding and RbcX-mediated assembly of cyanobacterial RuBisCO.**

Newly-synthesized RbcL is bound to GroEL (possibly after interaction with other chaperones) and becomes folded in the GroEL/GroES-cage upon ATP-hydrolysis (1). After folding and release from GroEL (2), RbcL is prone to aggregate in absence of assembly-chaperones such as RbcX (3). RbcX dimers can recognize and bind to the exposed C-terminal RbcL peptide (4). Monomers or spontaneously formed dimers of RbcL are most likely the initial association partners for RbcX, but the constitution of the assembly intermediates is still speculative (5). Assembly of RbcX-associated RbcL subunits results in a complex which is most likely RbcL<sub>8</sub>(X<sub>2</sub>)<sub>8</sub> (6). The small RbcS subunits assemble on top and bottom of the RbcL<sub>8</sub> core, thereby leading to the release of RbcX and formation of the active holoenzyme RbcL<sub>8</sub>S<sub>8</sub>.

The observed sequential operation mode of GroEL/GroES and RbcX seems to be optimized in the way that fast and efficient RbcX-mediated assembly of RbcL might elevate the turnover and release rate of GroEL/GroES by shifting the equilibrium to the assembly intermediates and products, thereby increasing the efficiency of the whole system to supply maximal amounts of assembled and active enzyme. In this way RbcX fulfills a function which could be beneficial regarding the problem of inefficient GroEL-substrate release during *in vitro* reconstitution of type I RuBisCO, as described above (chapter 5.2).

In contrast to other chaperones, such as GroEL or Hsp70, RbcX recognizes substrate specific structural features and it promotes the association of several folded protein subunits into a biologically functional oligomer (RuBisCO) rather than promoting protein folding directly. In this regard, RbcX can be defined as a specific assembly chaperone and is thus functionally reminiscent of a few other chaperones known to be required for the processes of nucleosome formation (e.g. nucleoplasmin) or proteasome assembly (e.g. proteasome maturation protein POMP/Ump1 or proteasome assembly factor PAC) (Akey and Luger 2003, Ellis 2006, Hirano *et al.* 2005, Witt *et al.* 2000). Yet the group of identified assembly chaperones is relatively small.

Notably, assembly of RuBisCO from *Synechococcus* sp. PCC7002, a species in which *rbcX* is encoded with both RuBisCO subunits in the *rbcLXS* operon, appears to be critically dependent on the presence of RbcX (Fig. 29, 30 and 49). In contrast, the assembly of RuBisCO from *Synechococcus* sp. PCC6301, which encodes *rbcX* outside the *rbcLS* operon, RbcX plays a supportive but not an essential role (Fig. 31). Similar observations of RbcX-independent assembly have been reported for *Synechococcus* sp. PCC7942, whose *rbcX* gene is also separated from *rbcLS* (Emlyn-Jones *et al.* 2006). If the localization of *rbcX* inside or outside the *rbcLS* operon is an evolutionary consequence of RbcX-dependent or -independent RuBisCO assembly, or if RuBisCO assembly in any photosynthetic organism can benefit of *rbcX* irrespective of its genomic localization, is speculative. It is also possible that divergent evolution of the RbcL sequence relaxed the requirement of RbcX function or enforced the need for further unknown assembly factors. In this regard it should be mentioned that highest yields of soluble and active enzyme were generally attained when elevated levels of both RbcX and GroEL/GroES were provided for RuBisCO expression in *E. coli* or in the *in vitro* translation system (Fig. 29 and 30). However, in cyanobacteria the amount of RbcX protein appears to be below detection limits and only little amounts of *rbcX* transcripts could be detected (Emlyn-Jones *et al.* 2006, Onizuka *et al.* 2004). Perhaps, the RbcX levels in plant chloroplasts are similarly low. Evidently, this is contrary to the high abundance of RuBisCO in these organisms or organelles. How this discrepancy is overcome or if further assembly factors next to RbcX are produced, is unclear.

### 5.3.2 The structural foundation for RbcX-mediated RuBisCO assembly

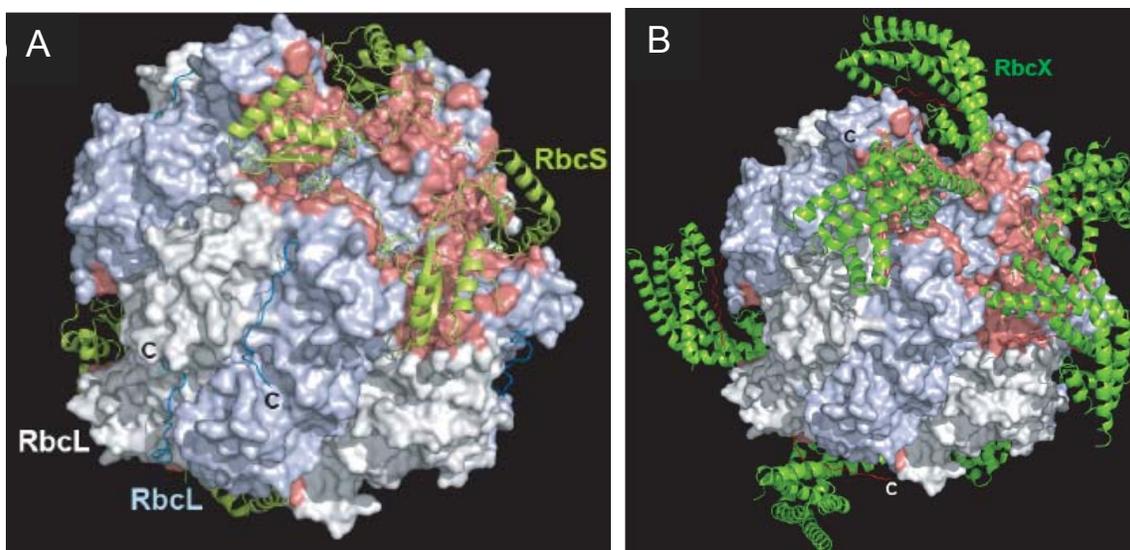
RbcX is an all  $\alpha$ -helical protein and functional as a dimer of approximately 30 kDa whose overall architecture seems to be conserved among cyanobacteria (Fig. 33, 36, 37 and 39). The dimer forms an arc-shaped complex with a central diagonal groove which could be shown to be capable of binding the C-terminus of RbcL subunits. This interaction requires the recognition motif EIKFEFDxxD in the RbcL C-terminus (residues 459 to 468 in Syn7002-RbcL). This peptide binds to RbcX in an extended conformation with the side chains F462 and F464 extending into conserved hydrophobic cavities of the groove, supported by hydrogen bonds and van-der-Waals contacts (Fig. 42 and 43).

Mutation of critical residues in the binding groove of RbcX resulted in disruption or attenuation of the RbcL-RbcX interaction and in misassembly or aggregation of RbcL (Fig. 40, 41 iii and 49 D). The same effect had C-terminal deletion or mutation of RbcL subunits, irrespective if they represented RbcX-dependent (Syn7002-RbcL) or RbcX-independent (Syn6301-RbcL) assembly types (Fig. 44, 45 and 49 F-H). This indicates on the one hand, that functional integrity of the RbcX binding groove is a prerequisite for the interaction with RbcL in order to keep it soluble and assembly-competent. On the other hand, integrity of the RbcL C-terminus with the recognition motif (particularly the presence of the two included phenylalanines) is not only critical for the initial recognition by and subsequent interaction with RbcX, but is itself required for RbcL assembly-competence and/or stabilization of assembly intermediates or RbcL<sub>8</sub> cores.

The C-terminal RbcL peptide is located on the outer surface of assembled RuBisCO and contributes to the structural regulation of its catalysis. Binding of ligands (substrates or inhibitors) to the empty active site of RuBisCO triggers the closure of the active site and conformational changes in the C-terminal strand. The latter then stabilizes the closed position of the active site's lid and is itself fastened in the closed state by a latch which involves D468 of the RbcX recognition motif (Duff *et al.* 2000). During RbcL assembly the C-terminal strand might be more detached from the RbcL subunits and the hydrophobic side chains might be more solvent exposed. Binding of RbcX could therefore protect and stabilize the C-terminus and thus the whole protein during assembly, thereby preventing uncontrolled misassembly or aggregation.

In addition to the central groove, a second region on RbcX with functional relevance for productive interaction with RbcL is formed by the polar surface areas around edges of the RbcX dimer. Mutation of critical conserved residues in these regions resulted in formation of soluble but misassembled RbcL complexes of high molecular weight (Fig. 40 and 41 ii). This implies that the peripheral corner regions of RbcX are essential to coordinate the proper positioning and suitable contact formation between the RbcL subunits of assembly intermediates, thereby allowing or enforcing the formation of RbcL<sub>8</sub> cores which are convertible into active RbcL<sub>8</sub>S<sub>8</sub> holoenzymes.

The combination of both, the central binding cleft and the polar peripheral regions of RbcX, appears to enable RbcX to accomplish the proper assembly of RbcL subunits. Anchoring of the RbcL C-terminus in the binding cleft is required to keep RbcL in a soluble state and to secure RbcL in an orientation relative to RbcX which is optimal for interaction with the peripheral RbcX binding surfaces. On this structural basis, optimal positioning and interconnection of the RbcL subunits is possible. In the resulting RbcL<sub>8</sub> complex, RbcX might be bridging adjacent RbcL dimers and masking parts of the surface which is covered by the globular domain of RbcS in the final holoenzyme complex (Fig. 51). Such overlapping binding areas of RbcX and RbcS could explain why Syn7002-RbcX cannot re-bind to RbcL after its displacement by RbcS upon RbcL<sub>8</sub>S<sub>8</sub> holoenzyme assembly (Fig. 46 B, i and C, iii). Although RbcS in RbcL<sub>8</sub>S<sub>8</sub> does not cover the C-terminal RbcL peptide, it is also possible, that binding of RbcS induces conformational conversions in RbcL<sub>8</sub>. These could stabilize the RbcL C-terminus so as to make the critical phenylalanine inaccessible to RbcX. The latter scenario could also prevent re-binding of RbcX either by itself or in concert with the steric overlap of recognition and binding regions for RbcS and RbcX.



**Figure 51. Crystal structure of Syn6301-RbcL<sub>8</sub>S<sub>8</sub> and model of the RbcL<sub>8</sub>(X<sub>2</sub>)<sub>8</sub>-complex.**

(A) RuBisCO of *Synechococcus* sp. PCC6301 bound to CABP (Newman *et al.* 1993). The 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 RbcX binding to the assembled RbcL<sub>8</sub> core complex using the same representation as in (A). The model is based on the structures of the Syn6301-RbcL<sub>8</sub>S<sub>8</sub> 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.

### 5.3.3 The benefit of dynamic RbcL-RbcX interaction

An intriguing aspect in the RbcX-mediated assembly of type I RuBisCO is the perfectly concerted sequence as well as the specific character of RbcL interaction with RbcX and RbcS, which allows the production of active RbcL<sub>8</sub>S<sub>8</sub> holoenzyme.

Whereas the affinity of Syn7002-RbcX to the C-terminal RbcL peptide was found to be relatively low, its affinity for folded and assembled RbcL was considerably higher (Fig. 43 A and B, Fig. 46 A, i). This suggests that the overall binding strength of RbcX is provided by both the central binding cleft (binding the C-terminal RbcL strand) and, to a substantial degree, by the peripheral surfaces (their binding sites on RbcL are yet speculative and remain to be determined). Importantly, the binding strength between RbcL and RbcX is weak enough to constitute a dynamic manner of interaction (Fig. 46 A, i). This causes not only exchangeability of RbcX, but it is also a prerequisite for the

efficient displacement of RbcX by RbcS in the final step of RbcL<sub>8</sub>S<sub>8</sub> holoenzyme formation (Fig. 46 B, i). Release of RbcX and the above discussed prevention of its re-binding to RbcL<sub>8</sub>S<sub>8</sub> allow the correct and stable positioning of RbcS on the RbcL<sub>8</sub> cores. The small subunits are not only essential for catalytic RuBisCO activity, but they are also required for maintaining the structural integrity of RbcL<sub>8</sub>S<sub>8</sub> by forming the dominant contacts between the four RbcL<sub>2</sub> dimers in the holoenzyme (Andrews and Ballment 1983, Curmi *et al.* 1992). In this regard, RbcS takes over the function of RbcX in stabilizing and positioning the RbcL subunits.

The functional significance of the dynamic nature in RbcL-RbcX interaction became evident when complexes between Syn7002-RbcL (from *Synechococcus* sp. PCC7002) were formed with heterologous RbcX from another cyanobacterial species, *Anabaena* sp. CA. Compared to Syn7002-RbcX, AnaCA-RbcX had a significantly higher affinity for Syn7002-RbcL or its C-terminal peptide (Fig. 43 C and D, Fig. 46 A, ii). As a consequence, stable RbcL/RbcX-complexes of high molecular weight (most likely RbcL<sub>8</sub>(X<sub>2</sub>)<sub>8</sub>) were formed which were not competent for assembly with RbcS to give the active RbcL<sub>8</sub>S<sub>8</sub> holoenzyme (Fig. 45 and 47 B, C). The binding strength of AnaCA-RbcX to RbcL was so high that RbcS was not able to displace the RbcX (Fig. 46 B, ii and Fig. 46 D, lane 6). Furthermore, it was observed that AnaCA-RbcX also bound to assembled RbcL<sub>8</sub>S<sub>8</sub> holoenzyme to give a complex of unknown composition containing RbcL, RbcX and RbcS (Fig. 46 C, iv and 47 E). The activity deficiency of the AnaCA-RbcX/RbcL/RbcS complexes suggests that the tight binding of AnaCA-RbcX to the RbcL C-terminus, which is involved in the structural regulation of catalysis (Duff *et al.* 2000), probably blocks the enzyme's active site and prevents proper substrate-binding and/or -conversion. The latter observations confirmed not only the biological relevance of dynamic RbcL-RbcX interaction for productive RuBisCO assembly, but also demonstrated the co-evolution of RbcX with RbcLS and the resulting limitations of functional compatibility between proteins from two evolutionary related but distinct species. However, in some instances exchangeability of RbcX between different species is possible, notably for species whose *rbcX* is encoded outside the *rbcLS* operon and whose RuBisCO assembly is supported by RbcX but not dependent on it (e.g. Syn7002-

RbcX supports assembly of Syn6301-RbcL and of Syn7942-RbcL) (Fig. 45, Emlyn-Jones *et al.* 2006).

### 5.3.4 Implications

The structural complexity of the hexadecameric type I RuBisCO might necessitate the assistance of a specific assembly chaperone compared to the dimeric type II RuBisCO which can assemble spontaneously. Indeed, homologs of the cyanobacterial assembly chaperone RbcX have been found in the genomes of numerous evolutionary distinct species with type I RuBisCO, such as higher plants (e.g. *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*), green algae (e.g. *Chlamydomonas reinhardtii*) or moss (e.g. *Physcomitrella patens*) (Fig. A1). Additionally, the RbcX recognition motif in the C-terminus of RbcL is conserved for type I but not for type II RuBisCO (Fig. 42 C and D). But this is not the only fact suggesting the plausibility that assembly of plant RuBisCO is a process assisted by chaperones, such as functional RbcX homologs. It is also characteristic for plants that their RbcL subunits are synthesized and assembled in the chloroplast stroma, whereas the nuclear encoded small subunits have to be imported into the chloroplasts and to be processed prior to association with the RbcL<sub>8</sub> cores (Gutteridge and Gatenby 1995). Hence, the availability of a factor like RbcX, which supports both assembly and stability RbcL<sub>8</sub>, would be beneficial, especially when production and import of sufficient amounts of RbcS is limited, e.g. due to stress situations which might temporally favor the synthesis of proteins different from RbcS.

Regarding the search for genetically manipulated plant RuBisCO with improved catalytic properties, which could be advantageous with respect to agricultural productivity, elucidation of obligate factors and conditions for efficient production of plant RuBisCO in heterologous *in vivo* or *in vitro* expression and screening systems would be extremely beneficial. Therefore, the possible requirement of assembly chaperones for proper assembly of this enzyme should be taken into account in future and is worth further investigation.

## 6 REFERENCES

- Agashe, V.R., Guha, S., Chang, H.C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., 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., Hartl, F.U. (2000). Roles of molecular chaperones in cytoplasmic protein folding. *Semin. Cell Dev. Biol.* *11*, 15-25.
- Akey, C.W., Luger, K. (2003). Histone chaperones and nucleosome assembly. *Curr. Opin. Struct. Biol.* *13*, 6-14.
- Andersson, I., Taylor, T.C. (2003). Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* *414*, 130-140. Review.
- Andrews, T.J., Whitney, S.M. (2003). Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants. *Arch. Biochem. Biophys.* *414*, 159-169. Review.
- 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., 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.
- Andrews, T.J., Ballment, B. (1983). The function of the small subunits of ribulose bisphosphate carboxylase-oxygenase. *J. Biol. Chem.* *258*, 7514-7518.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* *181*, 223-230.
- Apuya, N.R., Yadegari, R., Fischer, R.L., Harada, J.J., Zimmerman, J.L., Goldberg, R.B. (2001). The *Arabidopsis* embryo mutant *schlepperless* has a defect in the *chaperonin-60 $\alpha$*  gene. *Plant Physiol.* *126*, 717-730.
- 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*. John Wiley & Sons, NY.
- Baldwin, R.L., Rose, G.D. (1999). Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem. Sci.* *24*, 77-83.

- Baldwin, R.L. (1995). The nature of protein folding pathways: the classical *versus* the new view. *J. Biomol. NMR* 5, 103-109.
- Baneyx, F., Bertsch, U., Kalbach, C.E., van der Vies, S.M., Soll, J., 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.
- Barraclough, R., Ellis, R.J. (1980). Protein synthesis in chloroplasts. IX. Assembly of newly-synthesized large subunits into ribulose biphosphate carboxylase in isolated pea chloroplasts. *Biochim. Biophys. Acta* 608, 19-31.
- Ben-Zvi, A.P., Goloubinoff, P. (2001). Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *J. Struct. Biol.* 135, 84-93. Review.
- Bertsch, U., Soll, J. (1995). Functional analysis of isolated Cpn10 domains and conserved amino acid residues in spinach chloroplast co-chaperonin by site-directed mutagenesis. *Plant Mol. Biol.* 29, 1039-1055.
- Bertsch, U., Soll, J., Seetharam, R., 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. USA.* 89, 8696-8700.
- Berry, J.O., Nikolau, B.J., Carr, J.P., Klessig, D.F. (1986). Translational regulation of light-induced ribulose 1,5-bisphosphate carboxylase gene expression in amaranth. *Mol. Cell Biol.* 6, 2347-2353.
- Berry, J.O., Nikolau, B.J., Carr, J.P., Klessig, D.F. (1985). Transcriptional and post-transcriptional regulation of ribulose 1,5-bisphosphate carboxylase gene expression in light- and dark-grown amaranth cotyledons. *Mol. Cell Biol.* 5, 2238-2246.
- Boisvert, D.C., Wang, J., Otwinowski, Z, Horwich, A.L., Sigler, P.B. (1996). The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATP  $\gamma$ -S. *Nat. Struct. Biol.* 3, 116-121.
- Boston, R.S., Viitanen, P.V., Vierling, E. (1996). Molecular chaperones and protein folding in plants. *Plant Mol. Biol.* 32, 191-222. Review.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.
- Brahamsha B. (1996). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl. Environ. Microbiol.* 62, 1747-1751.

- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., Sigler, P.B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371, 578-586.
- Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* 107, 223-233.
- Bukau, B., Weissman, J., Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell* 125, 443-451.
- Bukau, B., Horwich, A.L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* 92, 351-366. Review.
- Cannon, S., Wang, P., Roy, H. (1986). Inhibition of ribulose biphosphate carboxylase assembly by antibody to a binding protein. *J. Cell Biol* 103, 1327-1335.
- Chatellier, J., Buckle, A.M., Fersht, A.R. (1999). GroEL recognises sequential and non-sequential linear structural motifs compatible with extended beta-strands and alpha-helices. *J. Mol. Biol* 292, 163-172.
- Checa, S.K., Viale, A.M. (1997). The 70-kDa heat-shock protein/DnaK chaperone system is required for the productive folding of ribulose-biphosphate carboxylase subunits in *Escherichia coli*. *Eur. J. Biochem.* 248, 848-855.
- Chen, L., Sigler, P.B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell* 99, 757-68.
- Chen, S., Roseman, A.M., Hunter, A.S., Wood, S.P., Burston, S.G., Ranson, N.A., Clarke, A.R., Saibil, H.R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* 371, 261-264.
- Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G., Dobson, C.M. (1999). Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. USA* 96, 3590-3594.
- Chory, J., Muller, E.D., Kaplan, S. (1985). DNA-directed *in vitro* synthesis and assembly of the form II D-ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospseudomonas sphaeroides*. *J. Bacteriol.* 161, 307-313.
- Chung, C.T., Niemela, S.L., Miller, R.H. (1990). One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. USA* 86, 2172-2175.

- Cloney, L.P., Bekkaoui, D.R., Feist, G.L., Lane, W.S., Hemmingsen, S.M. (1994). *Brassica napus* plastid and mitochondrial chaperonin-60 proteins contain multiple distinct polypeptides. *Plant Physiol.* *105*, 233-41.
- Cloney, L.P., Bekkaoui, D. R., 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., Wu, H.B., Hemmingsen, S.M. (1992a). Expression of plant chaperonin-60 genes in *Escherichia coli*. *J. Biol. Chem.* *267*, 23327-23332.
- Cloney, L.P., Bekkaoui, D. R., Wood, M.G., Hemmingsen, S.M. (1992b). Assessment of plant chaperonin-60 gene function in *Escherichia coli*. *J. Biol. Chem.* *267*, 23333-23336.
- 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. *Nucl. Acids Res.* *16*, 10881-10890.
- Curmi, P.M., Cascio, D., Sweet, R.M., Eisenberg, D., Schreuder, H. (1992). Crystal structure of the unactivated form of ribulose-1,5-bisphosphate carboxylase/oxygenase from tobacco refined at 2.0-Å resolution. *J. Biol. Chem.* *267*, 16980-16989.
- Dean, C., Pichersky, E., Dunsmuir, P. (1989). Structure, evolution and regulation of *RbcS* genes in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* *40*, 415-439. Review.
- de la Fortelle, E., Bricogne, G. (1997). Maximum-likelihood heavy atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* *276*, 472-494.
- Deuerling, E., Bukau, B. (2004). Chaperone-assisted folding of newly synthesized proteins in the cytosol. *Crit. Rev. Biochem. Mol. Biol.* *39*, 261-277. Review.
- Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., 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., Viitanen, P.V. (2000). Reconstitution of higher plant chloroplast chaperonin 60 tetradecamers active in protein folding. *J. Biol. Chem.* *275*, 11829-11835.
- Dill, K.A., Chan, H.S. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.* *4*, 10-19.
- Dobson, C.M. (2003). Protein folding and misfolding. *Nature* *426*, 884-890. Review.

- Dobson, C.M. (1995). Finding the right fold. *Nat. Struct. Biol.* 2, 513-517.
- Duff, A.P., Andrews, T.J., Curmi, P.M.G. (2000). The transition between the open and closed states of rubisco is triggered by the inter-phosphate distance of the bound bisphosphate. *J. Mol. Biol.* 298, 903-916.
- Ellis, J.R. (2006). Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem. Sci* 31, 395-401.
- Ellis, J.R. (2001). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* 11, 114-119.
- Ellis, R.J., Hartl, R.U. (1999). Principles of protein folding in the cellular environment. *Curr. Opin. Struct. Biol.* 9, 102–110. Review.
- Ellis, J.R. (1996). Discovery of molecular chaperones. *Cell Stress & Chaperones* 1, 155-160. Review.
- Ellis, R.J. (1990). Molecular chaperones: the plant connection. *Science.* 250, 954-959. Review.
- Emlyn-Jones, D., Woodger, F.J., Price, G.D., Whitney, S.M. (2006). RbcX can function as a rubisco chaperonin, but is non-essential in *Synechococcus* PCC7942. *Plant Cell Physiol.* 47, 1630-1640.
- Emsley, P., Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect D-Biol. Crystallogr.* 60, 2126-2132.
- English, R.S., Williams, C.A., Lorbach, S.C., Shively, J.M. (1992). Two forms of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Thiobacillus denitrificans*. *FEMS Microbiol. Lett.* 94, 111–119.
- Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph Model* 15, 132-134, 112-113.
- Evans, P.R. (1997). Scala. *Joint CCP4 and ESF-EACBM Newsletter* 33, 22-24.
- Ewalt, K.L., Hendrick, J.P., Houry, W.A., Hartl, F.U. (1997). *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell* 90, 491-500.
- Fenton, W.A., Kashi, Y., Furtak, K., Horwich, A.L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* 371, 614-61.
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431, 590-596.

- Fitchen, J.H., Knight, S., Andersson, I., Branden, C.-I., 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. USA* *87*, 5768-5772.
- Frydman, J., Nimmesgern, E., Ohtsuka, K., Hartl, F.U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* *370*, 111-117.
- Gatenby, A.A., Ellis, R.J. (1990). Chaperone function: the assembly of ribulose bisphosphate carboxylase- oxygenase. *Annu. Rev. Cell Biol.* *6*, 125-149. Review.
- Gatenby, A.A., van der Vies, S.M., Rothstein, S.J. (1987). Coexpression of both the maize large and wheat small subunit genes of ribulosebisphosphate carboxylase in *Escherichia coli*. *Eur. J. Biochem.* *168*, 227-231.
- Gatenby, A.A., van der Vies, S.M., Bradley, D. (1985). Assembly in *E. coli* of a functional multi-subunit ribulose bisphosphate carboxylase from a blue-green alga. *Nature* *314*, 617-620.
- Gatenby, A.A. (1984). The properties of the large subunit of maize ribulose bisphosphate carboxylase/oxygenase synthesized in *Escherichia coli*. *Eur. J. Biochem.* *144*, 361-366
- Glover, J.R., Lindquist, S. (1998). Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* *94*, 73-82.
- Goloubinoff, P., Gatenby, A.A., Lorimer, G.H. (1989a). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* *337*, 44-47.
- Goloubinoff, P., Christeller, J.T., Gatenby, A.A., Lorimer, G.H. (1989b). Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature* *342*, 884-889.
- Gouet, P., Courcelle, E., Stuart, D.I., Metz, F. (1999). ESPript: multiple sequence alignments in PostScript. *Bioinformatics* *15*, 305-308.
- Gurevitz, M., Somerville, C.R., McIntosh, L. (1985). Pathway of assembly of ribulosebisphosphate carboxylase/oxygenase from *Anabaena* 7120 expressed in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* *82*, 6546-6550.
- Gutteridge, S., Gatenby, A.A. (1995). Rubisco synthesis, assembly, mechanism and regulation. *Plant Cell* *7*, 809-819. Review.

- Gutteridge, S. (1991). The relative catalytic specificities of the large subunit core of *Synechococcus* ribulose biphosphate carboxylase/oxygenase. *J. Biol. Chem.* *266*, 7359-7362.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* *166*, 557-590.
- Hansen, W.J., Cowan, N.J., Welch, W.J. (1999). Prefoldin-nascent chain complexes in the folding of cytoskeletal proteins. *J. Cell Biol.* *145*, 265-277.
- Hartl, F.U., Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* *295*, 1852-1858. Review.
- Hartl, FU (1996). Molecular chaperones in cellular protein folding. *Nature* *381*, 571-579. Review.
- Hartman, F.C., Harpel, M.R. (1994). Structure, function, regulation and assembly of D-ribulose-1,5 biphosphate carboxylase/oxygenase. *Ann. Rev. Biochem.* *63*, 197-234. Review.
- Hayer-Hartl, M. (2000). Assay of Malate Dehydrogenase. *Chaperonin Protocols: Methods in Molecular Biology 140: Chaperonin Protocols* (Schneider, C., ed), Humana Press, Totowa, NJ, 127-132.
- Hayer-Hartl, M.K., Martin, J., Hartl, F.U. (1995). Asymmetrical interaction of GroEL and GroES in the ATPase cycle of assisted protein folding. *Science* *269*, 836-841.
- Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W., Ellis, R.J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* *333*, 330-334.
- Hernandez, J.M., Baker, S.H., Lorbach, S.C., Shively, J.M., Tabita, F.R. (1996). Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose biphosphate carboxylase/oxygenase from the chemoautotrophic bacterium *Thiobacillus denitrificans*. *J. Bacteriol.* *178*, 347-356.
- Hesterkamp, T., Bukau, B. (1996). The *Escherichia coli* trigger factor. *FEBS Lett.* *389*, 32-34. Review.
- Hill, J.E., Hemmingsen, S.M. (2001). *Arabidopsis thaliana* type I and type II chaperonins. *Cell Stress & Chaperones* *6*, 190-200.
- Hirano, Y., Hendil, K.B., Yashiroda, H., Iemura, S., Nagane, R., Hioki, Y., Natsume, T., Tanaka, K., Murata, S. (2005). A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature* *437*, 1381-1385.

- Hirohashi, T., Nishio, K., Nakai, M. (1999). cDNA sequence and overexpression of chloroplast chaperonin 21 from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* *1429*, 512-515.
- Horovitz, A., Fridmann, Y., Kafri, G., Yifrach, O. (2001). Review: allostery in chaperonins. *J. Struct. Biol.* *135*, 104-114. Review.
- Houry, W.A., Frishman, D., Eckerskorn, C., Lottspeich, F., Hartl, F.U. (1999). Identification of *in vivo* substrates of the chaperonin GroEL. *Nature* *402*, 147-154.
- Houtz, R.L., Portis, A.R. Jr. (2003). The life of ribulose 1,5-bisphosphate carboxylase/oxygenase-posttranslational facts and mysteries. *Arch. Biochem. Biophys.* *414*, 150-158. Review.
- Hubbs, A.E., Roy, H. (1993a). Assembly of *in vitro*-synthesized large subunits into ribulose biphosphate carboxylase/oxygenase is sensitive to  $Cl^-$ , requires ATP, and does not proceed when large subunits are synthesized at temperatures  $\geq 32^\circ C$ . *Plant Physiol.* *101*, 523-533.
- Hubbs, A.E., Roy, H. (1993b). Assembly of *in vitro* synthesized large subunits into ribulose-bisphosphate carboxylase/oxygenase. Formation and discharge of an  $L_8$ -like species. *J. Biol. Chem.* *268*, 13519-13525.
- Hubbs, A.E., Roy, H. (1992). Synthesis and assembly of large subunits into ribulose biphosphate carboxylase/oxygenase in chloroplast extracts. *Plant Physiol.* *100*, 272-281.
- Hunt, J.F., van der Vies, S.M., Henry, L., Deisenhofer, J. (1997). Structural adaptations in the specialized bacteriophage T4 co-chaperonin Gp31 expand the size of the Anfinsen cage. *Cell* *90*, 361-371.
- Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L., Deisenhofer, J. (1996). The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature* *379*, 37-45.
- Ishikawa, A., Tanaka, H., Nakai, M., Asahi, T. (2003). Deletion of a *chaperonin 60  $\beta$*  gene leads to cell death in the *Arabidopsis lesion initiation 1* mutant. *Plant Cell Physiol.* *44*, 255-261.
- Jackson-Constan, D., Akita, M., Keegstra, K. (2001). Molecular chaperones involved in chloroplast protein import. *Biochim. Biophys. Acta* *1541*, 102-113. Review.
- Jones, T.A., Zou, J.Y., Cowan, S.W., Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and location of errors in these models. *Acta. Crystallogr. A* *47*, 110-119.

- Kaiser, C.M., Chang, H.C., Agashe, V.R., Lakshimpathy, S.K., Etchells, S.A., Hayer-Hartl, M., Hartl, F.U., Barral, J.M. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature* *444*, 455-460.
- Karkehabadi, S., Peddi, S.R., Anwaruzzaman, M., Taylor, T.C., Cederlund, A., Genkov, T., Andersson, I., Spreitzer, R.J. (2005). Chimeric small subunits influence catalysis without causing global conformational changes in the crystal structure of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Biochemistry* *44*, 9851-9861.
- 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., Hartl, F.U. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* *122*, 209-20.
- Kessler, F., Blobel, G. (1996). Interaction of the protein import and folding machineries of the chloroplast. *Proc. Natl. Acad. Sci. USA* *93*, 7684-7689.
- Kitano, K., Maeda, N., Fukui, T., Atomi, H., Imanaka, T., Miki, K. (2001). Crystal structure of a novel-type archaeal rubisco with pentagonal symmetry. *Structure* *9*, 473-481.
- Kleywegt, G.T., Jones, T.A. (1994). A super position. *CCP4/ESF-EACBM Newsletter on Protein Crystallogr.* *31*, 9-14.
- Klumpp, M., Baumeister, W., Essen, L.O. (1997). Structure of the substrate binding domain of the thermosome, an archaeal group II chaperonin. *Cell* *91*, 263-270.
- Knight, S., Andersson, I., Branden, C.-I. (1990). Crystallographic analysis of ribulose-1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution. *J. Mol. Biol.* *215*, 113-160.
- Koumoto, Y., Shimada, T., Kondo, M., Hara-Nishimura, I., Nishimura, M. (2001). Chloroplasts have a novel Cpn10 in addition to Cpn20 as co-chaperonins in *Arabidopsis thaliana*. *J. Biol. Chem.* *276*, 29688-29694.
- Koumoto, Y., Shimada, T., Kondo, M., Takao, T., Shimonishi, Y., Hara-Nishimura, I., 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., Vigh, L. (2001). The chaperonins of *Synechocystis* PCC 6803 differ in heat inducibility and chaperone activity. *Biochem. Biophys. Res. Commun.* *289*, 908-915.
- Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. *Nature* *419*, 171-174.

- Kraulis, P.J. (1991). MOLSCRIPT - A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* *24*, 946-950.
- Kubota, H., Hynes, G., Willison, K. (1995). The chaperonin containing t-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol. *Eur. J. Biochem.* *230*, 3-16. Review.
- Kunisawa, J., Shastri, N. (2006). Hsp90 $\alpha$  chaperones large C-terminally extended proteolytic intermediates in the MHC class I antigen processing pathway. *Immunity* *24*, 523-534.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* *227*, 680-685.
- Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., Gierasch, L.M. (1993). Characterization of a functionally important mobile domain of GroES. *Nature* *364*, 255-258.
- Landry, S.J., Bartlett, S.G. (1989). The small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase and its precursor expressed in *Escherichia coli* are associated with groEL protein. *J. Biol. Chem.* *264*, 9090-9093.
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W., Hartl, F.U. (1992a). Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *EMBO J.* *11*, 4757-4765.
- Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., Hartl, F.U. (1992b). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* *356*. 683-689.
- Lansbury, P.T. Jr. (1999). Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. USA* *96*, 3342-3344.
- Lanzetta, P.A., Alvarez, L.J., Reinach, P.S., Candia, O.A. (1979). An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* *100*, 95-97.
- Larimer, F.W., Soper, T.S. (1993). Overproduction of *Anabaena* 7120 ribulose-bisphosphate carboxylase/oxygenase in *Escherichia coli*. *Gene* *126*, 85-92.
- Larimer, F.W., Machanoff, R., Hartman, F.C. (1986). A reconstruction of the gene for ribulose bisphosphate carboxylase from *Rhodospirillum rubrum* that expresses the authentic enzyme in *Escherichia coli*. *Gene* *41*, 113-20.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* *26*, 283-291.

- Lee, B., Read, B.A., Tabita, F.R. (1991). Catalytic properties of recombinant octameric, hexadecameric, and heterologous cyanobacterial/bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* *291*, 263-269.
- Lee, B., Tabita, F.R. (1990). Purification of recombinant ribulose-1,5-bisphosphate carboxylase/oxygenase large subunits suitable for reconstitution and assembly of active L<sub>8</sub>S<sub>8</sub> enzyme. *Biochemistry* *29*, 9352-9357.
- Lehel, C., Los, D., Wada, H., Gyorgyey, J., Horvath, I., Kovacs, E., Muata, N., 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., Fandrich, M., Klunker, D., Siegers, K., Lupas, A.N., Brown, J.R., Schiebel, E., Dobson, C.M., Hartl, F.U. (1999). MtGimC, a novel archaeal chaperone related to the eukaryotic chaperonin cofactor GimC/prefoldin. *EMBO J.* *18*, 6730-6743.
- Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACBM Newsletter* *26*.
- Levinthal, C. (1969). How to fold graciously. In DeBrunner, J.T.P. and Munck, E. (eds.), *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois*. University of Illinois Press, 22-24.
- Levy-Rimler, G., Bell, R.E., Ben-Tal, N., Azem, A. (2002). Type I chaperonins: not all rare created equal. *FEBS Lett.* *529*, 1-5. Review.
- Li, L.A., 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. *J. Bacteriol.* *179*, 3793-3796.
- Li, L.A., Tabita, F.R. (1994). Transcriptional control of ribulose 1,5-bisphosphate carboxylase/oxygenase activase and adjacent genes in *Anabaena* species. *J. Bacteriol.* *176*, 6697-6706.
- Liu, Y., Eisenberg, D. (2002). 3D domain swapping: As domains continue to swap. *Prot. Sci.* *11*, 1285-1299.
- Lorimer, G.H., Miziorko, H.M. (1980). Carbamate formation on the  $\epsilon$ -amino group of a lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by CO<sub>2</sub> and Mg<sup>2+</sup>. *Biochemistry* *19*, 5321-5328.
- Lubben, T.H., Donaldson, G.K., Viitanen, P.V., Gatenby, A.A. (1989). Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell* *1*, 1223-1230.

- Madueno, F., Napier, J.A., Gray, J.C. (1993). Newly Imported Rieske Iron-Sulfur Protein Associates with Both Cpn60 and Hsp70 in the Chloroplast Stroma. *Plant Cell* 5, 1865-1876.
- Maeda, N., Kitano, K., Fukui, T., Ezaki, S., Atomi, H., Miki, K., Imanaka, T. (1999). Ribulose biphosphate carboxylase/oxygenase from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1 is composed solely of large subunits and forms a pentagonal structure. *J. Mol. Biol.* 293, 57-66.
- Marcus, Y., Gurevitz, M. (2000). Activation of cyanobacterial RuBP-carboxylase/oxygenase is facilitated by inorganic phosphate *via* two independent mechanisms. *Eur. J. Biochem.* 267, 5995-6003.
- Martel, R., Cloney, L.P., Pelcher, L.E., Hemmingsen, S.M. (1990). Unique composition of plastid chaperonin-60:  $\alpha$  and  $\beta$  polypeptide-encoding genes are highly divergent. *Gene* 94, 181-187.
- Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., Hartl, F.U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* 352, 36-42.
- Mayhew, M., da Silva, A.C., Martin, J., Erdjument-Bromage, H., Tempst, P., Hartl, F.U. (1996). Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* 379, 420-426.
- Merrit, E.A., Bacon, D.J. (1997). Raster 3D : Photorealistic molecular graphics. *Methods Enzymol.* 276, 505-524.
- Miernyk, J.A. (1999). Protein folding in the plant cell. *Plant Physiol.* 121, 695-703. Review.
- Mitchell, P. (1966). Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biol. Rev. Camb. Philos. Soc.* 41, 445-502. Review.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta. Crystallogr. D Biol. Crystallogr.* 53, 240-255.
- Naylor, D.J., Hartl, F.U. (2001). Contribution of molecular chaperones to protein folding in the cytoplasm of prokaryotic and eukaryotic cells. *Biochem. Soc. Symp.* 68, 45-68.
- Nelson, D.L., Cox, M.M. (2005). *Lehninger Principles of Biochemistry*. Fourth edition, W.H. Freeman and Company, New York, p. 723-770.
- Netzer, W.J., Hartl, F.U. (1998). Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. *Trends Biochem. Sci.* 23, 68-73.

- Neupert, W. (1997). Protein import into mitochondria. *Annu. Rev. Biochem.* *66*, 863-917.
- Newman, J., Bränden, C.I., and Jones, T.A. (1993). Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Synechococcus* PCC6301. *Acta Cryst. D. Biol. Crystallogr.* *49*, 548-560.
- Newman, J., Gutteridge, S. (1993). The X-ray structure of *Synechococcus* ribulose-bisphosphate carboxylase/oxygenase-activated quaternary complex at 2.2-Å resolution. *J. Biol. Chem.* *268*, 25876-25886.
- Newman, J., Gutteridge, S. (1990). The purification and preliminary X-ray diffraction studies of recombinant *Synechococcus* ribulose-1,5-bisphosphate carboxylase/oxygenase from *Escherichia coli*. *J. Biol. Chem.* *265*, 15154-15159.
- Nishihara, K., Kanemori, M., Yanagi, H., Yura, T. (2000). Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Appl. Environ. Microbiol.* *66*, 884-889.
- Nishio, K., Hirohashi, T., 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.
- Onizuka, T., Endo, S., Akiyama, H., Kanai, S., Hirano, M., Yokota, A., Tanaka, S., 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 Physiol.* *45*, 1390-1395.
- Onizuka, T., Akiyama, H., Endo, S., Kanai, S., Hirano, M., Tanaka, S., Miyasaka, H. (2002). CO(2) response element and corresponding trans-acting factor of the promoter for ribulose-1,5-bisphosphate carboxylase/oxygenase genes in *Synechococcus* sp. PCC7002 found by an improved electrophoretic mobility shift assay. *Plant Cell Physiol.* *43*, 660-667.
- Parry, M.A., Andralojc, P.J., Mitchell, R.A., Madgwick, P.J., Keys, A.J. (2003). Manipulation of Rubisco: the amount, activity, function and regulation. *J. Exp. Bot.* *54*, 1321-1333. Review.
- Perrakis, A., Morris, R., Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* *6*, 458-463.
- Perutz, M.F. (1999). Glutamine repeats and neurodegenerative diseases: molecular aspects. *Trends Biochem. Sci.* *24*, 58-63. Review.
- Pierce, M.M., Raman, C.S., Nall, B.T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods* *19*, 213-221.

- Portis, A.R. (2003). Rubisco activase – Rubisco's catalytic chaperone. *Photosynth. Res.* *75*, 11-27.
- Portis, A.R. (1990). Rubisco activase. *Biochim. Biophys. Acta* *1015*, 15-28.
- Privalov, P.L. (1996). Intermediate states in protein folding. *J. Mol. Biol.* *258*, 707-725.
- Prusiner, S.B. (1997). Prion diseases and the BSE crisis. *Science* *278*, 245-251. Review.
- Pushkin, A.V., Tsuprun, V.L., Solovjeva, N.A., Shubin, V.V., Evstigneeva, Z.G., Kretovich, W.L. (1982). High molecular weight pea leaf protein similar to the *groE* protein of *Escherichia coli*. *Biochim. Biophys. Acta* *704*, 379-384.
- Ranson, N.A., Farr, G.W., Roseman, A.M., Gowen, B., Fenton, W.A., Horwich, A.L., Saibil, H.R. (2001). ATP-bound states of GroEL captured by cryo-electron microscopy. *Cell* *107*, 867-879.
- 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., Ranson, N.A., Gowen, B., Fuller, S.D., Saibil, H.R. (2001). Structures of unliganded and ATP-bound states of the *Escherichia coli* chaperonin GroEL by cryoelectron microscopy. *J. Struct. Biol.* *135*, 115-125.
- Roseman, A.M., Chen, S., White, H., Braig, K., Saibil, H.R. (1996). The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. *Cell* *87*, 241-251.
- Rospert, S., Glick, B.S., Jenö, P., Schatz, G., Todd, M.J., Lorimer, G.H., Viitanen, P.V. (1993). Identification and functional analysis of chaperonin 10, the *groES* homolog from yeast mitochondria. *Proc. Natl. Acad. Sci. USA* *90*, 10967-10971.
- Roy, H. (1989). Rubisco assembly: a model system for studying the mechanism of chaperonin action. *Plant Cell* *1*, 1035-1042. Review.
- Roy, H., Cannon, S. (1988). Ribulose biphosphate carboxylase assembly: What is the role of the large subunit binding protein? *Trends Biochem. Sci.* *13*, 163-165.
- Roy, H., Cannon, S., Gilson, M. (1988). Assembly of Rubisco from native subunits. *Biochim. Biophys. Acta.* *957*, 323-334. Review.
- Roy, H., Bloom, M., Milos, P., Monroe, M. (1982). Studies on the assembly of large subunits of ribulose biphosphate carboxylase in isolated pea chloroplasts. *J. Cell Biol.* *94*, 20-27.

- Rye, H.S., Burston, S.G., Fenton, W.A., Beechem, J.M., Xu, Z., Sigler, P.B., Horwich, A.L. (1997). Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. *Nature* 388, 792-798.
- Sambrook, J., Fritsch, E., Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Second Edition. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Schägger, H., von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* 199, 223-231.
- Schlegel, H.G. (1992). *Allgemeine Mikrobiologie*. 7. Auflage, Thieme Verlag, Stuttgart.
- Schlicher, T., Soll, J. (1996). Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. *FEBS Lett.* 379, 302-304.
- Schmid, F.X. (1990). Spectral methods of characterizing protein conformation and conformational changes. *Protein Structure. A Practical Approach* (Creighton, E.T., ed). IRL Press, Oxford.
- Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Viitanen, P., Lorimer, G., Buchner, J. (1994). Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science* 265, 656-659.
- Schneider, D. (2000). *Untersuchungen zur Struktur, Funktion und Regulation des cyanobakteriellen Cytochrom b<sub>6</sub>f-Komplexes*. Dissertation. Ruhr-Universität Bochum.
- Schneider, G., Lindqvist, Y., Lundqvist, T. (1990a). Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 1.7 Å resolution. *J. Mol. Biol.* 211, 989-1008.
- Schneider, G., Knight, S., Andersson, I., Branden, C.I., Lindqvist, Y., Lundqvist, T. (1990b). Comparison of the crystal structures of L<sub>2</sub> and L<sub>8</sub>S<sub>8</sub> Rubisco suggests a functional role for the small subunit. *EMBO J.* 9, 2045-2050.
- Schneider, G., Lindqvist, Y., Branden, C.I., Lorimer, G. (1986). Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 2.9 Å resolution. *EMBO J.* 5, 3409-3415.
- Schreuder, H.A., Knight, S., Curmi, P.M., Andersson, I., Cascio, D., Sweet, R.M., Branden, C.I., 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.
- Schultz, C.P. (2000). Illuminating folding intermediates. *Nat. Struct. Biol.* 7, 7-10.

- Sharkia, R., Bonshtien, A.L., Mizrahi, I., Weiss, C., Niv, A., Lustig, A., Viitanen, P.V., Azem, A. (2003). On the oligomeric state of chloroplast chaperonin 10 and chaperonin 20. *Biochim. Biophys. Acta* *1651*, 76-84.
- Shively, F.M., van Keulen, G., Meijer, W.G. (1998). Something from almost nothing: carbon dioxide fixation in chemoautotrophs. *Ann. Rev. Microbiol.* *52*, 192-230. Review.
- Shtilerman, M., Lorimer, G.H., 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., 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., Moarefi, I. (2000). 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., Horwich, A.L. (1998). Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.* *67*, 581-608. Review.
- Smith, S.A., Tabita, F.R. (2003). Positive and negative selection of mutant forms of prokaryotic (cyanobacterial) ribulose-1,5-bisphosphate carboxylase/oxygenase. *J. Mol. Biol.* *331*, 557-569.
- 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.
- Spreitzer, R.J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* *414*, 141-149. Review.
- Spreitzer, R.J., Salvucci, M.E. (2002). Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.* *53*, 449-475. Review.
- Stefani, M., Dobson, C.M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* *81*, 678-699. Review.
- Stoscheck, C.M. (1990). Quantitation of protein. *Methods Enzymol.* *182*, 50-68.
- Tabita, F.R. (1999). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynthesis Research* *60*, 1-28. Review.

- Tabita, F.R., Small, C.L. (1985). Expression and assembly of active cyanobacterial ribulose-1,5-bisphosphate carboxylase/oxygenase in *Escherichia coli* containing stoichiometric amounts of large and small subunits. *Proc. Natl. Acad. Sci. USA* *82*, 6100-6103.
- Tabita, F.R., McFadden, B.A. (1974a). D-ribulose 1,5-diphosphate carboxylase from *Rhodospirillum rubrum*. I. Levels, purification, and effects of metallic ions. *J. Biol. Chem.* *249*, 3453-3458.
- Tabita, F.R., McFadden, B.A. (1974b). D-ribulose 1,5-diphosphate carboxylase from *Rhodospirillum rubrum*. II. Quaternary structure, composition, catalytic, and immunological properties. *J. Biol. Chem.* *249*, 3459-3464.
- Tang, Y.C., Chang, H.C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M.J., Hartl, F.U., Hayer-Hartl, M. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* *125*, 903-914.
- Tcherkez, G.G., Farquhar, G.D., Andrews, T.J. (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. USA* *103*, 7246-7251.
- Terwilliger, T.C. (2000). Maximum-likelihood density modification. *Acta. Crystallogr. D Biol. Crystallogr.* *56 (Pt 8)*, 965-972.
- Terwilliger, T.C., Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta. Crystallogr. D Biol. Crystallogr.* *55 (Pt 4)*, 849-861.
- Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., 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., Frydman, J. (1999). *In vivo* newly translated polypeptides are sequestered in a protected folding environment. *EMBO J.* *18*, 85-95.
- Todd, M.J., Lorimer, G.H., Thirumalai, D. (1996). Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci. USA* *93*, 4030-4035.
- Towbin, H., Staehelin, T., Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* *76*, 4350-4354.
- Tsugeki, R., Nishimura, M. (1993). Interaction of homologues of Hsp70 and Cpn60 with ferredoxin-NADP<sup>+</sup> reductase upon its import into chloroplasts. *FEBS Lett.* *320*, 198-202.

- Tsuprun, V.L., Boekema, E.J., Samsonidze, T.G., 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* 289, 205-209.
- Vagin, A.A., Isupov, M.N. (2001). Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D Biol. Crystallogr.* 57, 1451-1456.
- Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* 93, 863-873.
- Van der Vies, S.M., Bradley, D., 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.
- Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* 229, 105-124.
- Viale, A.M., Arakaki, A.K. (1994). The chaperone connection to the origins of the eukaryotic organelles. *FEBS Lett.* 341, 146-151.
- Viale, A.M., Kobayashi, H., Takabe, T., Akazawa, T. (1985). Expression of genes for subunits of plant-type RuBisCO from *Chromatium* and production of the enzymatically active molecule in *Escherichia coli*. *FEBS Lett.* 192, 283-288.
- Viitanen, P.V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G.H., Gatenby, A.A., Soll, J. (1995). Functional characterization of the higher plant chloroplast chaperonins. *J. Biol. Chem.* 270, 18158-18164.
- Viitanen, P.V., Gatenby, A.A., Lorimer, G.H. (1992). Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of *Escherichia coli* proteins. *Protein Sci.* 1, 363-369.
- Viitanen, P.V., Lubben, T.H., Reed, J., Goloubinoff, P., O'Keefe, D.P., Lorimer, G.H. (1990). Chaperonin-facilitated refolding of ribulosebisphosphate carboxylase and ATP hydrolysis by chaperonin 60 (groEL) are K<sup>+</sup> dependent. *Biochemistry* 29, 5665-5671.
- Waldmann, T., Nimmegern, E., Nitsch, M., Peters, J., Pfeifer, G., Muller, S., Kellermann, J., Engel, A., Hartl, F.U., Baumeister, W. (1995). The thermosome of *Thermoplasma acidophilum* and its relationship to the eukaryotic chaperonin TRiC. *Eur. J. Biochem.* 227, 848-856.
- Walter, S., Buchner, J. (2002). Molecular chaperones – cellular machines for protein folding. *Angew. Chem. Int. Ed.* 41, 1098-1113.

- Wastl, J., Fraunholz, M., Zauner, S., Douglas, S., Maier, U.G. (1999). Ancient gene duplication and differential gene flow in plastid lineages: the GroEL/Cpn60 example. *J. Mol. Evol.* *48*, 112-117.
- Weiner, M.P., Costa, G.L., Schoettlin, W., Cline, J., Mathur, E., Bauer, J.C. (1994). Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene* *151*, 119-123.
- Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* *84*, 481-490.
- Wickner, S., Maurizi, M.R., Gottesman, S. (1999). Posttranslational quality control: folding, refolding, and degrading proteins. *Science* *286*, 1888-1893.
- Wildner, G.F. (1976). The use of 1-anilino-8-naphthalene sulfonate as fluorescent probe for conformational studies on ribulose-1,5-bisphosphate carboxylase. *Z. Naturforsch.* *31c*, 267-271.
- Wiseman, T., Williston, S., Brandts, J.F., Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* *179*, 131-137.
- Witt, E., Zantopf, D., Schmidt, M., Kraft, R., Kloetzel, P.M., Krüger, E. (2000). Characterisation of the newly identified human Ump1 homologue POMP and analysis of LMP7( $\beta$ 5i) incorporation into 20 S proteasomes. *J. Mol. Biol.* *301*, 1-9.
- Wolosiuk, R.A., Ballicora, M.A., Hagelin, K. (1993). The reductive pentose phosphate cycle for photosynthetic CO<sub>2</sub> assimilation: enzyme modulation. *FASEB J.* *7*, 622-637. Review.
- Wyatt, P.J. (1993). Light scattering and the absolute characterization of macromolecules. *Anal. Chim. Acta* *272*, 1-40. Review.
- Xu, Y., Horwich, A.L., Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroEL-(ADP)<sub>7</sub> chaperonin complex. *Nature* *388*, 741-750.
- Yang, T., Poovaiah, B.W. (2002). *Arabidopsis* chloroplast chaperonin 10 is a calmodulin-binding protein. *Biochem. Biophys. Res. Commun.* *275*, 601-607.
- Young, J.C., Agashe, V.R., Siegers, K., Hartl, F.U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* *5*, 781-791. Review.
- Young, J.C., Moarefi, I., Hartl, F.U. (2001). Hsp90: a specialized but essential protein-folding tool. *J. Cell Biol.* *154*, 267-273. Review.

Zabaleta, E., Oropeza, A., Assad, N., Mandel, A., Salerno, G., Herrera-Estrella, L. (1994). Antisense expression of chaperonin 60 $\beta$  in transgenic tobacco plants leads to abnormal phenotypes and altered distribution of photoassimilates. *Plant J.* 6, 425-432.

Zabaleta, E., Oropeza, A., Jimenez, B., Salerno, G., Crespi, M., Herrera-Estrella, L. (1992). Isolation and characterization of genes encoding chaperonin 60  $\beta$  from *Arabidopsis thaliana*. *Gene* 111, 175-181.

Zhang, N., Kallis, R.P., Ewy, R.G., Portis, A.R. Jr. (2002). Light modulation of Rubisco in *Arabidopsis* requires a capacity for redox regulation of the larger Rubisco activase isoform. *Proc. Natl. Acad. Sci. USA* 99, 3330-3334.

## 7 APPENDICES

### 7.1 Restriction enzymes

Enzyme	company	Recognition site (5' → 3')	buffer	BSA	temperature
<i>Bam</i> HI	New England Biolabs	GGATCC	NEB <i>Bam</i> HI	+	37 °C
<i>Ehe</i> I	Fermentas	GGCGCC	Y+/Tango	-	37 °C
<i>Hind</i> III	New England Biolabs	AAGCTT	NEB2	-	37 °C
<i>Nde</i> I	New England Biolabs	CATATG	NEB2	-	37 °C
<i>Nhe</i> I	New England Biolabs	GCTAGC	NEB2	+	37 °C
<i>Sph</i> I	New England Biolabs	GCATGC	NEB2	-	37 °C
<i>Xba</i> I	New England Biolabs	TCTAGA	NEB2	+	37 °C

**Table A1. Restriction enzymes.**

### 7.2 Oligonucleotides

Name	Gene	Sequence (5' → 3')	Purpose
<b>Syn6301-<i>rbcL</i></b>			
<i>rbcL</i> fw <i>Nde</i> I	Syn6301- <i>rbcL</i>	GGAGAGACCATATGCCCAAGACGCAATCTGCCGAGGCTATAAGG	<i>Nde</i> I, start
<i>rbcL</i> rev <i>Bam</i> HI	Syn6301- <i>rbcL</i>	ATAGTCAGAGGATCCCTTAGAGCTTGTCCATCGTTTCGAATTCG	<i>Bam</i> HI, stop
Syn6301_L_1-460_revNheBam	Syn6301- <i>rbcL</i>	GAGCAGGGATCCGCTAGCTTATTTCCAGAGGTCGAGGGCAGCAGC	aa 1-460, stop, <i>Bam</i> HI, <i>Nhe</i> I
<b>Syn6301-<i>rbcS</i></b>			
<i>rbcS</i> fw <i>Nde</i> I	Syn6301- <i>rbcS</i>	GGATTCTGACATATGAGCATGAAAACCTCTGCCCAAAGAGCGTGC	<i>Nde</i> I, start
<i>rbcS</i> rev <i>Bam</i> HI	Syn6301- <i>rbcS</i>	CAAGACAAAGGATCCCTTAGTAGCGGCCGGACGATGAACGATG	<i>Bam</i> HI, stop
6301-RbcS_N-FLAG, <i>Nde</i> I	Syn6301- <i>rbcS</i>	GCGACGCATATGGATTACAAAGACGATGACGATAAAGCGGGCATGAGCATGAAAACCTCTGCCCAAAG	<i>Nde</i> I, start, FLAG-tag
<b>Syn7002-<i>rbcL</i></b>			
Syn7002_L_fwNde_pET	Syn7002- <i>rbcL</i>	ACTAGTCATATGGTTTCAGACCAAATCTGCTGGGTTT	<i>Nde</i> I, start
Syn7002_L_revNheBam_pET	Syn7002- <i>rbcL</i>	GAGCAGGGATCCGCTAGCTTAGAGAGTGTCAACGGTATCGAATTC	<i>Bam</i> HI, <i>Nhe</i> I, stop
Syn7002_L_1-459_revNheBam	Syn7002- <i>rbcL</i>	GAGCAGGGATCCGCTAGCTTACTTCCACAGGTCGAGGGCGGCTGC	aa 1-459, stop, <i>Bam</i> HI, <i>Nhe</i> I
7002L_F462,464A_fw,uni	Syn7002- <i>rbcL</i>	CTGAATTGGCAGCCGCCCTCGACCTGTGGAAG	F462A
7002L_F462A_revPhos	Syn7002- <i>rbcL</i>	CAACGGTATCGAATTCGGCCTTGATTTC	F462A
7002L_F462,464A_fw,uni	Syn7002- <i>rbcL</i>	CTGAATTGGCAGCCGCCCTCGACCTGTGGAAG	F464A
7002L_F464A_revPhos	Syn7002- <i>rbcL</i>	CAACGGTATCGGCTTCGAACTTGATTTC	F464A
7002L_F462,464A_fw,uni	Syn7002- <i>rbcL</i>	CTGAATTGGCAGCCGCCCTCGACCTGTGGAAG	F462A/F464A
7002L_F462,464A_revPhos	Syn7002- <i>rbcL</i>	CAACGGTATCGGCTTCGGCCTTGATTTC	F462A/F464A
<b>Syn7002-<i>rbcS</i></b>			
Syn7002_S_fwNde_pET	Syn7002- <i>rbcS</i>	ACTAGTCATATGAAAACCTTACCTAAAGAAAAGCGT	<i>Nde</i> I, start
Syn7002_S_revBam_pET	Syn7002- <i>rbcS</i>	GAGCAGGGATCCCTTAGTAAACGGGTTTGGTTGGGCTT	<i>Bam</i> HI, stop
7002-RbcS_N-FLAG, <i>Nde</i> I	Syn7002- <i>rbcS</i>	GCGACGCATATGGATTACAAAGACGATGACGATAAAGCGGGCATGAAAACCTTACCTAAAGAAAAGC	<i>Nde</i> I, start, FLAG-tag
<b>Syn7002-<i>rbcX</i></b>			
Syn7002_X_fwNde_pET	Syn7002- <i>rbcX</i>	ACTAGTCATATGGAGTTTAAAAAAGTTGCGAAGGAA	<i>Nde</i> I, start
Syn7002_X_revBam_pET	Syn7002- <i>rbcX</i>	GAGCAGGGATCCCTTAGTCAAGATCTTCTGAAGTATC	<i>Bam</i> HI, stop
7002X_fwNde_N-FLAG	Syn7002- <i>rbcX</i>	GCGACGCATATGGATTACAAAGACGATGACGATAAAGCGGGCATGGAGTTTAAAAAAGTTGCG	<i>Nde</i> I, start, FLAG-tag
13GCT fw Phos	Syn7002- <i>rbcX</i>	GGCCATCGCTTTGCAAAGCTATTTGACC	T13A
13,14 rev Phos	Syn7002- <i>rbcX</i>	GTTTCCTTCGCAACTTTTTTAAACTCC	T13A
14GAA fw Phos	Syn7002- <i>rbcX</i>	GGCCATCACTGAACAAAAGCTATTTGACC	L14E
13,14 rev Phos	Syn7002- <i>rbcX</i>	GTTTCCTTCGCAACTTTTTTAAACTCC	L14E
7002X_Y17A_fw	Syn7002- <i>rbcX</i>	CCATCACTTTGCAAAGCGCTTTGACCTACCAAGCG	Y17A
7002X_Ymut rev Phos	Syn7002- <i>rbcX</i>	CCGTTTCCTTCGCAACTTTTTTAAACTCC	Y17A
7002X_Y20A_fw	Syn7002- <i>rbcX</i>	CCATCACTTTGCAAAGCTATTTGACCGCCAAGCG	Y20A
7002X_Ymut rev Phos	Syn7002- <i>rbcX</i>	CCGTTTCCTTCGCAACTTTTTTAAACTCC	Y20A
7002X_20mutII fw	Syn7002- <i>rbcX</i>	CAAAGCTATTTGACCCTTCAAAGCGGTGCGTC	Y20L
7002X_20mutII revPhos	Syn7002- <i>rbcX</i>	CAAAGTATGATGGCCGTTTCCTTCGCAAC	Y20L
7002X_Y17A_Y20L_fw Phos	Syn7002- <i>rbcX</i>	CCATCACTTTGCAAAGCGCTTTGACCTGCAAGCG	Y17A/Y20L
7002X_Ymut rev Phos	Syn7002- <i>rbcX</i>	CCGTTTCCTTCGCAACTTTTTTAAACTCC	Y17A/Y20L
7002X_Q29A_fw	Syn7002- <i>rbcX</i>	GTCTAATTAGTCAGGCCCTTAGTGAAACCAATCC	Q29A
7002X_Q29A_revPhos	Syn7002- <i>rbcX</i>	GCACCGCTTGGTAGGTCAAATAGCTTTGC	Q29A
7002X_E32A_fw	Syn7002- <i>rbcX</i>	GCAGCTTAGTGCAACCAATCCTGGACAGGCG	E32A
7002X_E32A_revPhos	Syn7002- <i>rbcX</i>	TGACTAATTAGACGCACCGCTTGGTAGG	E32A

Name	Gene	Sequence (5' → 3')	Purpose
7002X_T33A_fw	Syn7002- <i>rbcX</i>	CAGCTTAGTGAAGCGAATCTGGACAGGCGATTGG	<u>T33A</u>
7002X_33-35_revPhos	Syn7002- <i>rbcX</i>	CTGACTAATTAGACGCACCGCTTGGTAGG	T33A
7002X_N34A_fw	Syn7002- <i>rbcX</i>	CAGCTTAGTGAACCCGCGCTGGACAGGCGATTGG	<u>N34A</u>
7002X_33-35_revPhos	Syn7002- <i>rbcX</i>	CTGACTAATTAGACGCACCGCTTGGTAGG	N34A
7002X_P35A_fw	Syn7002- <i>rbcX</i>	CAGCTTAGTGAACCAATGCGGGACAGGCAATTGG	<u>P35A</u>
7002X_33-35_revPhos	Syn7002- <i>rbcX</i>	CTGACTAATTAGACGCACCGCTTGGTAGG	P35A
7002X_W40A_fw	Syn7002- <i>rbcX</i>	GACAGGCGATTGCGCTAGGAGATTCTCTAAACG	<u>W40A</u>
7002X_W40A_revPhos	Syn7002- <i>rbcX</i>	CAGGATTGGTTTCACTAAGCTGTCTGAC	W40A
7002X_Q51A_fw	Syn7002- <i>rbcX</i>	TCTAAACGTCATCCAATTGCGGAAAGTGATC	<u>Q51A</u>
7002X_Q51A_revPhos	Syn7002- <i>rbcX</i>	GAACTCTCTAGC AAAATCGCCTGTCCAGG	Q51A
7002X_E63A_fw	Syn7002- <i>rbcX</i>	GAAGCGATGATGCTAGCGAACAAAGAGCTCG	<u>E63A</u>
7002X_E63A_revPhos	Syn7002- <i>rbcX</i>	GAGGTAAAGATCACTTTCCTGAATTGG	E63A
70GCA_fw Phos	Syn7002- <i>rbcX</i>	CGTCTCGCAATCCTGACGTTGCGAG	<u>R70A</u>
70_rev Phos	Syn7002- <i>rbcX</i>	AGCTCTTGTGTTTCTAGCATCATCGC	R70A
7002X_V74A_fw	Syn7002- <i>rbcX</i>	CTCAGAATCCTGACGGCGGAGAAAACCTTGCGG	<u>V74A</u>
7002X_V74A_revPhos	Syn7002- <i>rbcX</i>	GACGAGCTCTTGTGTTTCTAGCATCATCGC	V74A
7002X_E76A_fw	Syn7002- <i>rbcX</i>	CTCAGAATCCTGACGGTGCGAGCAAACCTTGCGG	<u>E76A</u>
7002X_V74A_revPhos	Syn7002- <i>rbcX</i>	GACGAGCTCTTGTGTTTCTAGCATCATCGC	E76A
7002X_RE75LA_fw	Syn7002- <i>rbcX</i>	CTCAGAATCCTGACGGTGCTGGCAAACCTTGCGGAAG	<u>R75L/E76A</u>
7002X_RE75LA_revPhos	Syn7002- <i>rbcX</i>	GACGAGCTCTTGTGTTTCTAGCATCATCGC	R75L/E76A
7002X_F85A_fw	Syn7002- <i>rbcX</i>	GAGTTCGGAGGCGTTGCCAGAAAATGGTCC	<u>F85A</u>
7002X_F85A_revPhos	Syn7002- <i>rbcX</i>	CTTCCGCAAGGTTTTCTCGCACCGTCAG	F85A
7002X_P87A_fw Phos	Syn7002- <i>rbcX</i>	GGAGTTCGGAGTTTTTGGCGGAAAATGGTCC	<u>P87A</u>
7002X_P87A_rev Phos	Syn7002- <i>rbcX</i>	TTCCGCAAGGTTTTCTCGCACCGTCAGG	P87A
7002X_M89A_fw	Syn7002- <i>rbcX</i>	GGAGTTTTGCCAGAAGCGGCTCTCAGCAAATC	<u>M89A</u>
7002X_M89A_revPhos	Syn7002- <i>rbcX</i>	AGAACTCTTCCGCAAGGTTTTCTCGCAC	M89A
90GAA_fw Phos	Syn7002- <i>rbcX</i>	GAGTTTTGCCAGAAAATGGAACCTCAGCC	<u>V90E</u>
90_rev Phos	Syn7002- <i>rbcX</i>	CAGAACTCCTCCGCAAGGTTTTCTCGC	V90E
7002X_N98A_fw	Syn7002- <i>rbcX</i>	CCAAAATCAAGCAGTCCGCGGAAAACCATCGCC	<u>N98A</u>
7002X_N98A_revPhos	Syn7002- <i>rbcX</i>	CTGAGGACCATTCTGGCAAAAACCTCC	N98A
7002X_R102A_fw	Syn7002- <i>rbcX</i>	CAATGGAAAACCATGCGGCTCTTTATTAGAGCG	<u>R102A</u>
7002X_R102A_revPhos	Syn7002- <i>rbcX</i>	GACTGCTTGATTGGCTGAGGACCATTCTGG	R102A
7002X_E107A_fw	Syn7002- <i>rbcX</i>	CCGTCTTTATTAGCGGCTTAACTCAAGTTG	<u>E107A</u>
7002X_L107A_revPhos	Syn7002- <i>rbcX</i>	CGATGGTTTCCATTGGACTGCTGATTGGC	E107A
7002X_revBam_L83stop	Syn7002- <i>rbcX</i>	GAGCAGGGATCCTTACAGAACTCCTCCGCAAGGTTTTCTCGC	aa 1-83, <b>BamHI</b> , stop
7002X_L105stop_revBAM	Syn7002- <i>rbcX</i>	GAGGACGGATCCTTATAAAGAACGGCGATGGTTTCC	aa 1-105, <b>BamHI</b> , stop
7002X_revBam_L109stop	Syn7002- <i>rbcX</i>	GAGCAGGGATCCTTATAAACGCTCTAATAAAGAACGGCGATGG	aa 1-109, <b>BamHI</b> , stop
<b>AnaCA-rbcX</b>			
AnaCA_X_fwNde pET	AnaCA- <i>rbcX</i>	ACTAGTCAT <b>ATG</b> AACCTCAAGCAAATTCGGAAGATACAGCC	<b>NdeI</b> , start
AnaCA_X_revBam_pUC/pET	AnaCA- <i>rbcX</i>	GAGCAGGGAT <b>CCT</b> TAAGTAAAGGAAATAGCTTTTGACCAGCAC	<b>BamHI</b> , stop
AnaCA_X_fwNde_FLAG	AnaCA- <i>rbcX</i>	GCGACGCAT <b>ATG</b> GATTACAAAGACGATGACGATAAAGCGGCATGAACCTC AAGCAAATGGC	<b>NdeI</b> , start, <b>FLAG</b> -tag
<b>AtCpn60a</b>			
ara60alphafwNdeI	<i>AtCpn60a</i>	TAGTGTT <b>CATATG</b> AATGTAAGGAAATAGCTTTTGACCAGCAC	<b>NdeI</b> , start
alpha rev NheI_BamHI	<i>AtCpn60a</i>	GTGGGATCCGCGAGT <b>GTAG</b> TCTCC <b>TAC</b> ACCATGAGACCTCAGGAGCAGC AGC	<b>BamHI</b> , <b>NheI</b> , stop
<b>AtCpn60β</b>			
ara60betafwNdeI	<i>AtCpn60β</i>	CTGCAAT <b>CATATG</b> GCGAGCAAAGGAATTACATTCAACAAAG	<b>NdeI</b> , start
AtCpn60beta_revBamHI	<i>AtCpn60β</i>	GGTCGGGGAT <b>CCT</b> TAGTATCCGTAACCTGAGTTGTCCATTGG	<b>BamHI</b> , stop
<b>AtCpn20</b>			
a.th.cpn21_fw NdeI ATG	<i>AtCpn20</i>	TTGGTTGTCCAT <b>ATG</b> GCCTCTGTTGTTGCC	<b>NdeI</b> , start
at21_revBamHI	<i>AtCpn20</i>	AGCAGCCGGATCCAGCTTGTGTCAGATTAC	<b>BamHI</b>
ara21fwEhel	<i>AtCpn20</i>	GATATACATGGCGCTCTGTTGTTGCCCTAAGTATACTTCA	<b>Ehel</b>
ara21revHindIII	<i>AtCpn20</i>	AAAGATATAAAGCT <b>TCT</b> AAGAAAGTATAGCCATCACATCTGA	<b>HindIII</b> , stop
N-domain rev stop BamHI	<i>AtCpn20/N</i>	TGTCTCGGAT <b>CCT</b> TAACAATATCATCTTCTTGGAGGATAAGATG	<b>BamHI</b> , stop
C-domain fw ATG NdeI	<i>AtCpn20/C</i>	ACAGAGCAT <b>ATG</b> AAAGATCTCAAACCTTTGAATGACCGAGTC	<b>NdeI</b> , start
<b>AtCpn10</b>			
AtCpn10_fw44NdeI(ATG)	<i>AtCpn10</i>	ATCAAACAT <b>ATG</b> TCCACTAAATGGGAACCGACAAGGTTGTT	<b>NdeI</b> , start
AtCpn10_revBamHI	<i>AtCpn10</i>	TTGGACGGAT <b>CCT</b> ACTCAACGAGGGCCAACAAGTCACTCTC	<b>BamHI</b> , stop
ara10(aa44)fwEhel	<i>AtCpn10</i>	GGAGATATAGGGCCTCCACTAAATGGGAACCGACAAGGTT	<b>Ehel</b>
ara10(aa44)revHindIII	<i>AtCpn10</i>	TTAGCAGCCAAG <b>CTT</b> CACTCAACGAGGGCCAACAAGTCACT	<b>HindIII</b> , stop
<b>SoCpn20</b>			
s.ol.cpn21_fw NdeI ATG	<i>SoCpn20</i>	GTTGTT <b>CATATG</b> GCCTCAATAACCACATCG	<b>NdeI</b> , start
spinach21rev BamHI	<i>SoCpn20</i>	GTGTAGGAT <b>CCT</b> TAAGAAAGCACGCCATCACATCC	<b>BamHI</b> , stop
spinach21fw NcoI	<i>SoCpn20</i>	GATATAC <b>CTAGG</b> CTCAATAACCATCGAAGTAC	<b>NcoI</b>
s.ol.cpn21_rev HindIII	<i>SoCpn20</i>	GGGCTACAAT <b>CTAA</b> AGCTTCAACTGGCA	<b>HindIII</b>

**Table A2. Sequences of oligonucleotides (primers).**

Introduced endonuclease restriction sites are shown in bold. Start or stop codons are in italic and tags or mutation sites are underlined.

### 7.3 Expression of chloroplast chaperonins and co-chaperones

gene	chromosome	atg-code	expression: average leaves	expression : average roots
<i>AtCpn60α</i>	<b>2</b>	<b>at2g28000</b>	<b>1117.87</b>	<b>348.27</b>
<i>AtCpn60α</i>	5	at5g18820	8.43	4.47
<i>AtCpn60β</i>	<b>1</b>	<b>at1g55490</b>	<b>1347.37</b>	<b>143.00</b>
<i>AtCpn60β</i>	1	at1g26230	91.33	9.70
<i>AtCpn60β</i>	3	at3g13470	660.53	191.77
<i>AtCpn60β</i>	5	at5g56500	65.27	291.87
<i>AtCpn20</i>	<b>5</b>	<b>at5g20720</b>	<b>1553.13</b>	<b>643.40</b>
<i>AtCpn10</i>	<b>2</b>	<b>at2g44650</b>	<b>730.53</b>	<b>206.57</b>
<i>AtCpn10</i>	3	at3g60210	221.57	166.30

**Table A3. Expression of chloroplast chaperonin and co-chaperone genes in *A. thaliana*.**

Genes presented in bold have the highest expression level and were thus cloned, expressed and purified in this study. Data were kindly provided by Dr. E. Schleiff and Dr. T. Becker.

### 7.4 Standard amino acids

Amino acid			Side chain polarity	Side chain acidity or basicity	Mass (Da)	Aromatic or aliphatic	Hydrophobic
alanine	ala	A	nonpolar	neutral	89.09	-	x
arginine	arg	R	polar	basic	174.20	-	-
asparagine	asn	N	polar	neutral	132.12	-	-
aspartic acid	asp	D	polar	acidic	133.10	-	-
cysteine	cys	C	nonpolar	neutral	121.16	-	x
glutamic acid	glu	E	polar	acidic	147.13	-	-
glutamine	gln	Q	polar	neutral	146.15	-	-
glycine	gly	G	nonpolar	neutral	75.07	-	x
histidine	his	H	polar	basic	155.16	aromatic	-
isoleucine	ile	I	nonpolar	neutral	131.17	aliphatic	x
leucine	leu	L	nonpolar	neutral	131.17	aliphatic	x
lysine	lys	K	polar	basic	146.19	-	-
methionine	met	M	nonpolar	neutral	149.21	-	x
phenylalanine	phe	F	nonpolar	neutral	165.19	aromatic	x
proline	pro	P	nonpolar	neutral	115.13	-	x
serine	ser	S	polar	neutral	105.09	-	-
threonine	thr	T	polar	neutral	119.12	-	-
tryptophan	trp	W	nonpolar	neutral	204.23	aromatic	x
tyrosine	tyr	Y	polar	neutral	181.19	aromatic	x
valine	val	V	nonpolar	neutral	117.15	aliphatic	x

**Table A4. List of standard amino acids, abbreviations and properties.**

From Wikipedia encyclopedia ([www.wikipedia.org](http://www.wikipedia.org)).

## 7.5 Protein accession numbers

protein	species	accession Nr.
GroEL	Escherichia coli	P0A6F5
GroES	Escherichia coli	P0A6F9
DnaK	Escherichia coli	P0A6Z0
DnaJ	Escherichia coli	P08622
GrpE	Escherichia coli	Q1R8B1
AtCpn60 $\alpha$	Arabidopsis thaliana	P21238
AtCpn60 $\beta$	Arabidopsis thaliana	P21240
AtCpn20	Arabidopsis thaliana	O65282
AtCpn10	Arabidopsis thaliana	O80504
SoCpn20	Spinacia oleracea	Q02073
Gp31	Bacteriophage T4	P17313
RbcL	Synechococcus sp. PCC7002 (Agmenellum quadruplicatum)	Q44176
	Synechococcus elongatus PCC6301 (Anacystis nidulans)	P00880
	Microcystis aeruginosa NIVA-CYA 143	O88146
	Anabaena sp. PCC7120	P00879
	Aphanizomenon gracile 219	Q934E5
	Nodularia harveyana BECID29	Q5K2U4
	Trichodesmium erythraeum IMS101	Q10WH6
	Tychonema bourrellyi	O88105
	Gloeobacter violaceus PCC7421	Q7NIM7
	Arabidopsis thaliana	O03042
	Rhodospirillum rubrum	Q2RRP5
RbcS	Synechococcus sp. PCC7002 (Agmenellum quadruplicatum)	Q44178
	Synechococcus elongatus PCC6301 (Anacystis nidulans)	P04716
RbcX	Synechococcus sp. PCC7002 (Agmenellum quadruplicatum)	Q44177
	Synechococcus elongatus PCC6301 (Anacystis nidulans)	Q5MZ09
	Crocospaera watsonii	Q4C1R0
	Thermosynechococcus elongatus	Q8DIS6
	Microcystis aeruginosa NIVA-CYA 143	Q9R3Q3
	Synechocystis sp. PCC6803	Q55670
	Anabaena sp. PCC7120	Q44307
	Anabaena sp. CA	Q44212
	Aphanizomenon gracile 219	Q933W0
	Nodularia harveyana BECID29	Q5K2U3
	Planktothrix agardhii NIVA-CYA 127	Q61833
	Tychonema bourrellyi	O86944
	Gloeobacter violaceus PCC7421	Q7NIM6
	Arabidopsis thaliana, gene I	Q8L9X2
	Arabidopsis thaliana, gene II	Q94AU9
	Oryza sativa (japonica cultivar-group), gene I	Q84M38
	Oryza sativa (japonica cultivar-group), gene II	Q7XIH6
	Oryza sativa (japonica cultivar-group), gene III	Q7EZV3
	Zea mays (*)	EE045458
	Physcomitrella patens, gene I (**)	PPP_6613_C1
Physcomitrella patens, gene II (**)	PPP_11273_C1	
Chlamydomonas reinhardtii (***)	AV621347	

**Table A5. Protein accession numbers.**

Accession numbers for SwissProt. Only indicated accession numbers are valid for (\*) <http://www.maizegdb.org>, (\*\*) <http://www.cosmoss.org>, (\*\*\*) <http://est.kazusa.or.jp>.

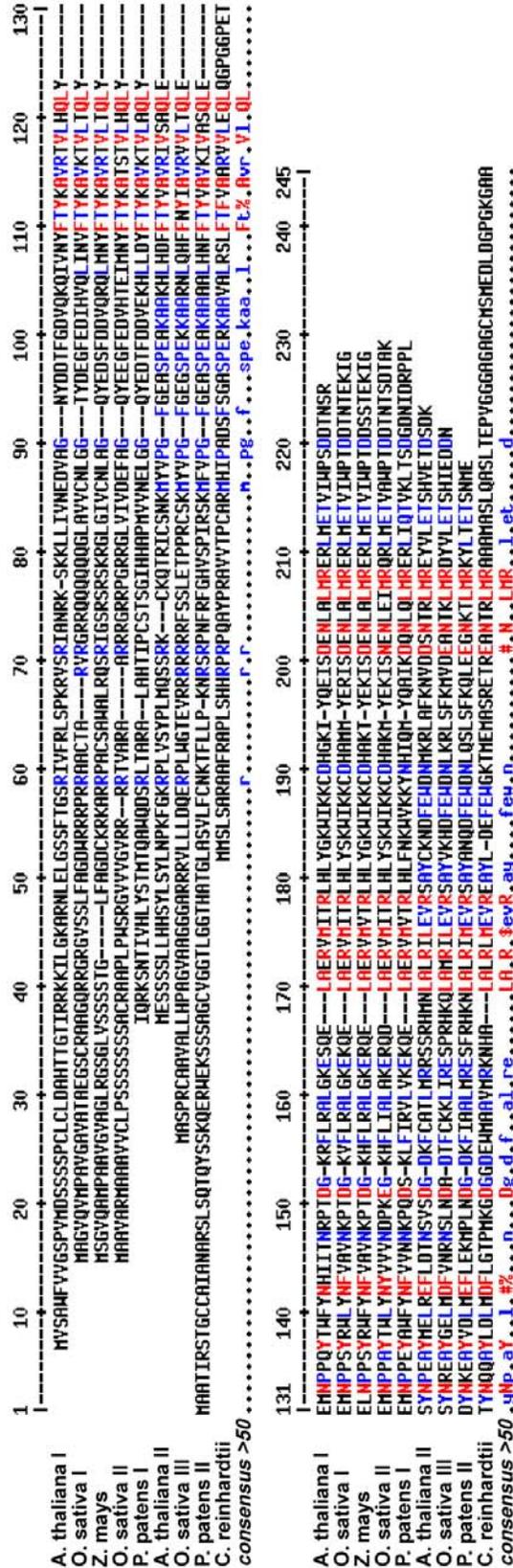
## 7.6 Crystallographic data

dataset	PtC14-2	native I	Scmet	native II	peptide cplx.	Y17AY20L	Q29 A	R70A	Anabaena
beamline	ESRF, BM14	SLS, X10SA	ESRF, ID14-4	ESRF, ID29	ESRF, ID29	ESRF, ID29	ESRF, ID29	ESRF, ID29	SLS, X10SA
wavelength (Å)	0.8793	0.9789	0.9795	1.000	0.976	1.000	1.000	1.000	1.072
space group	P4 <sub>2</sub> ,2	P4 <sub>2</sub> ,2	P1	I222	P4 <sub>2</sub> ,2	P4 <sub>2</sub> ,2	P4 <sub>2</sub> ,2	P4 <sub>2</sub> ,2	P6 <sub>5</sub>
cell dimensions, a, b, c (Å); α, β, γ (°)	93.42, 93.42, 407.75; 90, 90, 90	93.33, 93.33, 411.99; 90, 90, 90	70.36, 71.10, 113.16; 89.58, 81.10, 80.09	63.50, 69.34, 98.79 90, 90, 90	93.21, 93.21, 412.02; 90, 90, 90	93.45, 93.45, 411.54; 90, 90, 90	93.00, 93.00, 413.72; 90, 90, 90	93.54, 93.54, 413.88; 90, 90, 90	69.02, 69.02, 113.81; 90, 90, 120
resolution limits (Å)*	93.25 - 3.15 (3.32 - 3.15)	91.29 - 2.8 (2.95 - 2.8)	68.52 - 2.7 (2.85 - 2.7)	29.24 - 1.9 (2.0 - 1.9)	93.25 - 2.95 (3.11 - 2.95)	93.45 - 3.4 (3.58 - 3.4)	92.85 - 3.1 (3.27 - 3.1)	41.00 - 6.1 (6.43 - 6.1)	59.77 - 2.4 (2.53 - 2.4)
Rmerge **	0.081 (0.621)	0.080 (0.641)	0.101 (0.420)	0.089 (0.575)	0.080 (0.536)	0.089 (0.503)	0.077 (0.485)	0.107 (0.408)	0.078 (0.444)
I/sigma	15.9 (3.1)	26.7 (4.6)	11.6 (2.7)	11.5 (2.2)	11.7 (2.5)	9.5 (2.4)	10.6 (2.9)	6.7 (1.8)	10.7 (2.0)
multiplicity	7.0 (7.1)	14.5 (14.8)	3.9 (3.8)	4.2 (4.3)	3.7 (3.8)	3.4 (3.6)	3.5 (3.6)	2.0 (1.9)	3.6 (2.9)
completeness (%)	100 (100)	100 (100)	96.9 (96.3)	100 (100)	96.6 (97.6)	99.6 (99.8)	99.9 (99.9)	75.9 (75.9)	98.1 (88.2)
RbcX monomers /a.u.	6	6	12	2	6, 1 peptide	6	6	2	2
solvent content (%)	-	74.7	65.9	30.7	-	-	-	50.9	50.9
Wilson B-factor (Å <sup>2</sup> )	-	-	49.60	24.10	-	-	-	41.5	41.5
<b>Phasing</b>									
sites	6 (Pt)	-	37 (Se)	-	-	-	-	-	-
mean FoM	0.31 (SIRAS)	-	0.322 (SAD)	-	-	-	-	-	-
<b>Refinement</b>									
resolution range	-	20 - 2.8	20 - 2.7	20 - 1.9	20 - 2.95	20 - 3.4	20 - 3.1	20 - 2.5	20 - 2.5
reflections (test set)	-	43747 (2321)	53770 (2881)	16664 (887)	35788 (1913)	24596 (1319)	32237 (1726)	10128 (502)	10128 (502)
Rwork	-	0.240	0.212	0.197	0.235	0.253	0.231	0.241	0.241
Rfree	-	0.262	0.262	0.253	0.264	0.294	0.260	0.298	0.298
number of atoms	-	5105	10068	1893	5235	4727	4976	1769	1769
r.m.s.d. bonds (Å)	-	0.011	0.020	0.013	0.012	0.010	0.010	0.012	0.012
r.m.s.d. angles (°)	-	1.406	1.683	1.390	1.388	1.319	1.273	1.448	1.448
Ramachandran plot (%) ***									
% most favored region	-	93.0	91.6	95.0	93.5	87.8	93.9	93.6	93.6
% additionally allowed	-	6.8	8.3	5.0	6.5	11.9	6.1	5.9	5.9

\* values in parenthesis for outer shell; \*\* as defined in Scala; \*\*\* as defined in Procheck (Laskowski *et al.* 1993)

**Table A6. Crystallographic data collection, phasing and refinement statistics (Dr. Andreas Bracher)**

### 7.7 Alignment of RbcX from green alga, moss and higher plants



**Figure A1. Alignment of amino acid sequences of selected RbcX homologs.** Amino acid sequences of higher plants (*Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*), moss (*Physcomitrella patens*) and green alga (*Chlamydomonas reinhardtii*) were aligned using the MultAlin analysis program (Corpet 1988). High consensus level ( $\geq 90\%$ ) is depicted in red and low consensus level ( $\geq 50\%$ ) in blue. The protein accession numbers are listed in table A4.

## 7.8 Abbreviations

Units are expressed according to the international system of units (SI). One and three letter codes for amino acids are presented in table A4.

AA	acrylamide
aa	amino acid
ADP	adenosine 5'-diphosphate
Amp	ampicillin
Amp <sup>R</sup>	ampicillin resistance
AMP-PNP	adenosine 5'-( $\beta,\gamma$ -imido)triphosphate
Ana	<i>Anabaena</i>
ANS	1-anilino-8-naphthalene-sulphonate
APS	ammonium peroxodisulfate
At	<i>Arabidopsis thaliana</i>
ATP	adenosine 5'-triphosphate
BLAST	Basic Local Alignment Search Tool
$\beta$ -NADH	$\beta$ -nicotinamide adenine dinucleotide
bp	base pairs
BSA	bovine serum albumin
$^{\circ}\text{C}$	degree Celsius
$^{14}\text{C}$	carbon-14
Cam	chloramphenicol
Cam <sup>R</sup>	chloramphenicol resistance
CAPB	carboxy-arabinitol-1,5-bisphosphate
cDNA	copy DNA
CDTA	<i>trans</i> -1,2-diaminocyclohexane- <i>N,N,N',N'</i> -tetraacetic acid
ch	chloroplast
Cpn	chaperonin
CTAB	cetyltrimethylammoniumbromid
$\Delta$ (delta)	deletion
DE	diethylaminoethyl
DNA	deoxyribonucleic acid
DnaJ	bacterial Hsp40 chaperone
DnaK	bacterial Hsp70 chaperone
dNTP	2'-desoxyribonucleotide-triphosphate
DSP	dithiobis[succinimidylpropionate]
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
ESI	electro spray ionization
FFF	field flow fractionation
Fig.	Figure
fmet-tRNA	<i>N</i> -formyl-methionyl-tRNA
FPLC	fast protein liquid chromatography
FT	flowthrough

g	acceleration of gravity, 9.81 m/s <sup>2</sup>
GdnHCl	guanidinium hydrochloride
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
His <sub>6</sub>	histitin-tag
HRP	horseradish peroxidase
Hsp	heat shock protein
IPTG	isopropyl-β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
Kan	kanamycin
Kan <sup>R</sup>	kanamycin
K <sub>D</sub>	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
NTA	nitrioloaminetriacetic acid
OAc	acetate
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
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
rpm	revolutions per minute
Rr	<i>Rhodospirillum rubrum</i>

RT	room temperature
RTS	rapid transcription/translation system
<sup>35</sup> S	sulphur-35
SDS	sodiumdodecylsulfate
SeMet	selenomethionine
So	<i>Spinacia oleracea</i>
SeMet	Selenomethionine
sp.	species
Syn	<i>Synechococcus</i>
TCA	trichloroacetic acid
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
TEV	<i>Tobacco etch virus</i>
TF	trigger factor
Tris	trishydroxymethylaminomethan
tRNA	transfer-RNA
w/o	without
v/v	volume per volume
w/v	weight per volume

## 7.9 Publications

Effmert, U., **Saschenbrecker, S.**, Ross, J., Negre, F., Fraser, C.M., Noel, J.P., Dudareva, N., Piechulla, B. (2005). Floral benzenoid carboxyl methyltransferases: from *in vitro* to *in planta* function. *Phytochemistry* 66, 1211-1230. Review.

Pott, M.B., Hippauf, F., **Saschenbrecker, S.**, Chen, F., Ross, J., Kiefer, I., Slusarenko, A., Noel, J.P., Pichersky, E., Effmert, U., Piechulla, B. (2004). Biochemical and structural characterization of benzenoid carboxyl methyltransferases involved in floral scent production in *Stephanotis floribunda* and *Nicotiana suaveolens*. *Plant Physiol.* 135, 1946-1955.

**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.

## Lebenslauf

### Persönliche Daten

Sandra Saschenbrecker  
geboren am 19. Juli 1979 in Schwerin

### Wissenschaftlicher Werdegang

#### Hochschulreife

Juli 1998                      Abitur, Goethe-Gymnasium Schwerin

#### Hochschulstudium

Okt. 1998 – März 2003      Studium der Biologie and der Universität zu Rostock

Sept. 2000                      Diplomvorprüfung  
Fächer: Botanik, Zoologie, Mikrobiologie, Genetik,  
Physik, Chemie

Aug. 2001 – Sept. 2001      Forschungspraktikum in der Klinik für Chirurgie Rostock,  
unter Anleitung von PD Dr. Marian Löbler: *In situ*  
*Hybridisierung von mRNA an humanem Pankreasgewebe*

Okt. 2001 – Jan. 2002      Forschungspraktikum am Institut für Biowissenschaften an  
der Universität zu Rostock, unter Anleitung von Prof. Dr.  
Birgit Piechulla: *Überexpression der SAMT aus*  
*Stephanotis floribunda in der Hefe Pichia pastoris*

April 2002                      Diplomhauptprüfung  
Fächer: Biochemie (Hauptfach), Molekularbiologie,  
Mikrobiologie, Biotechnologie

Juli 2002 – März 2003      Diplomarbeit am Institut für Biowissenschaften an der  
Universität zu Rostock, unter Anleitung von Prof. Dr.  
Birgit Piechulla: *Charakterisierung der Genexpression*  
*eines BAMT-kodierenden Gens in Nicotiana suaveolens*

März 2003                      Diplom in Biologie

#### Dissertation

Juni 2003 – heute              Experimentelle Doktorarbeit in der Abteilung Zelluläre  
Biochemie von Prof. Dr. F. Ulrich Hartl am Max-Planck-  
Institut für Biochemie, Martinsried: *Folding and assembly*  
*of RuBisCO / Structural and functional characterization*  
*of the RuBisCO assembly chaperone RbcX*