Structure of green type Rubisco activase from Nicotiana tabacum

Mathias Michael Stotz

München 2012

Dissertation zur Erlangung des Doktorgrades

an der Fakultät Chemie und Pharmazie

der Ludwig–Maximilians–Universität

München

Structure of green type Rubisco activase from Nicotiana tabacum

---

vorgelegt von

Mathias Michael Stotz

aus Tübingen, Deutschland

München, 2012

Rule I

"*We are to admit no more causes of natural things, than such as are both true and sufficient to explain their appearances.*

   To this purpose the philosophers say, that Nature does nothing in vain, and more is in vain, when less will serve; for Nature is pleased with simplicity, and affects not the pomp of superfluous causes."

Sir Isaac Newton, The Mathematical Principles of Natural Philosophy

Acknowledgement

Contents

List of abbreviations

Δ . . . . Deletion

(w/v) . . . . Weight/volume

E. coli . . . . *Escherichia coli*

N. tabacum . . . . *Nicotiana tabacum*

AAA(+) . . . . ATPases associated with various cellular activities

ADP . . . . Adenosine 5'-diphosphate

AMP-PNP . . . . Adenosine 5'-(β,γ-imido)triphosphate

APS . . . . Ammoniumpersulfate

ATP . . . . Adenosine 5'-triphosphate

ATPγS . . . . Adenosine 5'-γ-thiotriphosphate

CA1P 2-carboxy-arabinitol 1-phosphate

CABP Carboxy-arabinitol-1,5-bisphosphate

CBB . . . . Calvin Benson Bassham

CCM] . . . . carbon concentrating mechanism

CD . . . . Circular dichroism

CIP . . . . Calf intestinal phosphatase

dA . . . . Desoxyadenosine

DNA . . . . Deoxyribonucleic acid

DSS . . . . Disuccinimidyl suberate

dT . . . . Desoxythymidine

DTT . . . . Dithiothreitol

ECL . . . . Enhanced chemoluminescence

EDTA . . . . Ethylenediaminetetraacetic acid

EM . . . . Electron microscopy

HEPES . . . . N-(2-hydroxyethyl)piperacin-N'-2-ethanesulfonic acid

His$_6$Ub .... Hexa-histidine tagged ubiquitin

HRP .... Horseradish peroxidase

Hsp .... Heat shock protein

IMAC .... Immobilised Metal Ion affinity chromatography

IPTG .... Isopropyl β-D-thiogalactopyranoside

ITC .... Isothermal titration calorimetry

LB .... Lysogenic broth media

MW .... Molecular weight

NAD$^+$ / NADH ...... Nicotinamide adenine dinucleotide oxidized / reduced

NADP / NADPH ..... Nicotinamide adenine dinucleotide phosphate oxidized / reduced

Ni-NTA .... Nickel-Nitrilotriacetic acid

NtRca .... *Nicotiana tabacum* Rubisco activase

OD .... Optical densitiy

PMSF .... Phenylmethylsulfonylfluoride

RbcL .... Rubisco large subunit

RbcS .... Rubisco small subunit

Rca .... Rubisco activase

RT .... Room temperature

Rubisco .... Ribulose 1,5-bisphosphate carboxylase/oxygenase

RuBP .... Ribulose 1,5-bisphosphate

SDS-PAGE .... Sodiumdodecylsulfate polyacrylamid gelelectrophoresis

SEC .... Size exclusion chromatography

SeMet .... Selenomethionine

SIRAS .... Single isomorphous replacement with anomalous scattering

TEMED .... N,N,N',N'-tetramethylethylenediamine

Tris .... Tris(hydroxymethyl)aminomethane

1 Summary

Ribulose 1,5-bisphosphate Carboxylase/Oxygenase (Rubisco) is a key enzyme of photosynthesis, it catalyses the carboxylation of ribulose 1,5-bisphosphate (RuBP), which is the main pathway for the fixation of atmospheric, inorganic CO_2 into organic biomass. Despite its central role for plant metabolism, the enzyme suffers from several shortcomings that necessitate its abundance in nature. Compared to other metabolic enzymes the catalytic rate of plant Rubisco is rather slow (~2-3 s^{-1}). Furthermore, a prominent side reaction is the oxygenation of the substrate RuBP. Importantly, Rubisco is also prone to inhibition under physiological conditions.

A lysine residue in the active site pocket of Rubisco must be carbamylated and bind a Mg^{2+}ion as cofactor for the enzyme to become catalytically active. Binding of RuBP to uncarbamylated Rubisco, or of 2-carboxy-arabinitol 1-phosphate (CA1P) to the carbamylated enzyme at night, results in the trapping of the sugar phosphate and inhibition of the enzyme. Inhibition of Rubisco provides a mechanism for the regulation of Rubisco activity in response to varying environmental conditions. The inhibited Rubisco complexes are stable and cannot reactivate spontaneously under physiological conditions. In plants, inhibited Rubisco must be reactivated in an ATP dependent manner by the protein Rubisco activase (Rca) an ATPase associated with various cellular activities (AAA+).

Rca catalyses the ATP-dependent remodeling of Rubisco and facilitates the release of the bound sugar phosphate but is not actively involved in carbamylation of the active site. Since Rca is essential for plant growth and required for maintaining Rubisco activation in vivo, it has been dubbed the "catalytic chaperone" of Rubisco. Interestingly, Rca forms polydisperse, heterogeneous oligomers in solution which vary in size between 50 kDa and 600 kDa. Additionally, Rca is heat labile and it has been proposed that it is a limiting factor for the photosynthetic potential of plants under moderate heat stress, which might become increasingly problematic as the Earth's temperature increases due to climate change. However, in the absence of a structure, the oligomeric state of the active enzyme as well as a mechanism for its Rubisco activation function remained elusive.

In the present study, the structure, the oligomeric state and the mechanism of Rubisco activase were

investigated. The crystal structure of an N- and C-terminally truncated Rca construct from *Nicotiana tabacum* was solved at a resolution of 2.95 Å by seleno-methionine single isomorphous replacement with an anomalous scatterer (SIRAS). The structure of the truncated subunit shows the classical AAA+ architecture consisting of a N-terminal nucleotide binding domain and a C-terminal 4-helix bundle. Whereas the nucleotide binding domain is folded in a non-classical Rossmann-fold typical for AAA+ proteins, the 4-helix bundle domain differs from the classical AAA+ fold topology. It features a small helical insertion, containing critical amino acid residues required for the recognition of the substrate Rubisco.

Guided by the structure key features of Rca from *Nicotiana tabacum* were targeted for mutation. Specifically, amino acid changes were introduced in the putative Rca oligomer interface and in the putative pore loop residues involved in the Rubisco remodeling. Mutations in the putative subunit-subunit oligomerisation interface generally disrupted oligomer formation of Rca in the nucleotide free state, but also identified an arginine to alanine mutation that lead to formation of stable hexamers in the nucleotide bound state while preserving wildtype activity. Negative stain electron microscopy of this mutant lead to a low resolution model of Rca in the hexameric state, which presumably reflects the active oligomer of the enzyme. Additional density was observed on the "top" side of the hexamer, which presumably corresponds to the N-terminal domain absent in the crystallised constructs. In analogy to other AAA+ domains, it is hypothesized that the N-terminal domain is acting as a substrate adaptor to Rubisco, consistent with reports of deletions and mutations in this domain. Mutation of amino acids in the loops facing the central pore in the hexameric state of Rca partially retained ATPase activity but showed greatly diminished Rubisco activation activity, indicating a role for the pore loops in Rubisco activation.

The combined structural as well as biochemical data presented in this thesis provides a framework for the detailed mechanistic analysis of Rubisco activation in plants. The data suggests that the active state of Rca is a hexamer. Initial Rubisco recognition is mediated by helix H9 and the N-terminal domain and remodeling of Rubisco might involve partial threading of Rubisco through the central pore by the pore loops.

2 Introduction

Proteins are abundant biological macromolecules made up of amino acids and named after the Greek word "proteios" which means elementary. They are found ubiquitously in all domains of life and are essential for almost all biological functions. Proteins act as biological "machines" serving diverse roles, ranging from structure forming assemblies to signaling to catalytic activities. The function of a protein is determined by its structure, which ultimately is encoded in the sequence of the 20 proteinogenic amino acids forming the unbranched protein chains. The elucidation of protein structures is therefore a method to investigate the biological function and mechanism of a protein.

2.1 Protein folding

Since the structure of a protein defines its function, proper folding of the newly-synthesized, disordered chain into its native, three dimensional structure is essential. The information for the structure of the native state is encoded in the primary sequence of a protein, as demonstrated by Christian Anfinsen in the early 1960's [1][2][3][4] through refolding experiments with Ribonuclease A. The main driving force of protein folding in solution is the burial of hydrophobic side chains in the interior of the protein and the entropy gain of solvent molecules (water) released during this process. Other stabilizing forces are salt bridges between opposite charges, hydrogen bonds and van der Waals contacts between atoms. The native state of a protein is therefore energetically lower than the unfolded state and usually represents the most stable state of the system [4]. However, since proteins usually require some degree of flexibility for their biological function, the native state is only 10-15 kcal/mol more stable than the unfolded state. Thus the temperature range for protein stability is relatively narrow[5].

Since random sampling of all possible conformations to reach the native state requires a timescale orders of magnitudes longer than measured refolding times of proteins in vitro, a pathway on which proteins fold must exist [6]. This can be illustrated by the multidimensional energy landscape of protein folding, called the folding "funnel" in which the coordinates of a conformational state (usually expressed

as the dihedral angles between the freely rotating bonds in a peptide) are plotted against the energy of that conformation. Figure 2.1 shows a simplified, two dimensional folding energy landscape. The unfolded protein U reaches the global energy minimum of the native state N on a certain pathway. Note that energy landscapes can be "rugged" and therefore have local energy minima that are associated with kinetically trapped intermediates. The "ruggedness" of the energy landscape determines the speed of folding, since more local minima lead to more kinetically trapped intermediates and therefore slower folding. While small proteins may fold very fast (timescale of μs), more complicated architectures often have a very rugged folding energy landscape and therefore fold very slow (timescale of min to h) or even fail to reach their native state in vitro. In these cases, partially folded intermediates often still display hydrophobic patches on the surface, by which they can aberrantly oligomerize and aggregate [7]. Furthermore, the intracellular environment is highly crowded, reaching protein concentrations of 300-400 g/l [8]. The occluded volume effect, while stabilizing native protein-protein interactions, also facilitates aggregation of partially unstructured protein states [8] [9] and even the formation of ordered amyloid aggregates, which are often associated with toxicity.
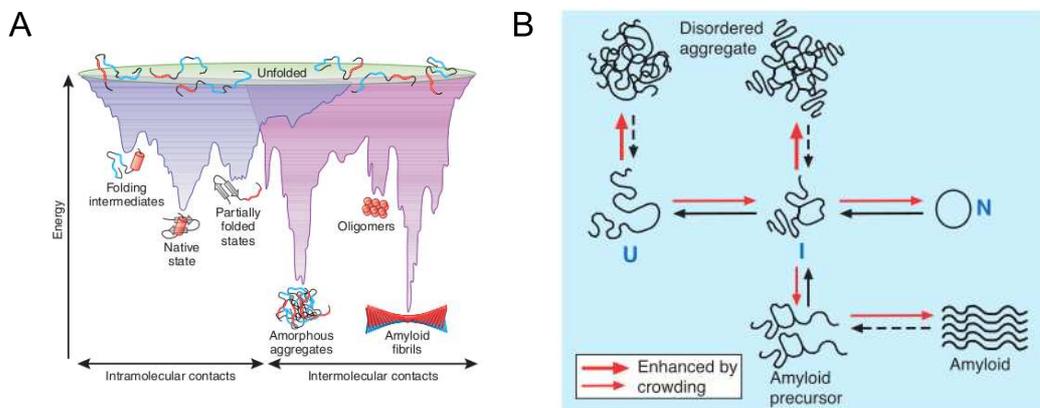
Figure 2.1: **Folding and aggregation pathways in the cell.**
A: Simplified two-dimensional energy landscape for a protein. Local minima correspond to kinetically trapped intermediate states and slow down the folding process. Aberrant intermolecular interactions can lead to formation of disordered or ordered (amyloid) aggregates. B: Folding and aggregation pathways in the crowded environment in the cell. The unfolded protein U can reach the native state N via the intermediate I, however, aggregation and amyloid formation are competing side reactions. Adapted from [10] and [11].

2.1.1 Molecular chaperones

Large multi-domain proteins often require molecular chaperones to fold in a biologically relevant timescale and to prevent potentially toxic aggregation (as recently reviewed in [10] [5] [7]). Since mutations in a protein can destabilise the native structure or influence the folding pathway, molecular chaperones can also act as an evolutionary buffer allowing for increased genetic variation (demonstrated for Hsp90 [12] as well as GroEL/ES [13]). The observation that chaperones can also facilitate the degradation of terminally misfolded client proteins (reviewed in [14]), led to an emerging role of molecular chaperones in protein quality control and homeostasis. Since a decline of protein homeostasis during aging has been proposed and aggregation-associated (neuro-) degenerative diseases are correlated with aging, regulation and mechanism of molecular chaperones are a field of intensive research [15] [16].

All molecules that interact with or assist proteins in acquiring their functional, native state without being present in the final structure are defined as molecular chaperones (as originally proposed by Ellis [17]). Proteins of this family of molecules are also often classified as heat shock proteins (Hsp) due to their increased expression after heat stress. Different conserved classes and mechanisms of molecular chaperones have been identified and are named after their respective molecular weights (small Hsp, Hsp40, Hsp60, Hsp70, Hsp90, Hsp100 and Hsp110). While small Hsp act as ATP independent, holdases that buffer aggregation, the Hsp40-Hsp70, the HSP60, the Hsp90, the Hsp100 and Hsp110 system require ATP for their specific functions (reviewed in [5], [10] and [7]). The Hsp100 chaperones belong to the ATPase associated with various cellular activities (AAA+) protein family (see 2.2) and actively remodel and unfold their substrates. They also actively disaggregate misfolded proteins in cooperation with the Hsp70 system [18] [5]. In the following sections the Hsp60/chaperonin system and the Hsp70 system are briefly discussed as examples for the complex mechanisms of molecular chaperones. The Hsp100 molecular chaperones are discussed in more detail in context of the AAA+ proteins.

2.1.1.1 The chaperonin system

The chaperonins are double-ring complexes with a molecular weight of 800-900 kDa. They function by ATPase regulated encapsulation of substrates of up to 60 kDa in the inner cavity formed by the rings. The chaperonins are classified into two groups (reviewed in [10]).

Group I chaperonins found in bacteria and organelles of bacterial origin in eukaryotes contain 14 subunits (GroEL, Hsp60, Cpn60) each in a back-to-back, staggered, heptameric double-ring and require a

cochaperone (Hsp10, Cpn10, Cpn20) for encapsulation of their substrate. The best studied family member is the bacterial chaperonin GroEL and its cochaperone GroES. GroEL is composed of 14 identical ~57 kDa subunits, arranged in two staggered heptameric rings. Each subunit consists of an equatorial domain harboring the ATP hydrolysis site, a intermediate domain linker and an apical domain [19] [20]. GroEL binds substrates by exposed hydrophobic residues in the apical domain of one ring. Subsequently, the ATP-regulated binding of the homoheptameric GroES triggers a conformational change and the release of the substrate into the inner hydrophilic cavity. The substrate is encapsulated for one ATP hydrolysis cycle (~10 s) before ATP binding to the opposite ring triggers the release of GroES and the substrate [21] [22] [23] [24] [25]. Partially folded substrates rapidly rebind GroEL if a single cycle of folding on the chaperonin is insufficient to reach the native state. The complete encapsulation of the substrate during the cycle prevents aberrant aggregation, which has led to the hypothesis of GroEL acting like an "Anfinsen cage", practically allowing the substrates to fold in infinite dilution [26] [27]. The recent finding that GroEL assisted folding can be faster than spontaneous refolding (in the absence of aggregation) challenged this conclusion [28] [29] [30]. Models have been proposed that explain this phenomenon by steric confinement and entropy reduction of the protein intermediates in the cage [30]. Interestingly, most of the ~250 identified protein substrates of GroEL are 20-50 kDa in size and have a complex α/β and $\alpha+\beta$ topology consistent with the encapsulation mechanism being required for folding of proteins with complex domain topologies [31].

Group II chaperonins are found in the cytosol of eukaryotes and contain two eight or nine membered heterooligomeric rings stacked directly back-to-back. Their apical domains contain an insertion that enables them to encapsulate their substrate without an additional cochaperone like GroES [32]. Among the most studied group II chaperonins are the archeal Thermosome and the eukaryotic TRiC/CCT. Whereas the archeal Thermosomes consist of a maximum of three homologous subunits arranged in a hexadecamer (e.g. Thermosome of *Thermoplasma acidophilum* [33]) or octadecamer (TF55 [34]), the TRiC/CCT consists of eight paralogous subunits in each ring [35]. Contrary to initial proposals that TRiC is specialised for the folding of only a small subset of cytosceletal, eukaryotic proteins [36] , it was recently reported that this chaperonin interacts with a broad range of polypeptides that function in many cellular processes [37]. Due to its slower ATPase rate, TRiC might have evolved to give these substrates a longer time to fold in encapsulation [38].

2.1.1.2 The Hsp70 system

The evolutionary conserved Hsp70 system does not rely on encapsulation of the substrates but instead binds and releases its substrates in an ATP dependent manner by recognizing to a hydrophobic peptide motive [39] [5]. Hsp70 is comprised of a N-terminal nucleotide binding domain (NBD) with a hexokinase like fold [40] and a C-terminal substrate binding domain (SBD) consisting of a β-sandwich domain with an α-helical lid [41]. Allosteric regulation of the binding affinity of the SBD by the nucleotide state of the NBD leads to sequential binding, holding and release of the substrate. Hsp70 acts in concert with Hsp40 and nucleotide exchange factors to accomplish this [5]. In the ATP state, the SBD and the NBD interact and the binding affinity to the substrate is low. Rapid on/off kinetics can be observed in this state. Although the structure of the complex of NBD and SBD in the ATP bound state is unknown, it is likely similar to the domain arrangement in the solved structure of Hsp110 [42] [43] [44]. Binding of Hsp40 and substrates stimulates the intrinsically weak ATPase rate. In the ADP state the NBD and SBD do not interact and display characteristics of the isolated domains in solution [45] [46]. Binding affinity for the substrate is high accompanied by slow on/off kinetics. Exchange of ADP to ATP catalysed by nucleotide exchange factors restarts the cycle and allows the substrate to be released.

2.2 AAA+ proteins

ATPases associated with various cellular activities (AAA) proteins are a subset of Walker (P-loop) ATPases [47] [48]. Originally defined as a protein family with a distinct set of conserved sequence features (Walker A, Walker B), the nomenclature has been extended and named AAA+ to encompass all structurally similar AAA proteins [49]. The two nomenclatures, AAA and AAA+, are not strictly distinguished in the literature and for simplicity examples of both groups will be discussed here as AAA+ proteins.

AAA+ proteins are found in all kingdoms of life and have a widely diverse set of functions ranging from motor proteins (Dynein) to DNA helicases (RuvB, SV40 helicase), to vesicle trafficking (Vps4, NSF) to protein unfolding (Clp, HslU) [52]. Their domain composition is usually very similar, however. Class I AAA+ proteins contain two AAA+ modules consisting of a nucleotide binding domain (NBD) and a 4-helix bundle domain in tandem, while the class II proteins only contain one module [53] [54]. Interestingly, it has been shown that the ATPase activity of one NBD of some class I AAA+ proteins is dispensable for the activity and it has been speculated that these NBDs are required for proper assembly
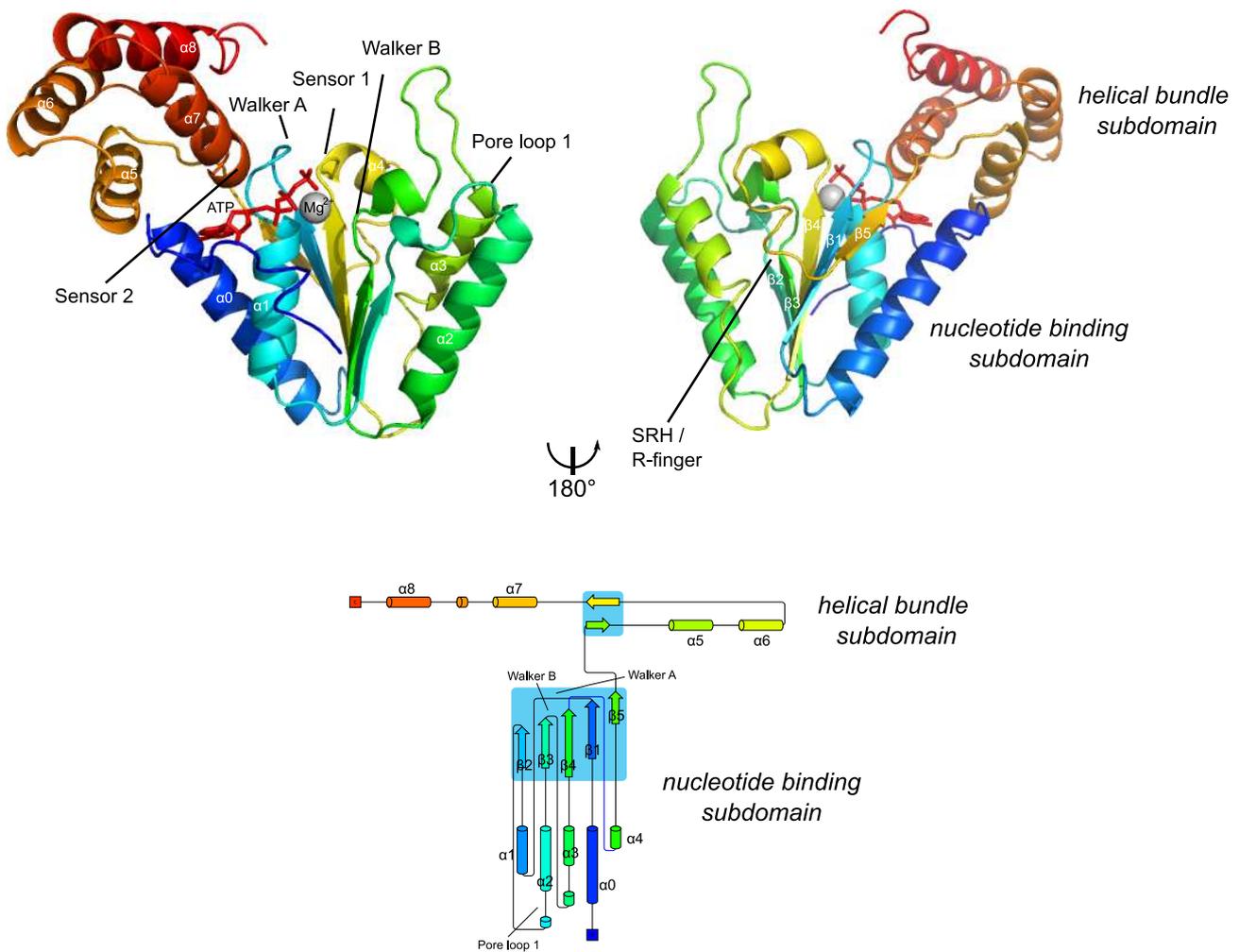
Figure 2.2: **Structure and topology of the D2 AAA+ module of the N-ethylmaleimide sensitive fusion protein (NSF).** (1D2N [50])

Structure and sequence features defining the AAA+ protein family are indicated. The structure is shown in a ribbon representation and colored in rainbow colors from N-terminus (blue) to C-terminus (red). ATP binding site (bound AMP-PNP) and Mg^{2+} are indicated in stick and sphere representations, respectively. Secondary structure elements are indicated. The topology of NSF corresponds to a representative AAA+ topology. The topology diagram of 1D2N was generated with Pro-origami [51]. α-Helices are represented as tubes, β-strands as arrows.

of the complex (e.g. NSF [50], p97 [55]).

In general, AAA+ proteins form homohexameric arrangements with a central pore, stabilised by the presence of nucleotide. The AAA+ modules of class I proteins therefore form two, stacked hexameric rings [55] [56], while class II proteins form single hexameric rings [57] [52].

Figure 2.2 shows the D2 AAA+ module of NSF as an example for a AAA+ structure. The nucleotide

binding domain features a central 5 stranded β-sheet in a Rosssmann fold-like, conserved topology with an inserted β-strand 4 between β1 and β3. The Walker A, Walker B and Sensor I sequence elements encode loops that are essential for nucleotide binding (Walker A) and hydrolysis (Walker B and Sensor I) within the NBD. Interestingly, the NBD usually also contains conserved arginine residues after α4 that act in trans in the assembled oligomer to form the nucleotide binding site, the so called arginine fingers. The arginine finger is proposed to facilitate allosteric communication and sense the nucleotide binding state of the adjacent subunit (reviewed in [58]). Another essential sequence element for the biological function of many AAA+ proteins is the pore loop sequence (Figure 2.2). A conserved sequence element after β2 faces the central pore and defines a loop. The conserved aromatic-hydrophobic-glycine motif is essential for substrate interactions but not for ATP binding or hydrolysis and is proposed to thread the substrate through the central pore [53]. It seems evident from studies on ClpX (protein unfoldase in proteolysis), RFC (clamp loader), Orc1 (DNA remodeler for replication initiation) as well as other AAA+ proteins that interactions with the substrate frequently regulate the ATPase rate of AAA+ proteins. This is often accomplished by a hydrogen bond between a conserved glutamate in the Walker B to an asparagine or serine/threonine at the C-terminus of β4, that inhibits ATP hydrolysis. Upon substrate binding a conformational change is predicted to break this hydrogen bond stimulating ATP hydrolysis [59][60].

The α-helical domain is structurally more diverse, but its core topology is a 4-helix-bundle domain. A conserved sequence motif called the Sensor II forms the N-terminal part of α7 and contains a conserved arginine necessary for ATP binding and hydrolysis [54].

AAA+ proteins are subclassified into clades based on their unique sequence and structural motifs [61] [54] . In general 7 clades can be distinguished: clamp loader clade, initiator clade, classical clade, helicase clade, HCLR clade, H2-insert clade and PS-II insert clade.

One of these clades, the so called classical clade comprises most of the initially identified AAA+ proteins [49] for example p97, Pex1, Pex6, the proteasomal AAA+ components, Katanin, Spastin and also Rubisco activase. Classical clade AAA+ proteins are structurally distinguished by a unique insertion of a small α-helical element between β2 and α2 in the NBD [61] [54]. Unique sequence features, such as the absence of an arginine in the Sensor 2 motif as well as two arginines in position for the arginine finger, further distinguish them from other AAA+ clades. Members of this class seem to be functionally unified by a protein remodeling activity, however, in different biological functions such as microtubule severing (Katanin, Spastin), peroxisome biogenesis (Pex1/Pex6) or proteolysis (proteasomal AAA+). An emerging common motive for these activities seems the mechanical remodeling of substrate proteins by

ATP-dependent conformational changes, mediated in part by the pore loop residues and the class defin-ing helical insertion [62] [53]. ATP hydrolysis is considered to mediate intersubunit rearrangements of the NBD and the C-terminal 4-helix bundle, that are in turn transmitted to other subunits in the hexam-eric complex by conserved contacts such as the Arginine fingers [63] [55]. Functional diversity seems to be mainly achieved by N-terminal domains outside the AAA+ module that often regulate substrate specificity and affinity. Indeed, Spastin recognizes the substrate tubulin by its N-terminal extension [57] and the classical clade family member p97 (Cdc48) functions in diverse biological processes such as the ERAD pathway or membrane fusion events of the endoplasmic reticulum and Golgi apparatus by recruiting several different adaptor proteins such as p47 (Shp1p) [64], p37 [65] or Ufd1/Npl4 [66] via its N-terminal domain.

2.2.1 AAA+ chaperones of the Hsp100 family

Unlike most other chaperones, the Hsp100 class of chaperones, not only prevent aggregation, they ac-tively unfold their substrates and even disaggregate misfolded proteins in cooperation with the Hsp70 system [18] [5]. Members of this class of chaperone comprise the Clp proteins in the bacterial cytosol and organelles of bacterial origin as well as the Hsp104 in the cytosol of lower eukaryotes and Hsp101 in plants. Interestingly, no homologue of Hsp100 has been identified in the cytosol of metazoa so far [5].

All Hsp100 proteins belong to the AAA+ protein class (see 2.2) and therefore share a common domain topology. Both class I (with two nucleotide binding domains; e.g. yeast Hsp104, bacterial ClpB) and class II (with one nucleotide binding domain; e.g. bacterial ClpX) types of AAA+ proteins are found in the Hsp100 chaperones [53]. The Hsp100 class proteins form homo-hexameric ring assemblies in the active state [52]. The bacterial Hsp100s (ClpA, ClpC, ClpX, HslU) are, furthermore, often associated with a protease (ClpP, HslV), thereby mediating unfolding of substrates for degradation.

The bacterial class II Hsp100, ClpX, is the best studied member of the Hsp100 family and functions in association with the protease ClpP to mediate substrate degradation. The general mechanism for substrate unfolding in the class II Hsp100 family seems to be the ATP dependent translocation of the substrate through the central pore. Essential for interaction with the substrate are conserved loops lining the inner, constricted pore formed by the hexameric arrangement (see also 2.2) [67] [68] [69]. ATP hydrolysis results in conformational changes in the pore loop and the adjacent subunits thereby leading to threading of the substrate [62]. The level and extent of cooperativity between the subunits is still a

a matter of debate. Recent studies on ClpX propose a semi probabilistic model of ATP hydrolysis and substrate threading [63] [62] .

The class I Hsp100 proteins, ClpB and Hsp104, on the other hand are not associated with a protease but cooperate with the Hsp70 system to mediate solubilization of protein aggregates [18] [70]. They contain an insertion in the C-terminal 4-helix bundle domain between the two NBDs. This so called M-domain is essential for the disaggregation activity of Hsp104 and ClpB, giving rise to a proposed "crowbar" like disaggregation mechanism [71] [72] [73] [74]. According to this mechanistic model, large aggregates are first pried apart by the M-domains in an ATP dependent manner, followed by further disassembly reactions possibly mediated by the pore loops and by the concerted action of the Hsp70 system. Interestingly, the position of the M-domains in the assembled Hsp100 complex is still a matter of debate, since different results have been obtained by electron microscopy (EM) studies on Hsp104 [75] [76] and ClpB [77]. While the M-domain is oriented inward and intercalated in the AAA+ hexamer in yeast Hsp104 it is oriented outward in the bacterial ClpB structures (compared in [78]).However, a recent crystal structure of the class II Hsp100 ClpC, which in contrast to Hsp104 or ClpB associates with the protease ClpP, shows the M-domain oriented outwards of the hexameric AAA+ arrangement, similar to the proposed arrangement in ClpB [56]. Loop sequences lining the inner pore are also conserved in class I Hsp100 proteins and elegant experiments in conjunction with ClpP have shown a translocase activity for ClpB, suggesting a threading mechanism of disaggregation [79].

Interestingly, yeast Hsp104 appears not only to be involved in the disaggregation of heat denatured proteins, it is also actively involved in the propagation of potentially beneficial prions (e.g. PSI$^+$) in yeast [80]. Recently, certain prion proteins (e.g. Sup35, which is known to mediate the PSI$^+$ prion state in budding yeast) have been implicated in the inheritance of beneficial traits ensuring the adaptability of the population to diverse conditions encountered in nature [81]. Hsp104 has been suggested to partition aggregated prion proteins to daughter cells and therefore is directly involved in the inheritance of these traits [81].

2.3 Photosynthesis

The process of light-dependent carbon dioxide fixation and conversion into organic sugar molecules is termed photosynthesis. After the discovery of Joseph Priestley in 1771, that "vegetation restored air that had been injured by respiration" it took about 70 years and the collaborative effort of Jan-Ingen

Housz, Jean Senebier, Théodore de Saussure and Robert Mayer to elucidate the light dependent fixation of CO_2 and release of oxygen. Photosynthesis is performed by bacteria, algae and plants and is one of the foundations of life on earth. Photosynthetically produced carbohydrates not only serve as energy source for these autotrophic organisms but also for all heterotrophic organisms such as animals that cannot produce organic (reduced) carbon containing molecules from inorganic (oxidized) carbon sources.

The general equation for the net reaction of photosynthesis is:

$$CO_2 + 2H_2A \xrightarrow{light} CH_2O + H_2O + 2A$$
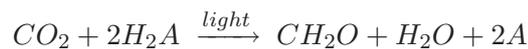
where A is a general electron donor. In oxygenic photosynthesis, water is the electron donor resulting in the formation of a 6 carbon sugar molecule and the release of oxygen:

$$6CO_2 + 12H_2O \xrightarrow{light} C_6H_{12}O_6 + 6H_2O + 6O_2$$

The compartmentalised process consists of the "light reactions" providing the reducing equivalents and ATP and the "dark reactions" that build up and reduce the carbohydrate chain. In eukaryotes, photosynthesis takes place in specialised organelles called chlororplasts, which are derived from an endosymbiotic event with a cyanobacterium [82]. Like mitochondria, chloroplasts are seperated from the cytosol by membrane layers. A permeable outer membrane and an almost impermeable inner membrane with a dedicated transport machinery enclose the intermembrane space. The compartment enclosed by the inner membrane is called stroma and contains, in contrast to mitochondria, a third membrane-enclosed compartment, the thylakoids. Thylakoids resemble a highly folded vesicle and are arranged in most organisms in stacks called grana. The thylakoid membranes are the place of the light reactions of photosynthesis, whereas the dark reactions take place in the stroma [83].

2.3.1 The "light reactions"

In the following paragraph the oxygenic electron transport light reactions of plant and cyanobacteria are explained briefly. These organisms employ a two center electron transport, whereas purple photosynthetic bacteria, heliobacteria, green sulfur bacteria and acidobacteria generally employ a one-center cyclic electron transport [84].

Plants and cyanobacteria convert light energy derived from absorbed photons to chemical energy in the form of ATP and reducing equivalents, and concomitantly oxidize water and release oxygen. This conversion involves a linear electron transport from water to NADPH and two light energy converting

reaction centers, Photosystem (PS) I and II. Energetically, this electron flow requires the sequential absorption of photons at PS I and PS II.

A schematic overview of the process is presented in Figure 2.3. Photons absorbed in the light harvesting complexes (LHC) associated with the photosystems are transferred to the special chlorophyll pair in PS II called P680 (for pigment with the maximum of absorption at 680 nm). Excitation of P680 leads to a rapid charge seperation to P680+ and ultimately electron transfer to the cytochrome b6f complex. Concomitantly, protons are transported into the thylakoid lumen. P680+ is a strong oxidant and subsequently electrons from the water splitting reaction are transferred to regenerate P680, which releases oxygen at the water splitting reaction center, the oxygen evolving complex (OEC). Plastoquinone together with the cytochrome b6f complex and plastocyanin transfer electrons to PS I. The special pair of chlorophyll in PS I is called P700 (for the maximum absorption at 700 nm). Similar to PS II, a rapid charge separation leads to the formation of P700+ after excitation and subsequently to regeneration of P700 by electrons delivered by plastocyanin. The electrons derived from the excited state of P700 are used to reduce ferredoxin in the chloroplast stroma which in turn then reduces NADP in the ferredoxin-NADP+ reductase (FNR) in a linear electron flow. However, part of the electrons derived from reduced plastoquinone can also be retransferred to cytochrome b6f to mediate the transfer of protons to the thylakoid lumen in a process called cyclic electron transport (reviewed in [85] [86]). The proton gradient that is generated across the thylakoid membrane by the actions of the water splitting reactions as well as the cytochrome b6f complex is used to synthesize ATP by the chloroplastic F_0F_1 ATPase in a process closely resembling the mitochondrial F_0F_1 ATPase.

Notably, in the past decades the complex multimeric enzymes of the light reactions have stimulated extensive structural analysis. The structure of PS I [87] as well as PS II [88] [89] [90] have been elucidated and the arrangement of the various cofactors have lead to a deeper understanding of the mechanism of the electron transport chain in the light reactions. Most recently, the structure of PS II has been solved at 1.9 Å and the structure of the Manganese cluster that mediates the water splitting reactions was revealed [91]. The enzymes involved in the light reactions are an example of how structural research on complex biochemical systems can ultimately contribute to a detailed mechanistic understanding.
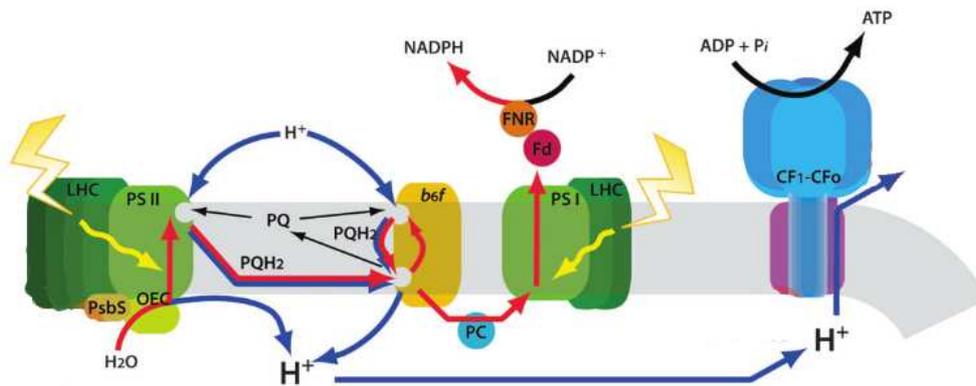
Figure 2.3: **Electron and proton flow in the light reactions of photosynthesis.**
Schematic representation of the complexes involved in the electron and proton flow in the light reactions. Linear electron flow follows along the red arrows. Electrons are derived from the oxidation of H_2O at the oxygen evolving complex (OEC) through PS II, plastoquinone (PQ/PQH$_2$). Bifurcated oxidation of PQH$_2$ occurs at the cytochrome b6f complex (b6f) where half of the electrons are linearly transferred to NADP+/NADPH coupled via plastocyanin (PC), PS I, ferredoxin (Fd), and ferredoxin-NADP+ oxidoreductase (FNR). The other half of the electrons returns to the PQH$_2$ pool. Protons released at the OEC and the cyclic reduction and oxidation of PQ/PQH$_2$, establish an electrochemical gradient of protons across the thylakoid membrane. This electrochemical gradient drives ATP synthesis from ADP at the chloroplast F_0F_1 ATPase. Adapted from [86].

2.3.2 The "dark reactions" / The Calvin Benson Bassham cycle

The energy equivalents (ATP) as well as the reducing equivalents (NADPH) generated in the light reactions are used to synthesize carbohydrates from inorganic CO_2 in the dark reactions in the chloroplast stroma in a metabolic pathway named after Melvin Calvin, James Benson and Andrew Bassham, who discovered this metabolic pathway. An overview of the pathway is shown in Figure 2.4. Briefly, the Calvin Benson Bassham cycle (CBB) can be subdivided into three phases: carbon fixation, reduction and regeneration of the substrate. In a first step, ribulose 1,5-bisphosphate (RuBP) is carboxylated by the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) to produce two molecules of 3-phosphoglycerate (see Chapter 2.4.1). In the reduction step, 3-phosphoglycerate is first ATP-dependently phosphorylated to 1,3-bisphosphoglycerate and subsequently reduced and dephosphorylated to glyceraldehyde 3-phosphate by NADPH and glyceraldehyde 3-phosphate dehydrogenase. In the regeneration step the majority of the glyceraldehyde 3-phosphate carbon molecules is used to regenerate RuBP, the substrate of Rubisco. This is accomplished in a series of isomerase, transaldolase and transketolase reac-

tions similar to the pentose phosphate pathway. In a final reaction the precursor ribulose 5-phosphate is ATP dependently regenerated to RuBP by phosphoribulosekinase. Overall the CBB cycle generates one molecule of glyceraldehyde 3-phosphate from 3 molecules of CO_2 at the "metabolic cost" of 9 molecules of ATP and 6 molecules of NADPH (reviewed in [92]).
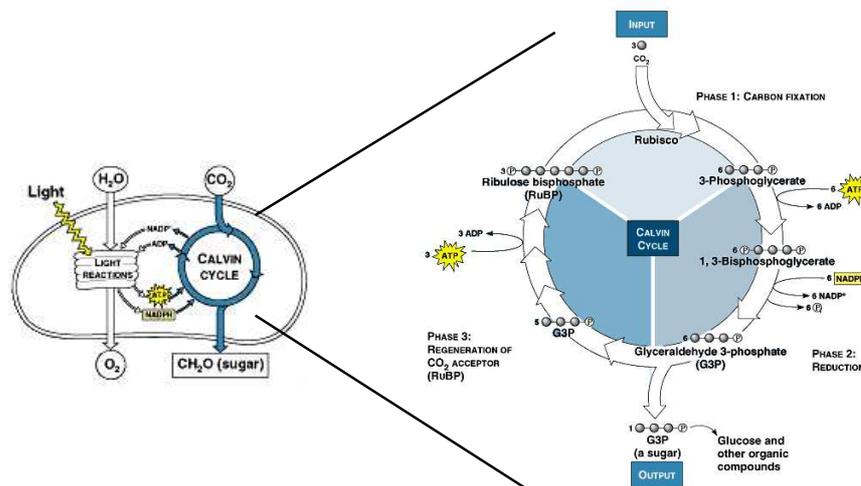
Figure 2.4: **The Calvin Benson Bassham cycle.**

Right: Schematic representation of the chloroplast showing the intricate connection of the light and dark reactions. Left: Schematic representation of the Calvin cycle. In phase 1 CO_2 is fixed by the action of Rubisco carboxylating RuBP to 2 molecules of 3-phosphoglycerate (3-PGA). 3-phosphoglycerate is subsequently phosphorylated and reduced in phase 2, resulting in glyceraldehyde-3-phosphate (G3P). In a series of isomerase, transaldolase and transketolase reaction G3P is regenerated to RuBP in phase 3. Adapted from [93].

2.3.3 Photorespiration

Photorespiration is the light and oxygen dependent release of CO_2 during oxygenic photosynthesis. Figure 2.5 shows a schematic overview of the photorespiratory pathway in contrast to the Calvin Benson Bassham cycle.

Rubisco not only carboxylates RubP, a prominent side reaction is the oxygenation of RubP to one molecule of 2-phosphoglycolate and one molecule 3-phophoglycerate. Photorespiration is mainly considered to be a salvage pathway for this prominent side reaction. 2-phosphoglycolate is toxic to the plant cell and has to be converted to 3-phophoglycerate in the photorespiratory C_2 cycle. This cycle ATP dependently recycles three carbon atoms of two 2-phosphoglycolate molecules and oxidizes one to CO_2. This oxidized carbon molecule is lost for photosynthesis and since oxygenation of RuBP is a

prominent side reaction of Rubisco, this net carbon loss can lead to a loss of every fifth carbon atom even in moderate conditions [94]. The photorespiratory pathway is comprised of 8 enzymes in four different compartments: chloroplast, cytosol, peroxisome and mitochondrion. (Reviewed in [95].)
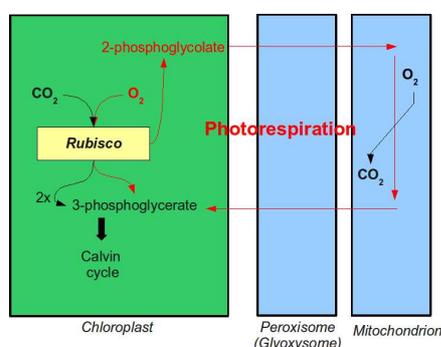
Figure 2.5: **Schematic overview of the photorespiratory pathway.**
Oxygenation of RuBP by Rubisco is a prominent side reaction, that leads to the formation of 2-phosphoglycolate. Since this molecule is toxic to the cell and cannot be used for further carbon fixation in the Calvin cycle it has to be reconverted to 3-phosphoglycerate by the photorespiratory salvage pathway. The pathway is comprised of 8 different enzymes in the chloroplast, the cytosol, the peroxisome and the mitochondrion. Since the oxygenation of glycolate to glyoxylate is catalysed in the plant peroxisomes, these organelles are often named glyoxisomes in plants.

Whether the photorespiratory pathway would be dispensable in the presence of a highly specific Rubisco is still a matter of ongoing debate. While the salvage of 2-phosphoglycolate seems essential for all photosynthetic organisms [96], it has been shown that plants grown under elevated CO_2 in conditions repressing the oxygenation side reaction show less improved growth than predicted by theoretical models [97]. This suggested that the photorespiratory pathway has another function than the mere salvage of 2-phosphoglycolate. Therefore several other hypotheses for the benefits of the photorespiratory pathway have been proposed. Since the peroxisomal enzymes for the photorespiratory pathway convey resistance to *Pseudoperonospora cubensis*, a devastating foliar disease pathogen, photorespiration has been implicated in plant defense [98]. Whereas the exact mechanism of this pathogen defense is not known, it has been speculated to be mediated by reactive oxygen species (H_2O_2, O_2^-) generated in the peroxisomes by the photorespiratory enzymes. Another more recent hypothesis is the involvement of the photorespiratory pathway in nitrate assimilation, that might explain the decrease in photosynthesis and growth of plants conducting C3 carbon fixation after long exposures (days to years) to carbon dioxide enrichment [99].

2.3.4 Carbon concentrating mechanisms

Although the photorespiratory pathway might be required for other functions than the salvage of 2-phosphoglycolate, plants, green algae and cyanobacteria have evolved carbon concentrating mechanisms (CCM) to increase the local concentration of CO_2 around Rubisco to avoid carbon loss due to the oxygenation of RuBP.

2.3.4.1 The carboxysome

The carboxysome is a cyanobacterial microcompartment enclosed by a proteinacous, icosahedral shell [100]. The shell self-assembles [101] and contains Rubisco and carboanhydrase. After the active uptake of HCO_3^-/CO_2 and accumulation of HCO_3^- in the cytosol the carboxysomal microcompartment ensures the local release of CO_2 around Rubisco by carboanhydrase [100], effectively raising the local CO_2 concentration.

2.3.4.2 The pyrenoid

The carbon concentrating mechanism of algae relies on the active transport of HCO_3^-/CO_2, but is evolutionary more diverse than the cyanobacterial CCM. Interestingly, in most algae (e.g. *Chlamydomonas rheinhardtii*) with a carbon concentrating mechanism Rubisco is located in a specialised microcompartment within the chloroplast lumen called the pyrenoid [102]. Whereas the carboxysome encloses Rubisco in a proteinacous shell, the pyrenoid shell is composed of starch like oligosaccharides [103]. Due to the diversity of employed CCM and the difficulty of biochemically characterizing oligosaccharide enclosed microcompartments the mechanistic details of CCM in algae are mostly unknown.

2.3.4.3 C4 and crassulacean acid metabolism (CAM) photosynthesis

Terrestrial plants (*Embryophyta*) have evolved three different photosynthetic pathways. C3 photosynthesis as described in previous paragraphs, C4 photosynthesis and the CAM pathway.

C4 photosynthetic plants show a special cellular anatomy called Kranz anatomy, characterised by an arrangement of the so called bundle sheath cells around the vascular tissue followed by the positioning of the mesophyll cells directly on the outer surface of the bundle sheath cells. In the mesophyll cells CO_2 (as HCO_3^-) is affixed to phosphoenolpyruvate (PEP) by PEP-carboxylase to oxaloacetate, which is then shuttled to the Rubisco containing bundle sheath cells as aspartate or malate. These C4 acid

intermediates (after which this photosynthetic pathway has been named) are then decarboxylated in the bundle sheath cells releasing CO_2 in the compartment of CO_2 fixation by Rubisco. The evolutionary origin and regulation of this spatial and mechanistic separation of initial CO_2 uptake and fixation is a field of research [104] [105] stimulated by initiatives for the transgenic incorporation of the C4 metabolism in crop plants (e.g. rice) [106] .

Whereas C4 photosynthesis reflects a spatial separation of CO_2 uptake and fixation, the crassulacean acid metabolism (CAM) pathway separates the two processes temporally. Evolved mostly in arid, hot conditions, CO_2 is initially affixed to PEP by PEP-carboxylase at night. The resulting oxaloacetate is converted to malate and stored in the vacuole. This makes it possible for CAM plants to close their stomata by day and reduce water loss, but release CO_2 for the fixation by Rubisco due to the decarboxylation of the stored malate [107].

2.4 Ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco)

Rubisco is the key enzyme of the dark reactions and arguably the most abundant protein on earth [108]. It catalyses the exergonic reaction of 1,5-ribulosebisphosphate (RuBP) and CO_2 to two molecules of 3-phosphoglycerate. Despite its crucial role in photosynthesis, its catalytic properties are rather poor when compared to other enzymes, which presumably necessitates its abundant expression [109]. For example Rubisco not only catalyses the carboxylation but also the namegiving oxygenation of RuBP, that leads to a net carbon loss and necessitates an energy consuming salvage pathway for 2-phosphoglycolate the photorespiration pathway (see 2.3.3). Additionally, the catalysed reaction proceeds comparably slow (kcat ~2-3 s^{-1}) and Rubisco is prone to inhibition by sugar phosphates. Optimization of Rubisco is predicted to enhance the photosynthetic properties of crop plants and has been a long standing goal of the research field (e.g. [110] [111]).

Interestingly, Rubisco from different species with only small variations in the sequences of the large and small subunits display different kinetic and affinity parameters. Indeed, compared to Rubisco of green plants, some of the Rubisco enzymes from red algae show a 2-3 fold higher substrate specificity for CO_2 over O_2 [112] [113]. Notably, Rubisco from the red algae *Griffithsia monilis* even displays kinetic characteristics superior to the best Rubiscos found in green plants [114].

2.4.1 The structure of Rubisco

Rubisco enzymes have been identified in most autotrophic species. So far four structural forms of Rubisco or Rubisco like enzymes have been identified in prokaryotes (bacteria, archea) and eukaryotes (algae, higher plants) [115].

Rubisco enzymes are always multimers and can contain a large (~52 kDa) and sometimes a small subunit (~15 kDa). Besides the sequence of the large subunit, the presence or absence of the small subunit is one classifying feature of Rubisco sequences and structures. Of the four types of Rubisco structures characterised so far, only form I contains a small subunit, whereas the oligomerization of the large subunits to at least dimers is retained in all forms. Figure 2.6 shows an overview of the different forms of Rubisco that are described in the following paragraphs. Only form I-III Rubiscos catalyse the name giving affixation of CO_2 or O_2 to RuBP [115].

Form I	Form II	Form III	Form IV
Spinacia oleracea	*Rhodospirillum rubrum*	*Thermococcus kodakarensis*	*Chlorobaculum tepidum*
(1RCX) top view	(5RUB) side view	(1GEH) top view	(1YKW) side view
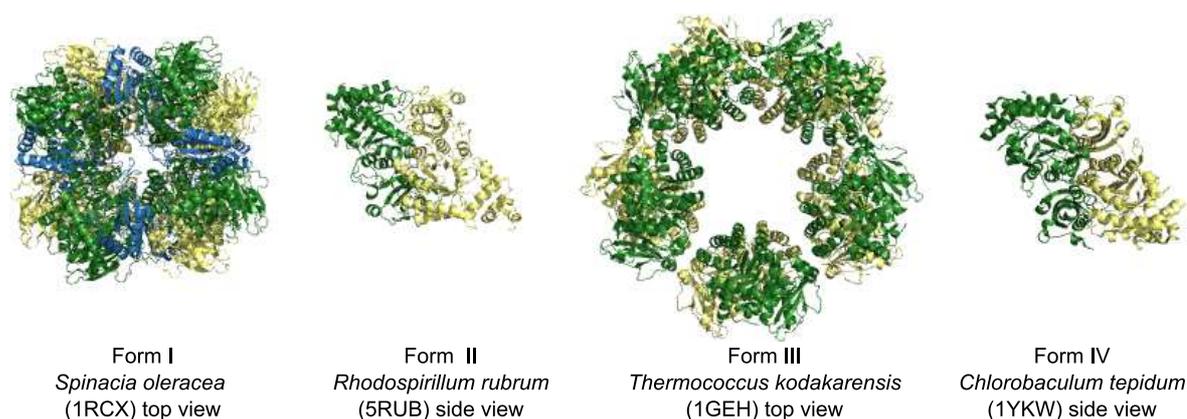
Figure 2.6: **Representative structures of the different forms of Rubisco.**

All forms are comprised of antiparallel dimers of large subunits. Form I forms tetramers of dimers and is the only form that requires the presence of the small subunit. Form II and Form IV are structurally highly similar and form a single antiparallel dimer of large subunits. Form III Rubisco forms tetramers or pentamers of dimers, shown here is the pentameric form. Large subunits are shown in in green and yellow, small subunits are shown in blue. All structures are shown in a ribbon representation and are drawn to scale. PDB accession numbers are indicated in brackets under the species name 1RCX [116], 5RUB [117], 1GEH [118] , 1YKW [119].

2.4.1.1 Form I

Form I Rubiscos are further subdivided into Form I A-D. Form I A and B are found in cyanobacteria, proteobacteria, green algae and plants and thus are commonly referred to as green-type Rubiscos, whereas

form I C and D are found mostly in proteobacteria and rhodophytes and thus are called red-type Rubiscos [115] [120]. While the branching of form I enzymes probably occurred before the first endosmybiotic event, phylogenetical analysis suggest that branching into the green and red lineage occurred after the branching in form I-IV [115].
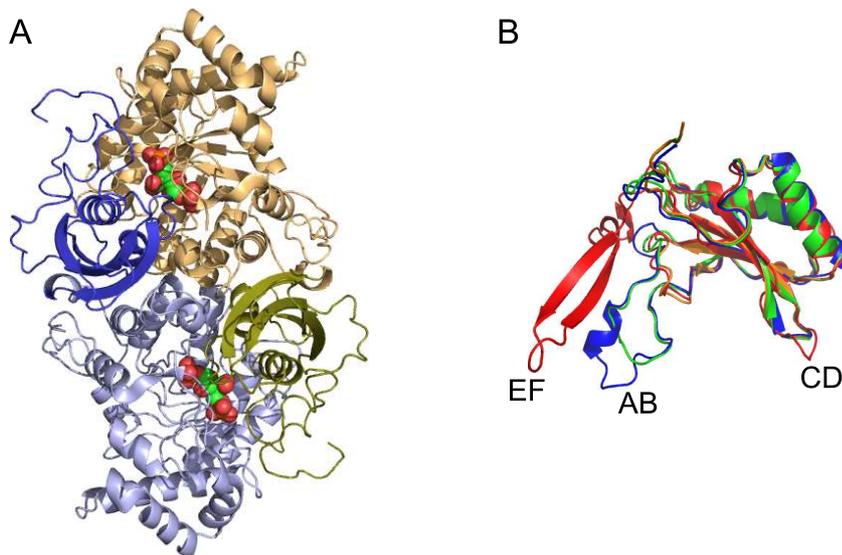
Figure 2.7: **Structure of the large and small subunit of Rubisco.**

A: Structure of the antiparallel dimer of form I Rubisco large subunits (from *Nicotiana tabacum*) shown in a ribbon representation colored in blue and brown shades respectively. N-terminal mixed β-sheet domains are shown in darker shades, C-terminal α/β-barrel domains are shown in lighter shades. Large and small subunits forming the hexadecameric L_8S_8 holoenzyme are not displayed. The structure of a tight binding inhibitor 2-carboxy-arabinitol-bisphosphate (CABP) bound to the active sites is shown in a ball and sticks representation with carbon atoms colored in green, oxygen atoms colored in red and phosphate atoms colored in orange. PDB accession number 4RUB [121]. B: Superposition of small subunit structures from different species. All structures are shown as ribbon representations. Structures are color coded according to their origin: *Nicotiana tabacum* green 4RUB [121], *Synechococcus elongatus PCC 6301* orange 1RBL [122], *Chlamydomonas reinhardtii* blue 1GK8 [123], *Galdieria partita* red 1BWV [124]. Variable loops are named according to their adjacent β-sheets. Not the variability of the AB loop, as well as the unique insertion of the additional β-sheets EF at the C-terminus of the small subunit of the red-type Rubisco from *Galdieria partita*.

All Form I enzymes are hexadecameric complexes consisting of a tetramer of antiparallel dimers of the large subunits capped by four small subunits at the top and bottom, respectively. This gives the complex a square prism '422' symmetry. The active site is formed in all bona fide Rubiscos (Form I-III) by a dimerization of the large subunits mediated by mostly charged interactions. The carboxy-terminal α/β-

barrel domain of one large subunit forms the active site with the amino-terminal mixed β-sheet domain of the dimerization partner [120] [125] (see Figure 2.7 A). The active site opens and closes during catalysis, mostly by movement of a loop (loop 6) and ordering of the extreme C-terminus of the enzyme [126].

While the structure of the large subunit is highly conserved in all Rubiscos, the structure of the small subunit is more diverse with the common core consisting of a four stranded anti-parallel β-sheet. The loop between strands A and B together with the extreme C-terminus show the highest sequence and structural divergence [127] (see Figure 2.7 B). Although the small subunit does not have a direct role in catalysis, chimeric Rubisco complexes show different kinetic properties e.g. in the specificity for the substrate CO_2 [127] [128].

2.4.1.2 Form II

Form II Rubisco consists of a single, antiparallel dimer of large subunits, forming two active sites by dimerization much like in form I Rubiscos. Found in Proteobacteria and some Dynoflagellates, form II Rubisco was used as a model protein for the catalysis and folding/assembly of Rubisco for decades, mainly because the structure of the enzyme of Rhodospirillum rubrum was solved already in 1986 [129] (later refined at higher resolution [117]) and the enzyme could be expressed heterologously.

2.4.1.3 Form III

Form III Rubisco sequences were first identified in large genomic scale sequencing studies in archaea [130]. They tend to form dimers similar to the form II enzyme or larger oligomeric states, for example tetramers or pentamers of dimers [131] [118] (PDB ID: 2CXE, 2CWX, 2D69 unpublished).

2.4.1.4 Form IV

Form IV Rubisco are not bona fide Rubiscos and show substitution in key active site residues involved in CO_2 fixation. They have been suggested to bind different substrates for enolase/lyase type reactions [132] of e.g. the methionine salvage pathway. The Rubisco-like proteins identified so far assemble into antiparallel dimers of large subunits and are structurally closely related to the form II enzymes [133].

2.4.2 Folding and assembly of form I Rubisco

Folding and assembly of the form I holoenzyme in plants requires the coordinated import of the small subunit from the cytosol into the chloroplast and the folding and assembly of the large subunit, which is synthesized in the cloroplast. Folding of the large subunit is dependent on encapsulation in the chloroplast chaperonin system (Cpn60, Cpn10, Cpn20 in plants). Although some cyanobacterial form I Rubiscos have been successfully expressed in *E. coli* using the heterologous GroEL/ES chaperonin [134] [135], the chaperonins of plant chloroplasts seem to be highly adapted for the folding of plant Rubisco. In addition, after the large subunit has been released from the chaperonin cage the assembly of the folded subunits requires the action of an assembly chaperone called RbcX [136] [137].

RbcX oligomerizes to an α-helical homodimer (RbcX$_2$) of two subunits each with a a molecular weight of 15 kDa [137]. Detailed biochemical experiments together with EM and crystal structures of the Rubisco RbcX$_2$ complex showed that RbcX$_2$ acts as a molecular stapler, binding to a C-terminal peptide of one subunit and to the N-terminal domain of an adjacent subunit, thus guiding and stabilizing proper RbcL dimer formation. Proper dimerization by binding of RbcX$_2$ also prevents back-binding to the chaperonin in vitro [138]. RbcX$_2$ is released by binding of RbcS to the assembly intermediate [137] [138]. Comparison of the crystal structures of the L$_8$S$_8$ holoenzyme with the crystal structure of L$_8$(X$_2$)$_8$ assembly intermediate lead to the hypothesis that binding of the small subunit displaces RbcX$_2$ by ordering of the '60ies loop' (residues 62-70) in the N-terminal domain of the RbcL subunit in a conformation that prevents contacts to RbcX [139]. The residues in the '60ies loop' form part of a wall of the active site and residues in this loop (Thr 62, Trp63) are directly involved in substrate binding. Thus binding of the small subunit presumably contributes to obtaining catalytic activity in vivo by stabilizing the loop residues in a catalytically competent conformation. Figure 2.8 shows an overview of the folding pathway of form I Rubisco. The complex and highly specialised assembly pathway shows that evolution of Rubisco might be highly restricted by the difficult folding pathway for the large subunits explaining the high conservation of Rubisco large subunit sequences as well as the rather poor catalytic properties of the enzyme.
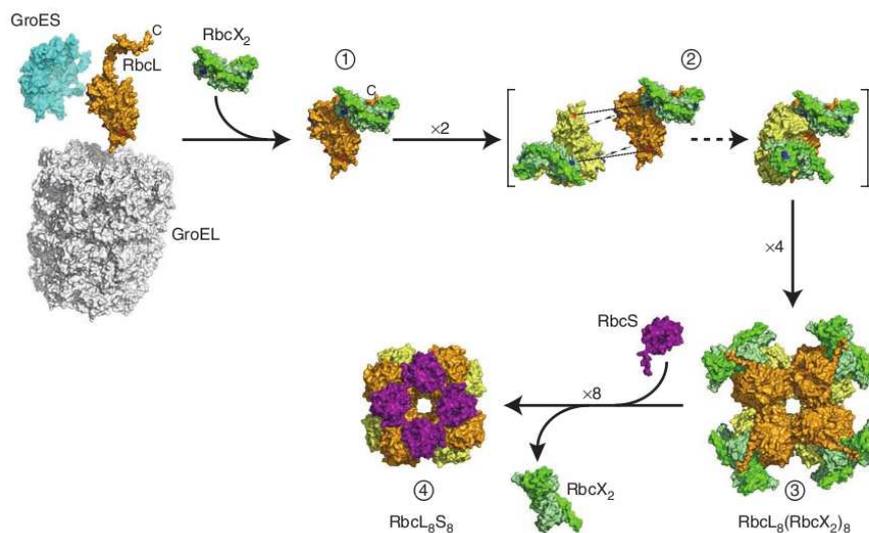
Figure 2.8: **Folding and assembly of form I Rubisco.**

After the GroEL/ES dependent folding of the large subunit, RbcX dimers mediate the correct dimer formation of the large subunit (1 and 2). The dimers subsequently assemble to tetramers while RbcX stays bound (3). Binding of the small subunit displaces RbcX and the active holoenzyme (L_8S_8) is formed (4). Adapted from [139].

2.4.3 Reaction mechanism

For catalysis the folded and assembled Rubisco holoenzyme requires the spontaneous carbamylation of an essential lysine residue (Lys 201) and subsequent binding of a Mg^{2+} ion in the active site located at the top of the TIM barrel in the C-terminal domain. This process is called activation of the enzyme [140] [141] [142] (Figure 2.9 A).

The reaction proceeds by binding of the substrate RuBP and closing of the active site by closure of loop 6 and ordering of the C-terminus, thus preventing the access of water molecules to the reactive intermediates (enolate) of the reaction. In an acid base catalysed reaction the ketone form of RuBP is isomerized to an enediol, which in turn makes an electrophilic attack of the gaseous substrates CO_2 or O_2 possible. Structural as well as theoretical models propose the carbamate of Lys 201 and Lys 175 as important bases and acids in the isomerization reaction [143] [144] [120]. As the product of the carboxylation is a β-ketonic acid, it quickly gets hydrolysed and enzymatically protonated to two molecules of 3-phosphoglycerate. The product of oxygenation similarly gets hydrolysed and protonated to one molecule of 3-phosphoglycerate and one molecule of 2-phosphoglycolate (as reviewed in [145]). An important indicator for the existence of the instable 6-carbon (carboxylation) or 5-carbon (oxygena-

tion) intermediates is the tight binding of the intermediate analogue 2-carboxy-arabinitol-bisphosphate (CABP), a potent competitive inhibitor of Rubisco. Figure 2.9 B shows an overview of the catalysed reaction, whereas Figure 2.9 A shows the structure of the active site.

The Mg^{2+} liganded by the carbamylated lysine 201, aspartate 203 and glutamate 204 and three water molecules in the apo structure plays a crucial role in postioning the substrates for catalysis. From crystal structures of the fully activated and CABP bound Rubisco holoenzyme it can be concluded that the substrate RuBP displaces two water molecules coordinated by Mg^{2+} and that the gaseous CO_2 would replace the third, positioning it for the electrophilic attack explaining the essential role of a metal ion for the catalysis of Rubisco [120].

The closure of the loop6 during the reaction cycle seems to depend mostly on the interphosphate distance of the bound substrate, however, the exact timing and trigger for closure remains poorly understood [126]. Interestingly, mutations in loop6 have shown, that pore loop sequence and rigidity/structure influence specificity of Rubisco, likely through the lysine residue 334 at the tip of the loop that is believed to position the gaseous substrate as well as lock the closed loop by a salt bridge to glutamate 60 in the N-terminal domain [146] [147]. The loop is further locked in the closed conformation by ordering of the C-terminus and formation of ionic contacts of aspartate 473 - the so called latch residue - to the main body of the enzyme [126] [148].

Presumably due to the complex catalytic mechanism several, "misfire" products of the Rubisco reaction are known. Products of catalytic 'misfire' arise at various points in the reaction pathway of both oxygenation and carboxylation. The misfiring product xylulose 1,5-bisphosphate (XuBP), for example, is formed by misprotonation of the initial enediol intermediate [149].
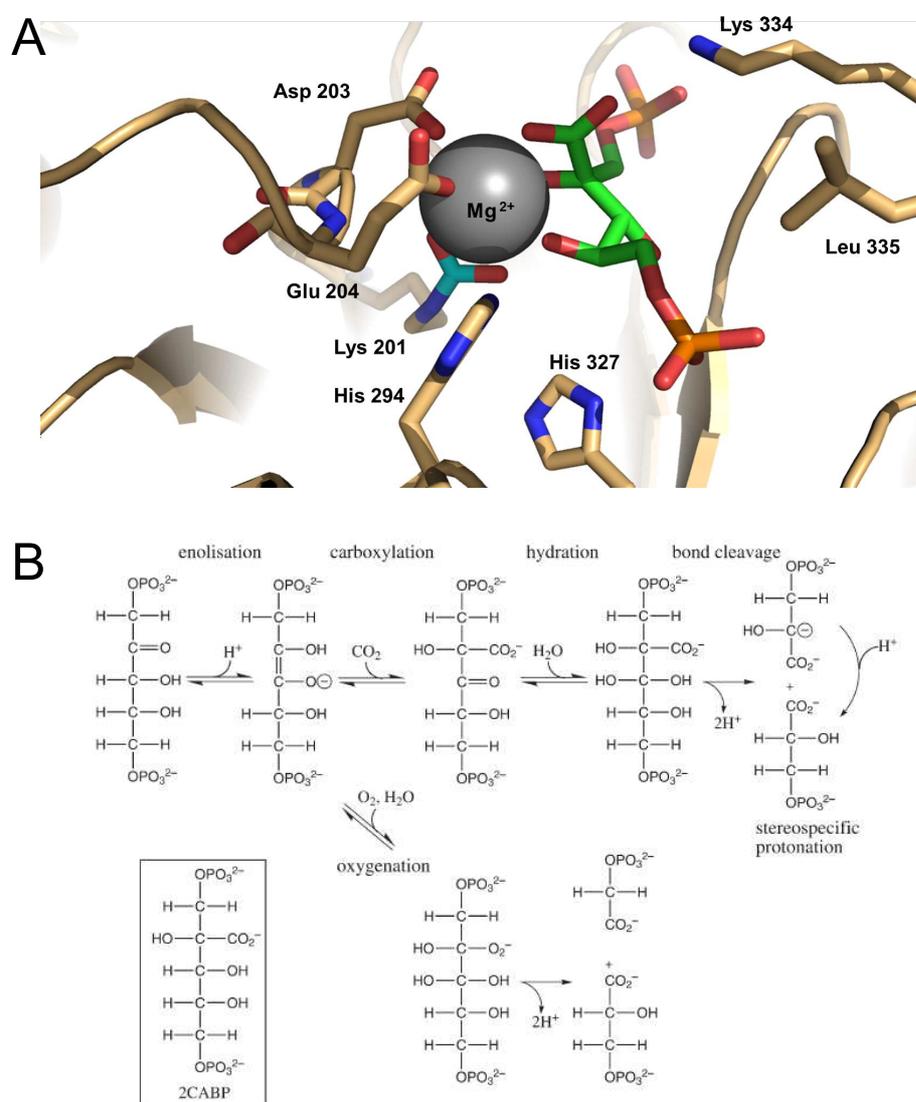
Figure 2.9: **The reactions catalysed by Rubisco.**

A: Structure of the activated active site of Tobacco Rubisco bound to the inhibitor CABP (PDB 4RUB [121]). CABP is represented as a stick model with carbon atoms colored in green. Selected active site residues are represented as a stick model with carbon atoms shown in light brown. The Mg^{2+} ion is shown as grey sphere and the carbamate group on Lys 201 is shown with cyan-colored carbon atom. Selected phosphate residues are colored orange, nitrogen atoms are colored blue and oxygen atoms are colored red. Structure of Rubisco large subunit is shown in light brown in a ribbon representation. B: Overview of the catalytic mechanism of Rubisco. Oxygenation and carboxylation of RuBP are shown. For comparison with the catalytic intermediates, the structure of the inhibitor CABP is also shown. Adapted from [120].

2.4.4 Regulation of Rubisco activity

Rubisco is a key enzyme in the dark reactions of photosynthesis and is - under suitable growth conditions - the rate-limiting step, making it the target for regulation of the carbon fixation cycle. Modulating expression and degradation of Rubisco is probably a means of long-term regulation. The effect of several oxidation dependent modifications of cysteine residues in Rubisco in the regulation of expression and degradation of the enzyme is currently being discussed [150]. However, short-term regulation of Rubisco is mostly achieved by a number of effector molecules such as phosphate or sulfate-anions, phosphorylated sugar molecules or NADPH [151]. Most of these effector molecules bind at the active site of Rubisco, although another binding site for anorganic phosphate has been identified in Rubisco at the so called 'latch' site [126] [152], the docking site for the extreme C-terminus of Rubisco. At high enough concentrations (>10mM), all effector molecules prevent Rubisco catalysis by competitive inhibition, however, at lower, sub-stoichiometric concentrations - which likely prevail in vivo - some effectors (e.g. phosphate or pyridoxalphosphate) greatly facilitate activation of the enzyme by an as yet poorly understood mechanism and thus positively affect Rubisco activity [151].
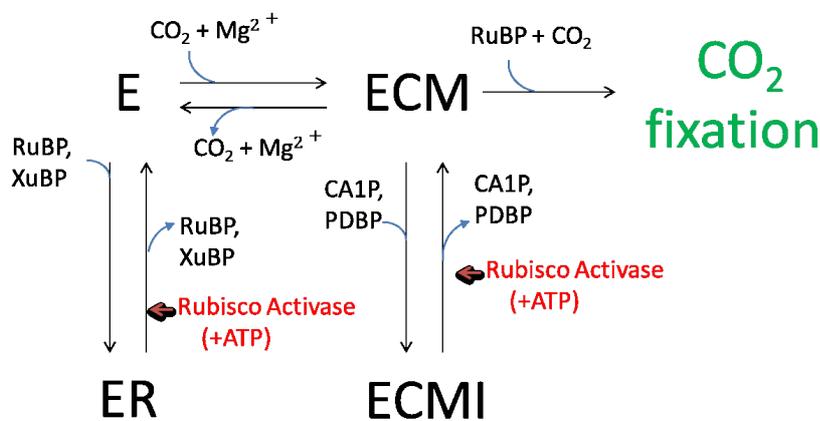
Figure 2.10: **Rubisco regulation by inhibitors and Rubisco activase.**
Rubisco (E) must be carbamylated and bind Mg^{2+} to be activated (ECM) for CO_2-fixation. Binding of the substrate RuBP or the misfiring product xylulose 1,5-bisphosphate (XuBP) to the inactivated enzyme result in a stable, inhibited complex (ER). Binding of the night time inhibitor carboxy-arabinitol 1-phosphate (CA1P) or another misfiring product, glycero-2,3-pentodiulose-1,5-bisphosphate (PDBP), to activated Rubisco similarly result in a stable, inhibited complex (ECMI). Inhibited complexes are ATP dependently reactivated by Rubisco activase [151].

Furthermore, inactivation and reactivation (decarbamylation and carbamylation) of Rubisco seem to

be a major factor in the physiological short term regulation of Rubisco [151]. The concentration of magnesium ions and the pH in the chloroplast stroma is for example tightly controlled [153]. The pH of chloroplast stroma increases in the light due to the action of the light reaction, which concomitantly increases the Mg^{2+} concentration by an unknown mechanism. Lower pH and Mg^{2+} concentrations in the dark presumably limit the amount of activated Rubisco at night [141]. In addition to that, binding of the substrate RuBP or the misfiring product xylulose-1,5-bisphosphate (XuBP) to the non-carbamylated form of Rubisco as well as binding of other inhibitory sugar-phosphates such as the night time inhibitor carboxy-arabinitol-1-phosphate (CA1P) to the carbamylated form, lead to the formation of stable complexes, that have to be remodeled for reactivation. The enzyme responsible for this remodeling function was identified in 1985 and accordingly named Rubisco activase (Rca) [154]. Figure 2.10 shows an overview of the Rubisco regulation by inhibitors. Since in vivo Rubisco activity critically depends on the action of Rubisco activase, there is mounting evidence that Rubisco activity is mostly controlled by the activity of Rubisco activase. Interestingly, the activity od Rubisco activase is controlled by redox-regulation in many plant species, which allows for indirect control of Rubisco activity in response to the redox status of the chloroplast stroma due to illumination [151] (see 2.5.2).

2.5 Rubisco activase (Rca)

The key to the discovery of Rubisco activase was the characterization of the *rca* mutant strain of *Arabidopsis thaliana* by Chris Somerville, Archie Portis Jr. and William Ogren. This strain only grew at high CO_2 concentration and accumulated RuBP because of a substantially decreased activation status of Rubisco [155] [156] (hence the strain was named *rca* for Rubisco activation). A combination of biochemical and genetic experiments lead to the finding that a protein, which later was named Rubisco activase, was absent in the Arabidopsis *rca* strain. This protein was subsequently discovered to be responsible for Rubisco activation [154]. Upon development of a suitable in vitro assay for this enzyme [154] it was discovered, that Rubisco activase is an ATP dependent [157] remodeler of Rubisco and therefore facilitates the release of bound sugar phosphates but is not actively involved in carbamylation or magnesium ion binding to the active site [158] (see also Figure 2.10). Subsequent mutagenesis studies (see 2.5.1) and preliminary EM studies of a putative Rubisco-Rubisco activase complex (see 2.5.3.1) gave rise to first mechanistic insights into the reactivation process. A large scale comparison of sequences of related ATPases identified activase as a member of the rapidly growing family of AAA+ proteins in 1999 [49].

Since the maintenance of Rubisco activity in vivo critically depends on the action of Rubisco activase, the mechanism and regulation of activase are the focus of ongoing research. Furthermore, Rubisco activase is a heat labile enzyme [159] [160], aproperty that might limit the photosynthetic capacity of plants under moderate heat stress. Therefore, improving activase stability is predicted to result in plants with improved photosynthetic properties and growth under moderate heat stress [161] [162] [163].

In addition to the main role of Rubisco activase in the reactivation of Rubisco, there are newly emerging controversial reports of additional roles of activase in plant metabolism. The finding that Rubisco activase is reversibly associated with the thylakoid membrane under moderate heat stress lead to the proposal of a classical chaperone function of activase for the components of thylakoid associated translation machinery [164]. However,the heat instability of activase itself [160] together with association of activase with cpn60 under moderate heat stress [165] argued against this dual role. Furthermore, it has been shown that the association of activase with the thylakoid membrane can be observed independent of heat stress, depending on the pH and the ATP levels in the stroma [166]. Another emergent new role is the involvement of activase in the jasmonate mediated leaf senescence pathway in plants. The jasmonate induced repression of Rubisco activase expression leads to rapid leaf senescence [167], which might be explained by a complete repression of the Calvin cycle under these conditions.

Proteins with sequence homology to Rubisco activase have also been identified in filament forming, heterocystous cyanobacteria [168] and subsequently named Rubisco activase-like proteins (rca-like). Inactivation of Rubisco from these species by known sugar-phosphates could not be observed in vitro [136], whereas in vivo the rca-like gene showed a diurnal expression pattern similar to plants [169]. In addition, a knockout strain for the rca-like gene of *Anabaena variabilis* required high CO_2 to grow [170], similar to a knockout of Rubisco activase in plants, consistent with an activase function in vivo. It is interesting to note that in heterocystous cyanobacterial colonies the bacteria can differentiate into two different cell types under nitrogen starvation. While one population of looses the ability for photosynthesis and becomes the nitrogen fixing cell type, the vegetative cells sustain photosynthesis and deliver carbohydrates to the nitrogen fixing cells. A requirement for high activation levels of Rubisco in these conditions seems likely, since the vegetative cells have to photosynthetically sustain the nitrogen fixing cells.

2.5.1 Putative mechanism of Rubisco activation

To elucidate the mechanism of Rubisco activase several mutagenesis studies have been conducted with the goal to identify critical residues involved in ATP binding and hydrolysis as well as activation function. The results of these studies will be discussed here in the context of the subdomains of Rubisco activase. Figure 2.11 shows the domain architecture of Rubisco activase and highlights residues critical for the function. Additionally, Figure 5.2 provides an overview of Rubisco activase sequences and highlights residues identified in mutational studies. In the absence of a structure of Rubisco activase most of these mutations were based on sequence-homology models of Rca to other AAA+ proteins [171].
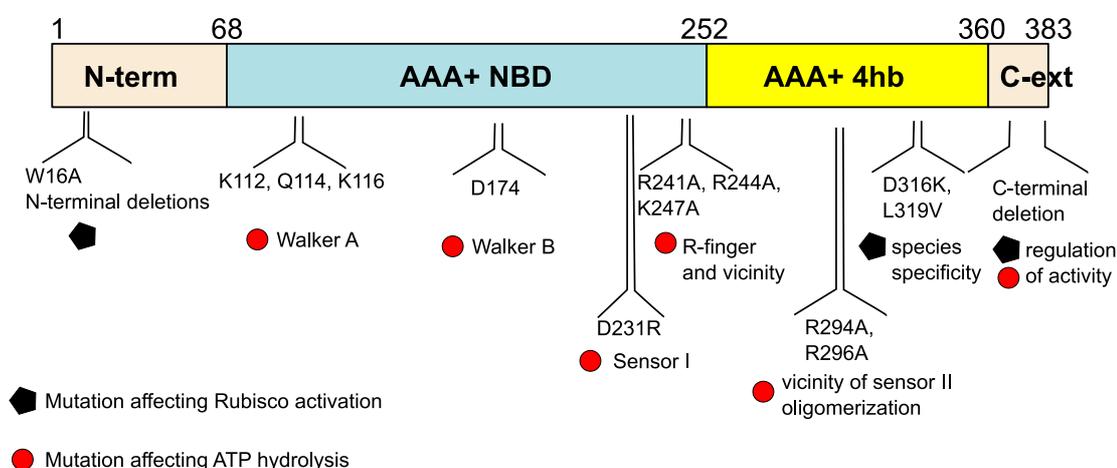
Figure 2.11: **Overview of mutational studies on Rubisco activase.**
Domain architecture of Rubisco activase shown as rectangular bar. (N-term: N-terminal domain. NBD: nucleotide binding α/β-subdomain. 4-hb: 4-helix bundle subdomain. C-ext: C-terminal extension.) Mutated residues are indicated together with their respective effect and sequence element.

2.5.1.1 The N-terminal domain

Rubisco activase features an N-terminal extension of 68 residues [171]. Truncation of the N-terminal 50 residues of tobacco activase lead to a loss of activase activity but retention of the ATPase activity [172]. The same result was independently observed for spinach activase, when only the N-terminal 12 residues were truncated. However, N-terminal truncation of 10, 8 or 4 residues lead to preservation of Rubisco activation [173]. Further mutagenesis clearly implicated a conserved tryptophan residue in the N-terminal domain (residue 16 in tobacco, residue 11 in spinach activase) as a critical residue for Rubisco activation [172]. In many proteins of the AAA+ family N-terminal domains serve as adaptors for other

(substrate-)proteins (see 2.2). Similar to these N-terminal adaptor domains the conserved tryptophane 16 might be part of a hydrophobic interaction interface with the substrate Rubisco.

2.5.1.2 The nucleotide binding subdomain

Even before activase was classified as a AAA+ protein, several mutagenesis studies aimed to identify critical residues for the ATPase function. Guided by experiments with crosslink-reactive nucleotide analogues [174] [175], the nucleotide hydrolysing domain as well as several critical residues for nucleotide binding and hydrolysis were identified. The identified residues mostly map to the well known features essential for AAA+ ATPase activity (see 2.2) such as Walker A loop (nucleotide binding K 112 [176], Q 114 [176], K 116 [177]), sensor 1 (nucleotide hydrolysis D 231 [178]), Walker B (nucleotide binding and hydrolysis D 174 [178]) and the arginine finger (intersubunit interaction and nucleotide binding R 241, R 244 [179] and K 247 in the vicinity [175]).

Interestingly, mutations of glutamine 114 (in the Walker A loop) to acidic amino acids (glutamate, aspartate) lead to slight stimulation of the ATPase activity, an increased Rubisco activation and reduced inhibition by ADP (see 2.5.2), indicating a possible role in coupling the energy of ATP hydrolysis to the protein remodeling activity [176] [180], as well as an involvement of this residue in ADP inhibition of activase.

2.5.1.3 The C-terminal 4-helix-bundle subdomain

The C-terminal 4-helix bundle domain contains residues involved in the substrate specificity of Rubisco activase. Rubisco activase from the plant family *Solonacaea* (e.g. tobacco) does not activate Rubisco from plants outside the family (e.g. spinach, *Chlamydomonas reinhardtii*) and Rubisco activase from *non-Solanacaea* does not activate *Solanacea* Rubisco [181]. A detailed comparison of sequences allowed the identification of residue substitutions in the *Chlamydomonas reinhardtii* Rubisco large subunit that mediate this species specificity. Mutation of proline 89 to arginine or alanine and mutation of aspartate 94 to lysine resulted in Rubisco that was no longer efficiently activated by spinach Rubisco activase but instead was activated by tobacco Rubisco activase [182] [183]. A similar comparison of activase sequences resulted in the identification of two residue substitutions in the C-terminal 4-helix bundle that also lead to a specificity conversion. Chimeric Rubisco activase of tobacco, in which parts of the C-terminal domain were replaced by the respective spinach sequence were able to activate spinach Rubisco,

but were impaired in the activation of tobacco Rubisco [184]. Specific substitution to the *non-Solanacaea* sequence of the residues aspartate 316 to lysine and leucine 319 to valine in tobacco resulted in the same effect. The analogous mutations in the *non-Solanacaea* cotton Rubisco activase to the *Solanacea* residues resulted in a conversion of the specificity as well [184]. These results indicated a direct physical interaction between the C-terminal subdomain of activase and an equatorial region on Rubisco. Since the species specificity reversing mutations on Rubisco and Rubisco activase seemed to be essentially charge switch mutations between the residues 94 in Rubisco large subunit (lysine for *Solanacea*, aspartate for *non-Solanacea*) and residue 316 (aspartate for *Solanacea*, lysine for *non-Solanacea*) in Rubisco activase, a direct salt bridge between these residues was implicated [184].

In addition to these findings related to the interaction between Rubisco and Rubisco activase, a comprehensive mutagenesis study of conserved arginine residues prompted the identification of two arginine residues in the Tobacco activase (R 294 and R 296) C-terminal domain important for nucleotide hydrolysis [179]. While mutation of arginine 294 to alanine lead to impaired nucleotide hydrolysis and impaired Rubisco activation it also affected subunit oligomerization. The mutation of arginine 296 to alanine only affected nucleotide hydrolysis and Rubisco activation and was therefore interpreted to be the Sensor 2 arginine involved in sensing the bound substrate as well as the nucleotide. It is interesting to note, however, that classical clade AAA+ proteins usually do not contain a sensor 2 arginine residue for nucleotide sensing (see 2.2).

2.5.1.4 The C-terminal extension

Compared to the core AAA+ subdomains, the C-terminal extension is less conserved in Rubisco activases and in many species (e.g. *Arabidopsis thaliana*) two Rubisco activase functional isoforms were identified that vary in the length of the C-terminal extension [185] [186]. Interestingly, deletion of the C-terminal 19 residues of the shorter isoform of spinach Rubisco activase lead to an increased Rubisco activation and slight reduction of the ATPase rate [173], confirming the earlier hypothesis that the long isoform might play a role in regulation of Rubisco activase [177] (see also 2.5.2).

2.5.2 Regulation of Rubisco activase

The finding that the ATPase activity of Rubisco activase is inhibited by ADP [157] [177] lead to the hypothesis that activase is controlled by the ATP/ADP ratio in the chloroplast stroma, enabling a modu-

lation of ATP hydrolysis of activase as well as the Rubisco activity in response to illumination. However, while the Rubisco activation status clearly responded to differences in illumination levels (becoming more active at higher light intensities), the RuBP, ATP or ADP levels in the chloroplast did not change significantly [187]. The increased sensitivity to ADP inhibition of the long isoform of activase [177], however, prompted further studies to investigate the regulatory function of the long C-terminal extension.

In many plants activase (e.g. *Arabidopsis thaliana*, *Spinacia oleracea*, *Oryza sativa*) has two isoforms, a short (40-42 kDa) and long isoform (~46 kDa) usually generated by alternative splicing, which vary in the length of the C-terminal extension. Mutation of two conserved cysteine residues in the C-terminal extension of the long isoform of Rca from *Arabidopsis thaliana* resulted in less inhibition by ADP and higher ATPase as well as Rubisco activation activities in the presence of a fixed amount of ADP in vitro. The same effect was observed in the presence of the reducing agent dithiothreitol (DTT) and thioredoxin-f but not DTT and thioredoxin-m [188]. Additionally, in vitro mixing experiments with long and short isoforms confirmed a regulatory influence of the long isoform on the short isoform. In oxidizing conditions the short isoform was completely inhibited by the presence of the long isoform in a 1:1 ratio at an ATP/ADP ratio of 1:3 [189]. Since the levels of reducing equivalents such as reduced thioredoxin-f increase with increasing light irradiation, it was hypothesized that redox-regulation of the long isoform regulates Rubisco activase activity in response to light and therefore also controls Rubisco activity. This regulatory pathway could be confirmed in vivo by expression of just one of the isoforms in the *Arabidopsis thaliana rca* mutant. Expression of the short isoform showed high Rubisco activation levels even in the dark, whereas expression of the long isoform showed a light responsive regulation of Rubisco activity [190]. The C-extension of the oxidized large isoform can be indirectly crosslinked to residues near the nucleotide binding pocket (peptides near Walker A, R-finger and peptides in the sensor 2 near to the nucleotide binding pocket). It was therefore postulated that the oxidized C-terminal extension of the long isoforms blocks binding of ATP, explaining the reduced binding affinity to ATPγS in the oxidized state [191] (see also 5.4).

Although the long isoform of Rubisco activase associated with the redox-regulation plays a pivotal role in regulating Rubisco activation in response to varying light conditions, it is interesting to note that in some plant species such as tobacco no long isoform has been identified so far.

Additionally, Rubisco activase expression levels in plants generally fluctuate in a diurnal manner, with the highest expression at the beginning of the light cycle and the lowest expression during the night [192].

This regulated expression closely resembles the diurnal activation cycle of Rubisco.

2.5.3 Oligomeric state

Intriguingly, Rubisco activase from different species polydispersely oligomerizes in a concentration and nucleotide dependent manner. Various oligomeric sizes from ~50 kDa to ~600 kDa have been described dependent on analysis method, Mg^{2+}, nucleotide and activase concentration [193] [194] [195] [196]. The specific ATPase as well as the Rubisco activation activity and the oligomeric size increase with increasing concentrations and in the presence of crowding agents (e.g. 5 % PEG 3350) [194]. This has been interpreted in conjunction with preliminary EM studies (see 2.5.3.1) to indicate that a large oligomer (>8 up to 16 subunits) is the active state of activase [195] [196]. This would be inconsistent with the typical active oligomer of a AAA+ protein, which is a hexamer [196]. Indeed, recent studies of the oligomerization state of a ATP hydrolysis deficient activase mutant (R294A) by mass spectrometry show a complex of Rca in the presence of Mg^{2+} and ATPγS with a mass corresponding to a hexamer [197].

Cysteine mutants in the N- and C-terminal domains of Rubisco activase coupled with bis-maleidoethane crosslinking, resulted in highly crosslinked species of cotton activase, implicating contacts between the N- and the C-terminal subunits in the oligomerization process [198].

2.5.3.1 Preliminary EM studies and structure of Rubisco activase

A preliminary negative stain EM study of a potential complex between Rubisco activase and Rubisco from tobacco in the presence of Mg^{2+} and ATPγS has been reported [199] in which the presence of Rubisco activase seems to increase the diameter of the Rubisco particles. This was interpreted as Rubisco activase surrounding Rubisco [171] [196], maintaining the four-fold symmetry of the Rubisco complex. Although the AAA+ chaperone RavA (a member of the MoxR subfamily) also surrounds its substrate mediated by an interaction domain [200], this arrangement is highly unusual for an interaction of a classical AAA+ protein with its substrate.

While this thesis was written and a manuscript was prepared for publication, the crystal structure of the C-terminal 4-helix bundle domain of activase from Creosote bush was solved at a resolution of 1.88 Å [201]. These results are compared to the results obtained in this study in section 5.1. Briefly, the structure comprises the C-terminal 4-helix bundle domain, with an extended third helix of the bundle.

This extended conformation seems to be stabilised by inter- and intramolecular contacts in the crystal, as discussed in section 5.1.

2.6 Aim of this study

Rubisco activase is a key enzyme of photosynthesis in eukaryotes and required for maintaining the activity of Rubisco. Indeed, activase has been dubbed a "catalytic chaperone" [171]. Furthermore, Rubisco activase is a heat labile enzyme [159] [160], a property that might limit the photosynthetic capacity of plants under moderate heat stress. It has been predicted that enhancing activase stability results in plants with improved photosynthetic properties and growth under conditions of moderate heat stress [161] [162] [163]. However, the mechanism of this AAA+ remodeling chaperone has remained elusive in the absence of a three dimensional structure. Furthermore, the oligomeric state of the active enzyme is still unknown.

The aim of this study, therefore, was the determination of the structure of Rubisco activase by X-ray crystallography. Subsequently, structure guided mutations coupled with negative stain electron microscopy should elucidate the mechanism of Rubisco reactivation as well as as the oligomeric state of the active enzyme. In addition, a detailed characterization of the interaction of Rubisco with Rubisco activase was undertaken.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Unless otherwise stated, chemicals were of *pro analysis* quality and purchased from Sigma-Aldrich (Steinheim, Germany) or Merck (Darmstadt, Germany). RuBP used in Rubisco activity assays was purchased from Sigma-Aldrich (Steinheim, Germany).

Table 3.1: Chemicals

Supplier	Chemical
Amersham Biosciences (Freiburg, Germany)	$Na^{14}CO_3$
Bio-rad (Munich, Germany)	Bradford assay dye reagent,
	AG1-X8 resin Cl^- form
Biozym (Hessisch Oldendorf, Germany)	Biozym LE Agarose
Difco (Heidelberg, Germany)	Bacto tryptone, Bacto yeast extract, Bacto-Agar
Fermentas (St. Leon-Rot, Germany)	GeneRuler 1 kb DNA Ladder
Invitrogen (Karlsruhe, Germany)	dNTP set, protein marker for SDS-PAGE,
	SYBR Safe DNA gel stain
	NuPAGE BT 4-16 % gradient gel
	NuPAGE Novex colloidal blue
J.M Gabler Saliter GmbH & Co. KG (Obergünzburg, Germany)	skim milk
MPI of Biochemistry, Peptide Synthesis Service (Martinsried, Germany)	oligopeptides

PEQLAB (Erlangen, Germany)	IPTG
Qiagen (Hilden, Germany)	Ni-NTA Agarose
Roche (Basel, Switzerland)	DTT
Roth (Karlsruhe, Germany)	Ampicillin, Rotiszint
Serva (Heidelberg, Germany)	Acrylamide-Bis, PMSF, SDS,
	Serva Blue R
Usb (Wooburn Green, UK)	ECL solutions

3.1.2 Enzymes, antibodies and kits

Table 3.2: Enzymes, antibodies and kits

Supplier	Reagent
Hampton Research (Aliso Viejo / CA, USA)	Crystallization screens and supplies
Jackson Immuno Research (Newmarket, UK)	HRP-coupled secondary antibodies
JPT Peptide Technologies GmbH (Berlin, Germany)	PepSpot Peptides Epitope mapping
Merck (Darmstadt, Germany)	Benzonase
New England Biolabs (Frankfurt am Main, Germany)	Restriction enzymes, T4 DNA Ligase
Promega (Mannheim, Germany)	PureYield Plasmid Midiprep System,
	Wizard Plus SV Minipreps
	DNA Purification System,
	Wizard SV Gel and PCR Clean-Up System,
	pGemTeasy vector system,
	Pfu Polymerase
Sigma-Aldrich (Steinheim, Germany)	anti-FLAG M2 mouse monoclonal antibody
Stratagene (Heidelberg, Germany)	Herculase Enhanced DNA Polymerase

3.1.3 Instruments

Table 3.3: Instruments

Supplier	Reagent
Abimed (Langenfeld, Germany)	Gilson Pipetman (2, 10, 20, 100, 200, 1000 μl)
Bachofer (Reutlingen, Germany)	Hybridization Oven
Beckman Coulter	centrifuges and rotors (CS-6R, J6-MI, Avanti)
	spectrophotometers (DU640, DU800),
	LS6500 multi-purpose scintillation counter
Bibby Scientific (Stone, UK)	Stuart rotating wheel
Biometra (Göttingen, Germany)	PCR thermocycler
Bio-Rad (München, Germany)	Chemidoc XRS,
	Mini Protean II electrophoresis cell,
	Power Pac 300,
	Horizontal agarose gel electrophoresis
	(Wide) Mini-SUB CELL GT
Drummond Scientific (Broomall , USA)	Pipet-aid
Eppendorf (Hamburg,Germany)	centrifuges (5415D, 5417R),
	Thermomixer comfort
Fisher Scientific (Schwerte, Germany)	Accumet Basic pH meter
GE Healthcare (München, Germany)	Äkta Purifier, Äkta Explorer,
	Ettan LC, prepacked chromatography columns
Hampton Research (Aliso Viejo, USA)	Crystallization supplies and tools
Hoefer Scientific Instruments (San Francisco, USA)	SemiPhor blotting systems
Mettler Toledo (Giessen, Germany)	Balances (AG285, PB602)
MicroCal (Northampton, USA)	MicroCal VP-ITC MicroCalorimeter
Millipore (Bedford, USA)	Amicon ultra centrifugal filter,
	Steritop vacuum filters
Misonix (Farmingdale, USA)	Sonicator 3000
New Brunswick Scientific	Innova 44 incubator / shaker
PEQLAB (Erlangen, Germany)	Nanodrop 1000

Roth (Karlsruhe, Germany))   ZelluTrans dialysis membrane

Scientific Industries, Inc. (Bohemia / NY, USA)   Vortex-Genie 2

Whatman GmbH (Dassel, Germany)   Whatman Protran nitrocellulose membrane

Wissenschaftlich Technische Werkstätten   pH meter

WTW (Weilheim, Germany)

3.1.4 Buffers and media

Media were prepared with deionised, double-distilled water and subsequently autoclaved. M9 minimal media was made with autoclaved or sterile filtered ingredients as indicated.

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
M9 minimal medium:	775 ml H_2O (autoclaved)
	20 ml 20 % (w/v) glucose (sterile-filtered)
	2 ml 1 M $MgSO_4$ (autoclaved)
	0.1 ml 1 M $CaCl_2$ (autoclaved)
	200 ml 5x M9 salts (autoclaved)
5x M9 salts:	64 g/l $Na_2HPO_4 \cdot 7H_2O$
	15 g/l KH_2PO_4
	2.5 g/l NaCl
	5 g/l NH4Cl

Buffers for protein purification were vacuum filtered through a Steritop filter membrane.

Coomassie destaining solution	10 % ethanol, 7 % acetic acid
Coomassie staining solution	0.1 % (w/v) Serva Coomassie Blue R-250, 40 % ethanol,
	8 % acetic acid

5x DNA loading dye	0.1 M EDTA, 0.1 % (w/v) SDS,
	0.025 % (w/v) bromphenol blue, 40 % glycerol
Electrophoresis migration buffer	120 mM Tris-HCl, 134 mM glycine,
	0.1 % SDS
Extraction buffer	50 mM Tris-HCl pH 7.4, 20 mM MgSO$_4$,
	20 mM NaHCO$_3$, 0.1 mM Na$_2$EDTA, 50 mM β-mercaptoethanol,
	1 mM PMSF
TAE	40 mM Tris-acetate, pH 8.5, 2 mM EDTA
TBS	50 mM Tris-HCl pH 8.0, 137 mM NaCl, 2.7 mM KCl
TBS-T	50 mM Tris-HCl pH 8.0, 137 mM NaCl, 2.7 mM KCl, 0.1 % Tween 20
PBS	12 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ / pH 8.0, 137 mM NaCl, 2.7 mM KCl
PBS-T	12 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 137 mM NaCl, 2.7 mM KCl, 0.1 % Tween 20
Ponceau S staining solution	0.1 % (w/v) Ponceau S, 5 % acetic acid
Pulldown buffer	50 mM Tris-HCl pH 8.0, 50 mM NaCl,
	50 mM imidazole, 5 mM MgCl$_2$
Purification buffer A	50 mM Tris-HCl pH 8.0,
	300 mM NaCl, 10 mM imidazole,
	5% glycerol
Purification buffer B	50 mM Tris-HCl pH 8.0,
	300 mM NaCl, 200 mM imidazole
Purification buffer C	50 mM Tris-HCl-HCl pH 8.0,
	10 mM NaCl

Regeneration buffer	62.5 mM Tris-HCl pH 6.7, 2 % (w/v) SDS,
	100 mM β-mercaptoethanol
Rubisco storage buffer	10 mM NaH$_2$PO$_4$ pH 7.6, 50 mM NaCl,
	1 mM EDTA, 10 % glycerol
5x SDS-sample buffer	225 mM Tris-HCl pH 6.8, 50 % glycerol
	5 % SDS (w/v), 5 % β-mercaptoethanol
	0.01 % bromphenolblue (w/v)
SEC buffer	20 mM Tris-HCl pH 8.0,
	50 mM NaCl, 2 mM MgCl$_2$
Western blot transfer buffer	50 mM Tris-HCl, 20 % methanol, 192 mM glycine

3.1.5 Strains

Table 3.4 gives an overview of used bacterial strains.

Table 3.4: Enzymes, antibodies and kits

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

3.1.6 Plasmids and oligonucleotides

Oligonucleotides used in this study are listed in Table 3.5. Oligonucleotides were purchased from Metabion (Martinsried, Germany).

Table 3.5: Primers used in this study

name	sequence
CrRca2for	CTCGGCGGTGGTGTGGCCTCTAGCCGCAAGCAGATG
CrRcarev	CTAGGATCCTTAGCGGCTGTAGGCCTCGGGCAG

AtRca2for	CTCCGCGGTGGTGTGAAAGAAGACAAACAAACCGTC
AtRca1rev	CTAGGATCCTCAAAAGTTGTAGACACAGGTTCC
Ats2rev	CTAGGATCCTTACTTGCTGGGCTCCTTTTC CTCTTTTTTCCGTAGAAAGTTCCACGGCC
Atlqcforw	GCCATCGGCCGTGGAACTTTCTAC
Atlqcrev	GTAGAAAGTTCCACGGCCGATGGC
NttruncN	CTCCGCGGTGGTAACTTGGACAACAAGTTG
T7term	GCTAGTTATTGCTCAGCGG
NtAct_Nterm_SacIIforw	CTCCGCGGTGGTGAACAAATA
NtAct_Nterm_EcoRIrev	CTAGAATTCCTAGTACTGGCGAAGACC
Pept_for	TAACTCCGCGGTGGTAGCGGCAGCGGCAGCAACGGCC TGCTGCTGCATATT
Pept_rev	CTACTAGAATTCTTACGCATGCATCGCGCGATGAATA TGCAGCAGCAGGCC
HueFLAGfor	CTGGTGTTGCGCCTCCGGGGTGGAATGGATTACAAAGACGAT GACGATAAAGCGGGCCGCGGTGGAGAACAAATAGAT
HueFLAGrev	ATCTATTTGTTCTCCACCGCGGCCCGCTTTATCGTCATCGTCTT TGTAATCCATTCCACCCCGGAGGCGCAACACCAG
W16Afor	AGTGACAGAGCGAAGGGTCTT
W16Arev	AAGACCCTTCGCTCTGTCACT
K92Dfor	GTTCACATCACCGATAACTTCTTGAAA
K92Drev	TTTCAAGAAGTTATCGGTGATGTGAAC
E138Afor	ATGAGTGCTGGAGCGTTGGAAAGTGGA
E138Arev	TCCACTTTCCAACGCTCCAGCACTCAT
E140Afor	GCTGGAGAATTGGCGAGTGGAAATGCA
E140Arev	TGCATTTCCACTCGCCAATTCTCCAGC
S141Afor	GGAGAATTGGAAGCGGGAAATGCAGGA
S141Arev	TCCTGCATTTCCCGCTTCCAATTCTCC
G142Vfor	GAATTGGAAAGTGTGAATGCAGGAGAG

G142Vrev	CTCTCCTGCATTCACACTTTCCAATTC
N143Afor	GAAAGTGGAGCTGCAGGAGAG
N143Arev	CTCTCCTGCAGCTCCACTTTC
A144Vfor	GAAAGTGGAAATGTGGGAGAGCCAGCC
A144Vrev	GGCTGGCTCTCCCACATTTCCACTTTC
Y188Afor	GGAACTACCCAAGCGACTGTCAACAAC
Y188Arev	GTTGTTGACAGTCGCTTGGGTAGTTCC
F232Afor	ACTGGTAACGATGCGTCCACATTGTAT
F232Arev	ATACAATGTGGACGCATCGTTACCAGT
Y236Afor	TTCTCCACATTGGCGGCTCCACTTATC
Y236Arev	GATAAGTGGAGCCGCCAATGTGGAGAA
R294Vfor	TTCGGTGCACTGGTAGCGAGAGTATAC
R294Vrev	GTATACTCTCGCTACCAGTGCACCGAA
R294Afor	TTCGGTGCACTGGCGGCGAGAGTATAC
R294Arev	GTATACTCTCGCCGCCAGTGCACCGAA
D299Afor	GCGAGAGTATACGCGGATGAAGTAAGG
D299Arev	CCTTACTTCATCCGCGTATACTCTCGC
D299Kfor	GCGAGAGTATACAAAGATGAAGTAAGG
D299Krev	CCTTACTTCATCTTTGTATACTCTCGC
ΔC(360)for	TTGGCTGACAAATAATAATACCTCAAAGAG
ΔC(360)rev	CTCTTTGAGGTATTATTATTTGTCAGCCAA
Y361Afor	TTGGCTGACAAAGCTCTCAAAGAGGCT
Y361Arev	AGCCTCTTTGAGAGCTTTGTCAGCCAA
ΔC(363)for	AAATACCTCAAATAGTAAGCACTTGGTGAT
ΔC(363)rev	ATCACCAAGTGCTTACTATTTGAGGTATTT

The expression plasmids pHUE, pHUENtRca and pHUsp2-cc [202] were a gift from Dr. S. Whitney (ANU, Australia).

3.2 Molecular biology methods

3.2.1 DNA analytical methods

DNA concentrations were determined by UV spectroscopy measured on a Nanodrop 1000 (PEQLAB). 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 is approximately 1.85 [203].

Authenticity of cloned constructs was confirmed by test restriction with suitable restriction enzymes and sequencing by the MPI sequencing service. For one sequencing reaction, approximately 5 μl of 100 ng/μl plasmid were needed. Sequencing chromatograms were visualised with the help of the program ApE (M. Wayne Davis, http://biologylabs.utah.edu/jorgensen/wayned/ape), and comparative sequence analysis was performed by alignment in this program.

Agarose gel electrophoresis was performed in 1 % TAE-agarose gels at 7-10 V/cm supplemented with SYBR-SAFE (1:5000) in 0.5x TAE buffer. Prior to loading samples on the gel, they were mixed with 5x DNA loading dye.

3.2.2 Competent *E. coli* cells and transformation

Chemically competent DH5α and BL21 cells were prepared by the CaCl$_2$ - method. Cells were grown to mid-log phase (OD 600 = 0.5) at 37°C in one liter of LB medium. After chilling the cells on ice for 10 min, the culture was centrifuged for 30 min at 3300 xg at 4°C. The supernatant was removed and the cells were resuspended gently in 20 ml ice-cold 0.1 M CaCl$_2$, 15 % glycerol. After incubation on ice for 20 min, 50 μl of the cell suspension were aliquoted in chilled sterile reaction tubes, frozen in liquid nitrogen and stored at -80°C. For transformation, 50 μl of chemically competent cells were thawed on ice and mixed with ~100 ng plasmid DNA. After incubation on ice for 30 min, cells were heat shocked at 42°C for 90 s and allowed to cool down before 200 μl LB-medium were added. Cells were incubated with shaking at 37°C for 30 min before cells were streaked out on LB-Agar plates supplemented with 200 μg/ml ampicillin. A single colony of the transformants was picked and inoculated in liquid medium for further analysis.

3.2.3 Plasmid and DNA-fragment purification

For the amplification of plasmid DNA *E. coli* DH5α were grown for 8-16 h in LB medium containing the appropriate antibiotics at 37°C with moderate shaking [203]. Plasmids were subsequently purified

via anion exchange chromatography using the Wizard Plus SV Miniprep DNA Purification System, according to the manufacturer's instructions. In order to purify or isolate DNA-fragments after agarose gel electrophoresis or enzymatic reactions the Wizard SV Gel and PCR Clean- Up system was used.

3.2.4 Polymerase chain reaction (PCR)

Amplification of DNA fragments was achieved by PCR. To amplifiy genes from plasmids a high fidelity polymerase mix named Herculase (Agilent) was used, while Pfu polymerase (Promega) was used for the insertion of mutations by quick change PCR.

3.2.4.1 DNA amplification by PCR

DNA amplification from plasmids or annealed oligonucleotides was carried out in the following reaction and with the temperature cycles shown in Table 3.6. Primers used shown in Table 3.5.

PCR reaction :   2 μl Herculase 10x buffer

2 μl 2 mM dNTP

0.2 μl (1.2 U) Herculase Polymerase mix

1-30 ng template DNA

0.5 μl 10 μM forward primer

0.5 μl 10 μM reverse primer

total volume 20 μl

Table 3.6: Temperature cycles for DNA amplification by PCR

Temperature (°C)	time (min)	cycles
95	2	1
95	0.3	
50	0.5	35
72	1 min/kb	
72	10	1
4	∞	1

For cloning in pGemTeasy the amplified DNA was desoxyadenosine (dA) tailed by addition of 0.2 μl

Taq polymerase and 2 μl 2 mM dATP and incubation at 72°C for 30 min.

3.2.4.2 Site directed mutagenesis

For site directed mutagenesis of existing plasmids the following PCR reactions and cycle conditions were used. Primers were self complementary and contained the desired mutation (see Table 3.5). After the PCR 0.5 μl DpnI were added to 10 μl of the reactions and incubated at 37°C for 30 min to digest the methylated template plasmid before transformation into *E. coli* DH5α.

PCR reaction :   2 μl Pfu 10x buffer

2 μl 2 mM dNTP

0.38 μl (1.2 U) Pfu Polymerase

300 ng template plasmid

0.72 μl 10 μM forward primer

0.72 μl 10 μM reverse primer

total volume 20 μl

Table 3.7: Temperature cycles for DNA amplification for site directed mutagenesis

Temperature (°C)	time (min)	cycles
95	2	1
95	0.5	
55	1	18
68	10	
68	15	1
4	∞	1

3.2.5 Restriction digest, ligation and cloning strategy

3.2.5.1 Restriction digest and ligation

500 ng - 2 μg DNA were digested with 20-60 U of the respective restriction enzymes in 20-50 μl reactions in buffer according to manufacturers recommendations for 3 h at 37°C. Digested plasmid vector DNA was dephosphorylated with calf intestinal phosphatase (CIP) according to the manufacturer's instructions

prior to further use. Linearized plasmids, as well as plasmid inserts were separated by agarose gel electrophoresis. For further use linearised plasmids or plasmid inserts were purified by the Wizard Plus SV Miniprep DNA Purification System.

Purified insert and plasmid DNA eluted in deionized water were used for ligation. Approximately 100 ng of digested plasmid DNA (pHUE) and an approximately five-fold excess of complementarily digested DNA insert were incubated with 200 U of T4 DNA ligase in a total volume of 10 μl. The reaction was performed in T4 DNA ligase reaction buffer for 1-12 h at room temperature (RT). Afterwards the reaction mix was transformed in *E. coli* DH5α cells. Successful ligation was confirmed by restriction analysis and subsequent DNA sequencing.

The pGemTeasy vector system (Promega) was used according to manufacturers instructions, for ligation of dA tailed PCR products.

3.2.5.2 Precipitation of DNA for sequential restriction enzyme digests

To avoid loss of DNA as a result multiple agarose gel electrophoresis purifications, the DNA was precipitated after treatment with the first restriction enzyme as follows:

5 μl 3 M ammoniumacetate (pH 7) and 100 μl 100 % ethanol were added to 50 μl of the respective restriction enzyme reaction, mixed and incubated on ice for 20 min. The precipitated DNA was spun down at 4°C at 16100 xg for 10 min, the supernatant was removed and the pellet was washed with 300 μl 70 % Ethanol. After the pellet was spun down again, the supernatant was removed and the pellet was dried in a centrifugal evaporator (Concentrator 5301, Eppendorf). For further restriction enzyme digestion the pellet was redissolved in 25 μl of deionized water.

3.2.5.3 General cloning strategy

In this study the expression vector pHUE was used for the heterologous expression of protein constructs in *E. coli* [202]. Proteins are expressed with an N-terminal fusion to hexa-histidine tagged ubiquitin (His$_6$Ub), which improves solubility and enables easy protein purification by immobilised metal-ion affinity chromatography (IMAC). Cleavage of the tag by the deubiquitinating enzyme Usp2 leaves the native N-terminus of the protein of interest.

DNA was amplified from plasmids or annealed oligonucleotides with a 5' primer with the 5'-overhang sequence CTC CGC GGT GGT coding for a SacII restriction site as well as the C-terminal amino acids

of ubiquitin [202]. The 3'primer was designed with a 3'-overhang encoding for a BamHI or EcoRI restriction site, both of which are found in the 3' multiple cloning site of the vector pHUE (primers see Table 3.5). For sequencing and subsequent cloning the dA tailed constructs were first ligated into pGemTeasy via the desoxythymidine (dT) overhang of the linearized vector according to manufacturers instructions. Transformation into *E. coli* DH5α and subsequent colony selection, growth and isolation yielded enough plasmid with high quality for sequencing and final cloning into the expression vector pHUE.

pGemTeasy with inserted sequences were digested with SacII and EcoRI or BamHI either simultaneously or sequentially and the purified inserts were ligated into the pHUE vector digested with the same enzymes and dephosphorylated by CIP. After transformation in DH5α, ampicillin resistant colonies were picked, grown in liquid medium supplemented with ampicillin and the correct insert was confirmed by sequencing. Mutations in the inserts were generated by site directed mutagenesis, unless indicated otherwise.

The coding sequence of *Chlamydomonas reinhardtii* Rubisco activase (CrRca) was amplified with the primer pair CrRca2for and CrRcarev from the plasmid Rca-pET11a (C.Liu unpublished), cloned into the pGemTeasy vector and sequentially digested with SacII and BamHI for cloning into the vector pHUE. The expression plasmid was named pHUECrRca.

The coding sequence of the short isoform of *Arabidopsis thaliana* Rubisco activase (AtsRca) was amplified with the primer pair Atrca2for and Ats2rev from the plasmid pProEx(AT_RCAK9) (A. Schmeinck unpublished), cloned into the vector pGemTeasy and sequentially digested with SacII and BamHI and inserted into the pHUE vector. The expression plasmid was named pHUEAtsRca.

The coding sequence of the long isoform of *Arabidopsis thaliana* Rubisco activase (AtlRca) was amplified with the primer pair Atrca2for and AtRcalrev from the plasmid pProEx(AT_RCAg1) (A. Schmeinck unpublished) and cloned into the vector pGemTeasy. The internal SacII site of AtlRca was removed with site directed mutagenesis with the primer pair Atlqcfor and Atlqcrev and sequentially digested with SacII and BamHI and inserted into the vector pHUE. The expression plasmid was named pHUEAtlRca.

The coding sequence of the N-terminally truncated Rubisco activase (residues 68-383) from *Nicotiana tabacum* (ΔN NtRca) was amplified with the primer pair NttruncN and T7term from the plasmid pHUENtRca [202]. The PCR product was subsequently directly digested with SacII and EcoRI and ligated into the vector pHUE. The expression plasmid was named pHUEΔN NtRca.

The coding sequence of the N-terminal domain of Rubisco activase (residues 1-67) from *Nicotiana*

tabacum (Nterm) was amplified with the primer pair NtAct_Nterm_SacIIforw and NtAct_Nterm_EcoRIrev from the plasmid pHUENtRca [202]. The PCR product was subsequently dA tailed, cloned into the vector pGemTeasy and subsequently digested with SacII and EcoRI and inserted into the vector pHUE.

Mutations in pHUENtRca or pHUEΔNNtRca were generated by site directed mutagenesis with the appropriate primer pairs (see Table 3.5). The expression vector pHUEFLAGNtRca for the expression of N-terminally FLAG tagged proteins was generated from pHUENtRca by site directed mutagenesis with the primer pair HueFLAGfor and HueFLAGrev by Oliver Müller-Cajar (unpublished). Other pHUE-FLAG expression vectors were generated by restriction digest with SacII and EcoRI and ligation with complimentary inserts.

The coding sequence of a peptide motif of the Rubisco large isoform (sequence NGLLLHIHRAMHA) was amplified together with an N-terminal serine glycine linker (sequence SGSGS) by primer overlap extension PCR with the partially complementary primers Pept_for and Pept_rev. The PCR product was dA tailed, cloned into the vector pGemTeasy and subsequently digested with SacII and EcoRI and inserted into pHUE.

3.3 Protein biochemistry

In the following sections all protein biochemical methods used in this study will be summarized.

3.3.1 Protein analytical methods

3.3.1.1 Protein sequence alignment

Protein sequences were aligned using the programs ClustalW [204] or MultAlign [205]. Color coded figures of alignments with indications of secondary structure elements were generated with the program Espript [206].

3.3.1.2 Protein quantification

Protein mixtures were quantified by a Bradford (BioRad protein assay) assay after calibration against known amounts of bovine serum albumin. 5 μl of appropriately diluted protein solution (1:5 to 1:20) were added to 795 μl ddH$_2$O and 200 μl Bradford reagent, mixed and incubated for 5 min before the absorbance of the sample at 595 nm was measured on a DU640 photometer (Beckman Coulter).

Pure proteins were quantified by their absorbance at 280 nm using the extinction coefficients listed in Table 3.8 measured on a Nanodrop 1000 (PEQLAB). Extinction coefficients were calculated with the program ProtParam [207]. Extinction coefficients of point mutations not listed were the same as for the respective wildtype NtRca.

Table 3.8: Extinction coefficients used for proteins in this study

Protein	extinction coefficient (M^{-1} cm^{-1})	molecular weight (Da)
NtRca	39880	42759
ΔN68 NtRca	29910	35161
ΔC360 NtRca	38390	40418
ΔC363 NtRca	39880	40822
ΔN68 ΔC360 NtRca	28420	32820
NtRca W16A	34380	42644
N-term	9970	7616
His_6Ub	1490	10597
His_6Ub-Peptide	1490	12568
NtRubisco (active site)	113220	67222

3.3.1.3 SDS-PAGE

In order to analytically separate proteins based on apparent molecular weight, discontinuous sodium-dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) [208] was used. Gels were prepared as described in Table 3.9, furthermore, isopropanol was layered over the separating gel while the gel-matrix polymerised. Samples were mixed with 5x SDS-sample buffer, boiled for 5 min at 95 °C, and centrifuged for 1 min at 4 °C at 16100 xg. Gel electrophoresis was carried out in Mini Protean (Bio-rad) electrophoresis chambers in the presence of electrophoresis migration buffer employing a starting constant voltage of 100 V followed by 140-160 V for the rest of the run.

Samples for subsequent mass spectrometry were run on precast NuPAGE BT 4-16 % gradient gels (Invitrogen) according to manufacturers instructions at 200 V for 1 h 10 min.

Table 3.9: SDS-polyacrylamide gel preparation

Solutions (4 gels)	6 % gel separating	12.5 % gel separating	16 % gel separating	4 % gel stacking
deionized water	10.8 ml	6.4 ml	4.1 ml	6.1 ml
1.5 M Tris-HCl pH 8.8	5 ml	5 ml	5 ml	-
0.5 M Tris-HCl pH 6.8	-	-	-	2.5 ml
10 % (w/v) SDS	200 μl	200 μl	200 μl	100 μl
30 % (v/v) acrylamide (AA) 0.8 % bis-AA	4 ml	8.4 ml	10.7 ml	1.3 ml
10 % (w/v) ammonium per-oxo-disulfat	100 μl	100 μl	100 μl	50 μl
TEMED	6 μl	6 μl	6 μl	10 μl

3.3.1.4 Coomassie blue staining of SDS-polyacrylamide gels

Coomassie blue dye integrates with cationic, nonpolar, and hydrophobic side chains of a polypeptide. Coomassie blue staining was used for the visualization of protein bands (>500 ng) in polyacrylamide gels. Protein bands were fixed and stained in the gel matrix by incubating the gels for 1 hr with Coomassie blue staining solution. Background staining was removed by incubation with Coomassie destaining solution for 1-3 h.

Precast NuPAGE gradient gels for mass spectrometry were stained with NuPAGE Novex colloidal blue staining solutions according to the manufacturers instructions.

3.3.1.5 Western blotting and immunodetection

Western blot followed by specific immunodetection [209] was employed for specific identification of protein bands after SDS-PAGE. Proteins were transferred to a nitrocellulose membrane in a semidry western blot (SemiPhor) in transfer buffer at a current of 40 mA per gel for 1 h. Transferred proteins were stained with Ponceau S to ensure efficient transfer and destained with water before proceeding with the immunodetection.

Nitrocellulose membranes were blocked for 1 h at RT in PBS-T with 5 % (w/v) milk powder. The primary antibody anti-FLAG M2 (mouse monoclonal, Sigma) was diluted 1:5000 in PBS-T with 5 % (w/v) milk powder and incubated with the membrane for 1 h at RT. The membrane was washed twice

with PBS-T and subsequently incubated in PBS-T for 15 min. PBS-T was exchanged three times and the membrane incubated in fresh PBS-T for 5 min. The secondary antibody peroxidase-conjugated goat anti-mouse (Jackson Immuno Research) was diluted 1:10000 in PBS-T with 5 % (w/v) milk powder and incubated with the membrane for 1h at RT. The membrane was washed again (as above) and bound HRP coupled antibodies were detected by enhanced chemoluminescence (ECL) staining with a 1:1 mix of Rodeo ECL detection reagent 1 and Rodeo ECL detection reagent 2 (usb). Chemiluminescence was detected on a Chemi-Doc XRS (Bio-rad) gel documentation system.

3.3.1.6 Circular Dichroism (CD) -spectroscopy

Structural integrity of purified Rca and Rca mutants was analysed by CD-spectroscopy. Proteins were desalted into buffer (20 mM K_2HPO_4/KH_2PO_4 pH 8) using the Bio-Spin6 chromatography column (Bio-Rad) according to manufacturers instructions. Protein concentration was adjusted to a final concentration of 0.1 mg/ml and the CD spectra were recorded on a JASCO CD-spectrometer (average of 4 passes, 250 nm to 190 nm, 0.1 nm data pitch, 20°C). CD-spectra were fitted with the CONFIT algorithm on the SMP56 base dataset to estimate the secondary structure content. Structural integrity of point mutant proteins was verified by CD

3.3.2 Test expression of proteins

E. coli BL21 cells carrying the pHUE plasmid for expression of the respective proteins were grown in LB supplemented with 200 μg/ml ampicillin at 37°C in an Innova44 incubator (New Brunswick Scientific) at moderate shaking (220 rpm) overnight. 5 ml LB media supplemented with 200 μg/ml ampicillin was inoculated with 50 μl of the overnight culture and grown at 37°C to an optical density at a wavelength of 600 nm (OD600) of ~0.8. The culture was subsequently shifted to the temperature to be tested (37°C, 23°C, 18°C) and induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). After the indicated times (2 h , 4h or ~12 h) cells were harvested by flash freezing in liquid N_2. Cells were subsequently resuspended in 500 μl lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazol, 5 % glycerol) supplemented with 1 mg/ml lysozyme and 1 mM phenylmethysulfonylfluorid (PMSF), incubated at 4°C for 40 min and lysed on ice by sonication. 80 μl of the lysate were added to 20 μl 5xSDS sample buffer for analysis of total protein expression. 100 μl of the lysate were clarified by centrifugation (16100 xg, 15 min, 4°C) and 80 μl of the supernatant were added to 20 μl 5xSDS sample buffer for analysis

of soluble protein expression. The residual supernatant was discarded and the pellet was washed with 200 μl deionized water and spun down again (16100 xg, 15 min, 4°C). The pellet was redissolved in 125 μl 1xSDS sample buffer for analysis of insoluble protein expression. Samples were analysed by 12.5 % SDS PAGE.

3.3.3 Expression and purification

Unless otherwise stated all protein purification procedures were carried out at 4 °C. All heterologously expressed proteins in this study were expressed in *E. coli* BL21 as a His$_6$Ub-fusion constructs using the pHUE expression vector [202]. *E. coli* BL21 cells carrying the respective pHUE plasmid were grown in LB supplemented with 200 μg/ml ampicillin at 37°C in an Innova44 incubator (New Brunswick Scientific) at moderate shaking (220 rpm) to an optical density at a wavelength of 600 nm (OD600) of ~0.8. The culture was then incubated at 23°C for ~20 min and expression was induced with 0.5 mM IPTG for ~16 h at 23°C. The *E. coli* cells were harvested after this incubation period by centrifugation in J6-Mi centrifuge with JS 4.2 rotor (Beckman Coulter) at 3300 xg for 40 min. Cell pellets were either flash frozen in liquid Nitrogen and stored at -80 °C or directly lysed as described in section 3.3.3.2.

3.3.3.1 Expression of selenomethionine (SeMet) labeled △N68 NtRca

Protein for SeMet labeling was expressed in *E. coli* BL21(DE3) harboring the pHUE△NNtRca plasmid grown in M9-minimal-medium with 200 μg/ml ampicillin at 37°C in an Innova44 incubator (New Brunswick Scientific) at moderate shaking (220rpm) to an OD600 of ~0.8. At this point, the following amino acids were added: 125 μg/ml L-Lys, 100 μg/ml L-Phe, 100 μg /ml L-Tyr, 50 μg/ml L- Ile, 50 μg/ml L-Leu, 50 μg/ml L-Val and 60 μg/ml L-SeMet. Cells were allowed to grow for 15 min to shut down L-Met biosynthesis before the temperature was reduced to 23°C and protein synthesis was induced with 0.5 mM IPTG for ~16 h.

3.3.3.2 Purification of His$_6$Ub-fusion constructs of Rubisco activase

His$_6$Ub-fusion constructs were captured by immobilised metal affinity chromatography (IMAC) and after cleavage of the ubiquitin moiety further purified by anion exchange chromatography.

Harvested cells from 1 l of bacterial culture were resuspended in 20 ml buffer A (50 mM Tris-HCl pH 8.0, 30 mM NaCl, 10 mM imidazole, 5 % glycerol) supplemented with 1 mM PMSF, 1 mg/ml lysozyme,

5 U/ml Benzonase. The cells were disrupted by ultrasonication with a Misonix Sonicator 3000 (20x 20 s bursts,70 W, 75 s cooling on ice between bursts). Following removal of cell debris by centrifugation in an Avanti J-25 with JA25.50 rotor (Beckman Coulter) for 40 min at 4 °C and 32 800 xg, the supernatant was applied to a gravity flow column packed with Ni-NTA resin (Qiagen) equilibrated with buffer A. The His_6Ub-fusion protein was eluted using buffer B (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 200 mM imidazole). The ubiquitin moiety was cleaved overnight at 23°C using the deubiquitylating enzyme Usp2 [202]. The protein solution was dialysed at 4°C in a ZelluTrans 10-12 kDa dialysis bag (Roth) against buffer C (50 mM Tris-HCl pH 8.0, 10 mM NaCl) and applied to a MonoQ (GE) column; proteins were eluted with a linear salt gradient to 500 mM NaCl. Fractions containing Rca were combined and concentrated, 5 % glycerol added and the protein stored in aliquots at -80 °C after flash-freezing in liquid N_2.

Rca for X-Ray crystallographic studies was purified further by gel-filtration (Superdex 200, GE) in buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl). Fractions of the main oligomeric peak were pooled.

Protein purification of the SeMet-labeled protein was performed as for unlabeled protein, except that buffers contained 10 mM β-mercaptoethanol.

3.3.3.3 Purification of other His_6Ub-fusion constructs

Hi_6Ub-fusion constructs of a peptide motif of the Rubisco large subunit of *Nicotiana tabacum* with an N-terminal linker with the sequence SGSGS (His_6Ub-Peptide) were purified by IMAC as described above. Following elution from the column packed with Ni-NTA resin (Qiagen), fractions containing the protein were pooled and dialysed against pulldown buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 50 mM Imidazol, 5mM $MgCl_2$).

His tagged ubiquitin used as control in the pulldown experiments was purified from pHUENtRca, after cleavage of the ubiquitin moiety from Rca by Usp2, the His_6Ub does not bind to the Anion exchange (MonoQ) column and was therefore directly pooled and concentrated from the flowthrough of this column.

For storage glycerol was added to the concentrated samples to a final concentration of 5 %, samples were flash frozen in liquid N_2 and stored at -80°C.

3.3.3.4 Purification of Rubisco from *Nicotiana tabacum* leaves

Leaves from *Nicotiana tabacum* (common tobacco) were a kind gift from Prof. H.-U. Koop and S. Kirchner (LMU Botany, Munich). The *Nicotiana tabacum* Rubisco (NtRbcL$_8$S$_8$) was purified by crystallization from bulk extract essentially as previously described [210].

Briefly 40 g of leaves from *Nicotiana tabacum* were frozen in liquid N$_2$ and plant cells were lysed by manual grinding with mortar and pestle. The fine powder was extracted with 100 ml extraction buffer (50 mM Tris-HCl pH 7.4, 20 mM MgSO$_4$, 20 mM NaHCO$_3$, 0.1 mM Na$_2$EDTA, 50 mM β-mercaptoethanol and 1 mM PMSF) containing 1 g Casein, 10 ml glycerol and 2 g polyvinylpolypyrrolidone (PVPP). The extract was filtered through four layers of Miracloth (Calbiochem) and clarified by centrifugation (27000 xg, 4°C, 10 min). The supernatant was passed through an AG1-X8 Cl$^-$ (Bio-Rad) anion exchange column equilibrated in the extraction buffer. The flowthrough from this column was brought to 10 % PEG 3350 by the dropwise addition of 60 % PEG 3350 in extraction buffer. Precipitated protein was removed by centrifugation (27000 xg, 4°C, 10 min) and the supernatant was brought to 20 % PEG 3350 by the dropwise addition of 60 % PEG 3350 in extraction buffer. After 10 min incubation on ice the precipitated protein was pelleted by centrifugation (27000 xg, 4 °C, 10 min) and thoroughly redissolved in 5-8 ml of the extraction buffer. The protein solution was again cleared by centrifugation (27000 xg, 4 °C, 10 min) to remove any insoluble, precipitated protein. The supernatant was subsequently dialysed against a low salt buffer (25 mM Tris-HCl pH 7.2, 0.1 mM EDTA) for 3-5 days at 4°C in a ZelluTrans 10-12 kDa dialysis bag (Roth) with frequent buffer changes. Rubisco protein crystals were removed by gentle centrifugation (1400 xg, 10 min) and redissolved directly in storage buffer (10 mM Na$_2$HPO4 pH 7.6, 50 mM NaCl, 1 mM EDTA and 10 % (v/v) glycerol) and the protein solution was flash-frozen in liquid N$_2$ and stored at -80°C.

3.3.4 Protein crystallization and structure determination

3.3.4.1 Crystallization

Initial crystal screening was carried out by the crystallization facility at the MPI of Biochemistry. 100 nl protein solution were mixed with 100 nl reservoir solution in sitting drop vapor diffusion at 4°C and RT. The precipitant kits Index (Hampton), crystal platform magic 1 and crystal platform magic 2 (MPI of Biochemistry, crystal facility) were used for initial screening. Initial crystals of full length NtRca were obtained from a selected condition of crystal magic 1 (12-18% PEG 3350, 50 mM Tris-HCl pH 8),

whereas improved crystals of ΔN68 NtRca were obtained from a selected condition of the Index screen.

Full-length NtRca crystals were grown using the hanging drop vapor diffusion method by mixing 2 μl protein sample (5 mg/ml) with 2 μl of reservoir solution at 18°C and equilibrating against a reservoir solution of 500 μl (12-18 % PEG 3350, 50 mM Tris-HCl pH 8).

ΔN68 NtRca and ΔN68ΔC360 NtRca crystals were grown using the hanging drop vapor diffusion method by mixing 2 μl protein sample (5 mg/ml) with 2 μl of reservoir solution at 18 °C and equilibrating against a reservoir solution of 500 μl. Hexagonal rod-shaped crystals were obtained after two to three weeks with a precipitant containing 50 mM MES-Na pH 6.0, 350 mM magnesiumformate. For cryoprotection the original drop was mixed 1:1 with mother liquor (50 mM MES-Na pH 6.0, 450 mM magnesiumformate) containing 35 % glycerol and equilibrated against mother liquor containing 35 % glycerol.

3.3.4.2 Data collection, structure solution and refinement

Diffraction data were collected at the ESRF Grenoble on beamlines ID14-4, ID23-1 and ID23-2 at 100 K. The high resolution dataset of ΔN68ΔC360 NtRca was collected with an additional low resolution pass with a heavily attenuated beam to prevent overloaded pixels on the detector. SeMet incorporation into the crystallised protein was verified by collecting an X-ray fluorescence excitation spectrum in the energy range of the KI absorption edge of selenium (12614-12693 eV). Anomalous signal from SeMet-labeled protein was collected at the maximum SeMet absorption peak at 0,9795 Å and f' and f'' were estimated with the program Chooch [211]. A strategy for data collection was calculated in iMosflm [212] after autoindexing of two wedges of reflection images 90° apart.

Diffraction data were integrated in XDS [213] and POINTLESS, SCALA [214] [215] and Truncate [216] were used to convert the integrated datasets to CCP4 format. High and low resolution pass datasets for ΔN68ΔC360 NtRca were scaled together and merged in SCALA.

The structure of ΔN68 NtRca was solved by single isomorphous replacement with anomalous scattering (SIRAS) from the dataset of native and SeMet labeled protein crystals using direct methods in ShelXD [217]. Nine Selenium sites were found and further refined in SHARP [218]. After density modification in Resolve [219] with an estimated solvent content of ~54 % the resulting density map was readily interpretable and a structural model was manually built using Coot [220]. Density for the residues 361-383 was not discernible. The final model was refined against native data of the crystals of

ΔN68 ΔC360 NtRca using iterative refinement in REFMAC5 [221] and model building in Coot. Non glycine residues facing solvent channels without detectable side chain density were modeled as alanines. The final model contains two Ramachandran outliers according to the program PROCHECK [222]. The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 3T15. Figures were generated with PyMOL [223].

3.3.5 Analytical size-exclusion chromatography

Analytical size-exclusion chromatography of proteins (50 μl of 4.7 μM protomer) was performed on Superdex 200 PC3.2/30 column equilibrated in SEC buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM MgCl$_2$) on an Ettan LC (GE) system (flow 50-100 μl/min). Protein elution from the column was detected by the absorbance at 280 nm. The column was calibrated with globular proteins of known molecular weight (ovalbumin 43 kDa, conalbumin 75 kDa, aldolase 158 kDa, ferritin 440kDa, thyroglobulin 669kDa).

3.3.6 ATPase assay

ATPase activity was assayed spectrophotometrically using a coupled assay monitoring the oxidation of reduced Nicotinamide adenine dinucleotide (NADH) [224] in buffer (100 mM Tricine pH 8.0, 2 mM DTT, 5 mM MgCl$_2$, 2 mM phosphoenolpyruvate, 0.5 mM NADH, 2 mM ATP, 20 U /ml pyruvate kinase, 30 U/ml lactate dehydrogenase) with a final Rca concentration of 2.35 μM Rca protomer (0.1 mg/ml) at 20°C. The oxidation of NADH, which is directly coupled to the generation of ADP, was followed by the reduction in the absorbance at 340 nm on a DU 800 spectrophotometer (Beckman Coulter).

3.3.7 Rubisco activation assay

Rubisco activase activity was assayed with modifications of the established two-step assay [225].

Purified NtRbcL$_8$S$_8$ was incubated at 50°C for 20 min and subsequently inactivated by desalting on a NAP5 (GE) column in a CO$_2$-free, N$_2$-sparged buffer containing 20 mM Tricine pH 8.0, 150 mM NaCl and 0.2 mM EDTA, followed by the addition of 1 mM RuBP to form the Rubisco-RuBP (ER) complex. The ER complex was incubated at RT for 20 min before activation was initiated in buffer containing 100 mM Tricine pH 8.0, 5 mM NaHCO$_3$, 5 mM MgCl$_2$, 50 mM DTT, 1 mM ATP, 3 mM creatine phosphate, 14 U/ml creatine phosphokinase and Rca and/or mutant enzymes at 20°C. Final concentrations in the

assay were 11.9 μM NtRbcL$_8$S$_8$ active sites (0.8 mg/ml) and 4.7 μM Rca protomer (0.2 mg/ml). After various times, aliquots were removed and Rubisco enzyme activity determined by a 1 min Rubisco activity assay. Aliquots were incubated in buffer containing 50 mM Tris-HCl pH 8.0, 20 mM MgCl$_2$, 10 mM NaHCO$_3$/NaH^{14}CO$_3$ (~25 Bq/nmol),2 mM RuBP for 1 min before the reaction was stopped by the addition of 20 % formic acid. Samples were dried down completely to evaporate residual CO$_2$/^{14}CO$_2$ and the amount of fixed ^{14}CO$_2$ was measured by addition of 200 μl deionized water and 1 ml Rotiszint eco plus (Roth) and subsequent ^{14}C-scintillation counting on a LS6500 (Beckman Coulter).

The 100 % activation activity of wildtype NtRca was ~12.5 $\pm$ 2 nmol Rubisco sites/min/mg protein. Mutant enzymes were quantified by comparing the initial rate of activation during the first 5 min to wildtype NtRca.

3.3.8 Disuccinimidyl suberate (DSS)-crosslinking

DSS is a noncleavable crosslinker that contains an amine-reactive N-hydroxysuccinimide ester at each end of an 8-carbon spacer arm. It crosslinks free amines for example at the N-terminus of the protein or at lysine side chains that are within 12 Ådistance of each other. Purified proteins were buffer exchanged to 20 mM HEPES pH 8, 150 mM NaCl, 2 mM MgCl$_2$ using the Bio-Spin6 chromatography column (Bio-Rad) according to manufacturers instructions. Proteins were crosslinked in the presence or absence of the nucleotide analogue ATPγS on ice. The protein concentration, DSS concentration (given as fold-excess over the total protein concentration) and the time for the crosslinking reaction are given in the results section. The crosslinking reaction was stopped by the addition of 1 M Tris-HCl pH 8 to a final concentration of 50 mM. Samples were analysed by 6 % SDS-PAGE and immunoblotting.

3.3.9 Peptide membrane interaction analysis

A custom peptide membrane displaying peptides of the *Nicotiana tabacum* Rubisco large subunit was obtained from JPT Peptide Technologies GmbH, Berlin. This membrane redundantly displays the sequence of the large subunit as tridecapeptides overlapping in 11 residues. The 13 residues long peptides were and N-termially acetylated and C-terminally covalently bound to a cellulose-PEG-membrane by an additional β-alanine. Each of the 232 peptide spots (0.37 cm x 0.37 cm) carried approximately 5 nmol of peptide.

Before the membrane was used for the first time it was rinsed twice in methanol. The membrane

was subsequently incubated three times in TBS for 10 min each at RT. To prevent unspecific protein interactions the membrane was blocked with TBS with 1 % (w/v) milk powder for 3 h. Following another washing step with TBS for 10 min, the membrane was then incubated overnight at 4°C with 10 μg/ml N-terminally FLAG tagged protein (as indicated in the results section) in TBS containing 1 % (w/v) milk powder. Subsequently the membrane was washed thee times in TBS-T for 5 min at RT and incubated in TBS / 1 % (w/v) milk powder containing the primary antibody anti-FLAG M2 (Sigma) diluted 1:2000 for 45 min. Following a washing step with TBS-T the membrane then was incubated for 45 min in a 1:5000 dilution of the secondary antibody peroxidase-conjugated goat anti-mouse (Jackson Immuno research) in TBS / 1 % (w/v) milk powder. The membrane was washed again three times in TBS-T and HRP coupled antibodies were detected by enhanced chemiluminescence (ECL) staining with a 1:1 mix of Rodeo ECL detection reagent 1 and Rodeo ECL detection reagent 2 (Usb). Chemiluminescence was detected on a Chemi-Doc XRS (Bio-rad) gel documentation system.

Controls were performed before each incubation with FLAG-tagged protein following the same protocol but without the FLAG-tagged protein in the overnight incubation to confirm the complete removal of bound proteins (see below) as well as to control for unspecific binding of the primary and secondary antibody to the membrane.

After each incubation and detection the peptide membrane was regenerated by incubation in a buffer containing SDS and β-mercaptoethanol. Following a washing step with water (three times, 10 min) the membrane was incubated in regeneration buffer at 50°C four times for 30 min each. The membrane was subsequently incubated in 10x PBS (three times, 20 min) followed by a washing step in TBS-T (20 min) and TBS (five times, 10 min). This procedure results in the removal of proteins and antibodies bound to the peptides.

3.3.10 Pulldown experiments

200 μl NiNTA resin (Qiagen) resuspension was spun down (10000 xg, 1 min, 4°C) and washed three times with 1 ml of pulldown buffer . The resin was subsequently resuspended in 200 μl of pulldown buffer. Subsequently the resin material was loaded with the bait (2 mg of His_6Ub-Peptide) or the control (2 mg His_6Ub) in pulldown buffer for 30 min at 4 °C on a rotating wheel (Bibby Scientific). After incubation the resin was washed twice with 1 ml pulldown buffer and resuspended in 200 μl of this buffer. Subsequently 1 mg of NtRca was added as prey and nucleotide or nucleotide analogues were

added to a final concentration of 1 mM. The mix was incubated for 2 h at 4°C on a rotating wheel. Subsequently the resin material was washed three times with 400 μl pulldown buffer with the respective nucleotide or nucleotide analogue added to a final concentration of 1 mM. Bound protein was eluted by incubation of the NiNTA resin in 300 μl of buffer B (50 mM Tris-HCl, 300 mM NaCl, 200 mM imidazole). Eluted protein samples were analysed by 16 % SDS-PAGE followed by Coomassie blue staining.

4 Results

4.1 Protein purification

4.1.1 Purification of Rubisco activase

Rubisco activase from *Nicotiana tabacum* (NtRca) was purified according to a previously reported procedure [202] with minor modifications. All Rca constructs and mutants, used in this study, were heterologously expressed in *E. coli* BL21 as a N-terminal fusion with a hexa-histidine tagged Ubiquitin. The presence of the tag improved the solubility of the respective proteins upon expression in *E. coli*.
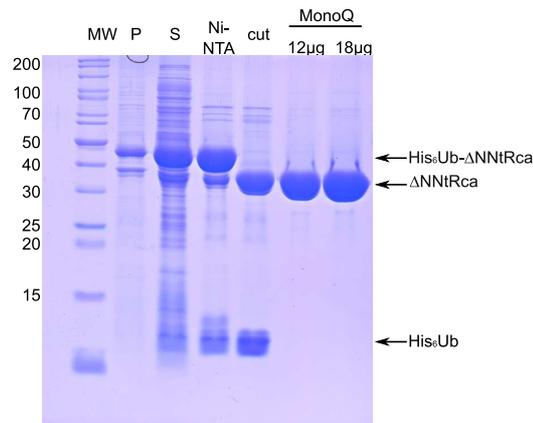
Figure 4.1: **Representative purification of ΔN68 NtRca.**

12.5 % SDS-PAGE analysis. Molecular weight of marker bands and location of the respective Rca bands are indicated. Successive IMAC and anion exchange chromatography yielded highly pure protein (lanes MonoQ). Usp2 mediated cleavage of the N-terminal tag was complete (lane "cut").

MW: molecular weight marker. P: Pellet. S: Soluble. Ni-NTA: Pooled protein fractions after IMAC. Cut: Protein after cleavage of the N-terminal tag by Usp2. MonoQ: Different amounts of pooled protein fractions after MonoQ.

A representative example of the purification of ΔN68 NtRca is presented in Figure 4.1. Immobilised metal ion affinity chromatography (IMAC) using Ni-Nitrilotetraacetic acid (Ni-NTA) Agarose (Qiagen), followed by anion exchange chromatography using a MonoQ column (GE) yielded pure protein as anal-

ysed by SDS-PAGE . Samples for crystallography were further purified by size exclusion chromatography using a Superdex200 GL10/300 (GE). The average yield was ~20-30 mg of pure protein from a 1 l culture of *E. coli* BL21 grown in LB media. CD-spectra of point-mutant proteins were identical to respective spectra of wildtype proteins, indicating that they were stably folded.

4.1.2 Expression tests for Rubisco activase from other species

Rca sequences from other species (*Arabidopsis thaliana, Chlamydomonas rheinhardtii*) were cloned into the pHUE expression vector and the soluble expression of the N-terminally Histidine tagged Ubiquitin fusion protein was assayed. Figure 4.2 shows a representative overview of the expression analysed by SDS-PAGE.
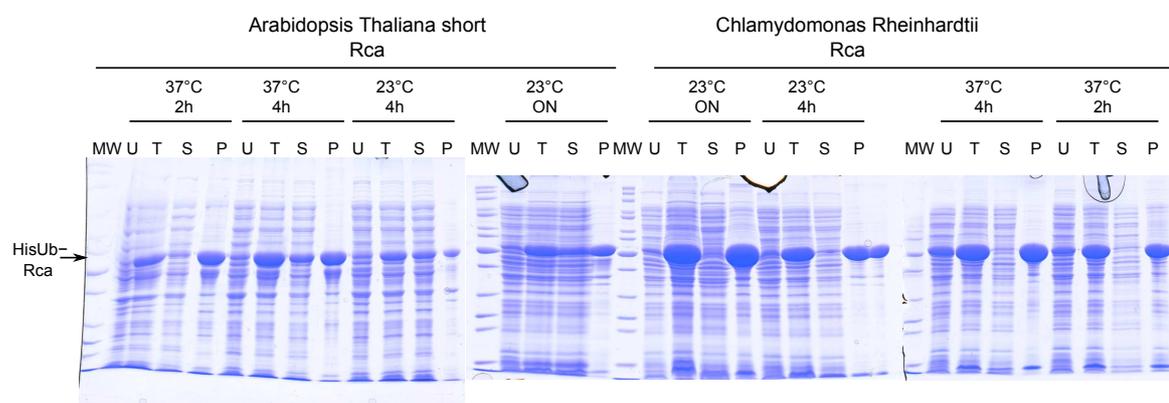
Figure 4.2: **Test-expressions of Rca from other species.**

12.5 % SDS-PAGE analysis. Location of the Rca bands are indicated. The conditions for expressions are indicated above the lanes. Expression was induced with 0.5 mM IPTG and equivalent volumes were loaded for each sample. While *Arabidopsis thaliana* Rca was soluble under the tested conditions, Rca from *Chlamydomonas reinhardtii* was not.

MW: molecular weight marker. U: Total uninduced. T: Total induced. S: Soluble. P: Pellet. ON: overnight

While Rca from *Arabidopsis thaliana* was expressed in soluble form at 37°C and 23°C, Rca from *Chlamydomonas rheinhardtii* was insoluble under all conditions tested (Figure 4.2 Pellet).

4.1.3 Purification of *Nicotiana tabacum* Rubisco

Nicotiana tabacum Rubisco was purified from ~8-10 week old plant leaves by successive polyethyleneglycol (PEG) precipitation and crystallisation in a low salt buffer, as described previously [210]. Figure 4.3

shows a representative overview of the purification. Tobacco Rubisco quantitatively crystallised when dialysed against low salt buffer for 3-5 days (see Figure 4.3) and was confirmed to run as a high molecular weight complex on clear native PAGE (data not shown). *N. tabacum* Rubisco appeared to be pure as analysed by SDS-PAGE (Figure 4.3).The average yield was ~10 mg pure protein from 40 g of fresh leaf material.
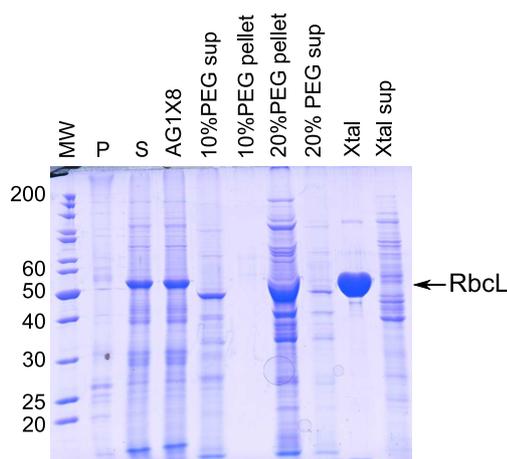

Figure 4.3: **Representative purification of Rubisco.**

12.5 % SDS-PAGE analysis. Molecular weight markers as well as the location of the band for the Rubisco large subunit are indicated. Rubisco crystallised quantitatively from the redissolved 20 % PEG 3350 pellet and appeared to be pure (lanes Xtal and Xtal sup).

MW: Molecular weight marker. P: Pellet after lysis and initial filtration. S: Soluble after lysis. AG1X8: Flowthrough after removal of secondary metabolites by AG1-X8 (Cl⁻ form) column. 10 % PEG sup: Supernatant of precipitation by 10 % PEG 3350. 10 % PEG pellet: Pellet of precipitation by 10 % PEG 3350 (redissolved in an equivalent volume). 20 % PEG pellet: Pellet of precipitation by 20 % PEG 3350 (loaded in an equivalent volume). 20 % PEG sup: Supernatant of precipitation by 20 % PEG 3350. Xtal: Redissolved crystals after dialysis against low salt buffer. Xtal sup: Supernatant of crystallisation.

4.2 Crystal structure of Rubisco activase from *Nicotiana tabacum*

4.2.1 Crystallisation of Rubisco activase

Initial crystals in trials with full length NtRca appeared after 4 weeks of hanging drop equilibration against a mother liquor containing 12-18 % PEG 3350, 50 mM Tris-HCl pH 8. Only clusters of small, needle-like, hexagonal rods could be obtained even after streak seeding or additional screening for suit-

able additives with Hampton additive screen HR2-428 (Figure 4.4 A). To test whether partial degradation of the full length protein in the drop was necessary for formation of crystals, the obtained crystals were taken up in mother liquor, spun down, residual liquid was taken off and the redissolved pellet was loaded on a pre-cast NU-PAGE BT (4-16 %) gradient SDS-PAGE gel together with a fresh full length protein control (Figure 4.5). Indeed, the crystallised protein appeared to be a degradation product of full length NtRca. Bands of the gel were subjected to in-gel tryptic digest and analysed by mass spectrometry (LTQ-Orbitrap) at the Core Facility of the MPI for Biochemistry (Dr. Cyril Boulegue). Additionally, a sample of the crystals was dissolved in MilliQ-H_2O for N-terminal Edman sequencing performed by Reinhard Mentele (AG Lottspeich, MPI Biochemistry).
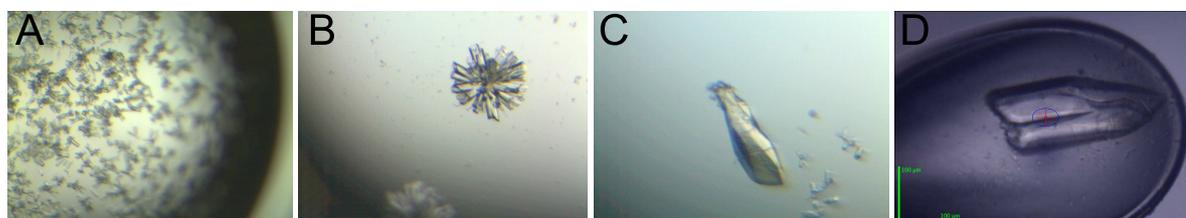

Figure 4.4: **Representative images of protein crystals obtained for different Rca constructs.**
The crystal shown in D diffracted to 2.95Å and produced the dataset used for final refinement of the structure.
A: initial crystals obtained for full length NtRca. B: Representative cluster of hexagonal rods obtained for ΔN68 NtRca. C: Single crystal of ΔN68 NtRca. D: Crystal of ΔN68ΔC360 NtRca mounted at ID23-1 beamline at ESRF Grenoble . Green scalebar corresponds to 100 μm.

Analysis by mass spectrometry indicated the absence of a N-terminal sequence in the detected peptides of the crystallised sample, whereas the peptide coverage of the fresh protein sample was close to 100% (see Figure 4.6).

Edman sequencing confirmed the N-terminus of the crystallised sample to start with QYN or NLD corresponding an N-terminal truncation of 65 or 67 residues respectively. This was consistent with the missing peptide region as analysed by mass spectrometry (see supplementary Figure 4.6).Based on these findings a mutant of NtRca was engineered that lacked the first 67 residues (ΔN68 NtRca, number corresponding to the first amino acid in the construct). Crystals of this construct appeared after several days as clusters of big hexagonal rods in hanging drops equilibrated against a mother liquor containing 0.25-0.45 M magnesium formate, 50 mM MES pH 6 (Figure 4.4 B). Rod like, single crystals could easily be detached from the clusters and diffracted to a resolution of up to 3.3Å (Figure 4.4 C and Table 4.1). The structure could be solved by Se-Met SIRAS (see 4.2.2) and an initial model was built in the resulting
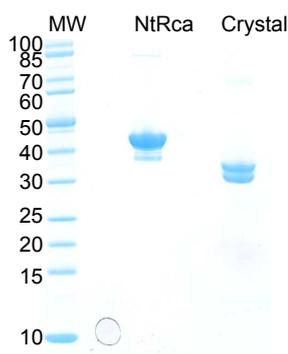
Figure 4.5: **Initial Rca crystals contained an N-terminally truncated fragment of full length Rca.**

NU-PAGE BT (4-16 %) gradient SDS-PAGE gel of full length Rca in comparison with the redissolved crystallised protein. Molecular weights of the marker bands are indicated.

MW: molecular weight marker. NtRca: full length Rca. Crystal: redissolved, crystallised protein

experimental density.

Figure 4.6: **Comparison of peptide coverages of a fresh protein sample and the crystallised sample.**

The fresh protein sample is indicated as NtRca, the crystallised sample as NtRca Xtal. Peptide sequences identified by mass spectrometry colored in green. Sequences in black were not identified. N-terminal sequences not identified in the crystallised sample were interpreted as N-terminal truncation consistent with N-terminal Edman sequencing.

The last 23 C-terminal residues were not resolved in the electron density most likely due to their flexibility. To further improve resolution, a second construct was engineered corresponding to a N-terminal truncation of 67 residues and a C-terminal truncation of 23 residues (ΔN68 ΔC360 NtRca, numbers corresponding to the first and last amino acids in the construct). Isomorphous crystals of this truncated protein were obtained in hanging drops equilibrated against a mother liquor containing 0.25-0.45 M magnesium formate, 50 mM MES pH 6 (Figure 4.4 D) and diffracted to a resolution of up to 2.95Å. The final model for the structure of NtRca was built and refined against a high resolution dataset obtained from these crystals.

4.2.2 Structure solution and refinement

Table 4.1 provides an overview of dataset quality, phasing power and refinement statistics. Rca crystallised in space group $P6_5$, a hexagonal space group with a screw axis. The structure was initially solved by SeMet-SIRAS at a resolution of 3.3Å. The selenium substructure was solved using the ShelX program package (ShelXC, ShelXD) [217] and contained 9 selenium atom positions. The positions were further refined in SHARP [218] and the phasing power was sufficient to result in a readily interpretable map after density modification with Resolve [219]. Models were built iteratively using manual building in Coot [220] and refinement cycles in REFMAC [221]. Final refinement and model building was carried out against a high resolution dataset of ΔN68ΔC360 NtRca (AB3036) that diffracted to 2.95 Å. Due to the steep diffraction intensity decrease with resolution, the high resolution datasets contained many saturated reflections at low resolution. Therefore additional datasets were taken with a separate low resolution pass at lower beam intensity to collect complete low resolution data.

Interestingly, the Wilson B-factor of all collected datasets was unusually high (e.g., 113Å^2 for AB3036), which is reflected in the average B-factor of the protein residues for the final model of 118 Å^2 and the overall limited resolution. Consistent with the thermal instability of the Rca protein, this points to a high degree of flexibility of the protein, which might have hampered previous attempts to elucidate the protein structure.

The final model contained one molecule of ΔN68ΔC360 NtRca per asymmetric unit. Residues 142-144, 177-190, 208-218 and 235-236 were not discernible in the electron density and are probably disordered. Non glycine side chains facing solvent channels were modelled as alanine, when no density for the side chain could be identified. The final model overall displays satisfying geometry (Table 4.1) and only contains two Ramachandran outliers according to the criteria of the program PROCHECK [222].

Table 4.1: Dataset quality, structure solution and refinement

Dataset	AB3036 (native)	AA9291 (native)	AB3005 (Se-Met, anomalous)
construct	ΔN68ΔC360 NtRca	ΔN68 NtRca	ΔN68 NtRca
beamline	ID23-1	ID23-2	ID14-4
wavelength	1.0332	0.8726	0.9795
space group	P6$_5$	P6$_5$	P6$_5$
cell a,b,c (Å)	103.64, 103.64, 56.68	104.27, 104.27, 56.74	104.5, 104.5, 56.59
cell α,β,γ(°)	90, 90, 120	90, 90, 120	90, 90, 120
resolution (Å) *	56.71-2.95 (3.11-2.95)	48.06-3.30 (3.48-3.30)	47.95-3.31 (3.49-3.31)
Rmerge *, **	0.062 (0.505)	0.046 (0.473)	0.053 (0.612)
I/sigma *, **	14.6 (2.7)	16.6 (2.5)	19.1 (2.7)
multiplicity *	4.0 (3.7)	3.7 (3.8)	5.5 (5.3)
completeness (%) *	99.9 (100.0)	100.0 (100.0)	99.7 (98.1)
solvent content (%)	54.1	56.3	55.8
Phasing		SeMet-SIRAS	
sites		9	
phasing power cen./accen.		2.174 / 1.777	
FOM cen./accen.		0.435 / 0.409	
Refinement			
resolution range *	20-2.95 (3.025-2.95)		
reflections (test set)	7051 (350)		
Rwork *	0.226 (0.288)		
Rfree *	0.296 (0.395)		
number of atoms	1919		
B-factor (protein) (Å^2)	118		
r.m.s.d. bonds (Å)	0.014		
r.m.s.d. angles (°)	1.449		
Ramachandran plot ***			
% most favoured	87.2		
% additionally allowed	10.0		

*Values in parenthesis for outer shell. ** As defined in Scala. *** As defined in PROCHECK.

4.2.3 The structure of Rubisco activase

NtRca is composed of a 67 residue N-domain, followed by a AAA+ module and a 23-residue C-terminal extension (Figure 4.7 A). The N-domain as well as the 23-residue C-terminal extension are not part of the crystallised construct (4.2.1).
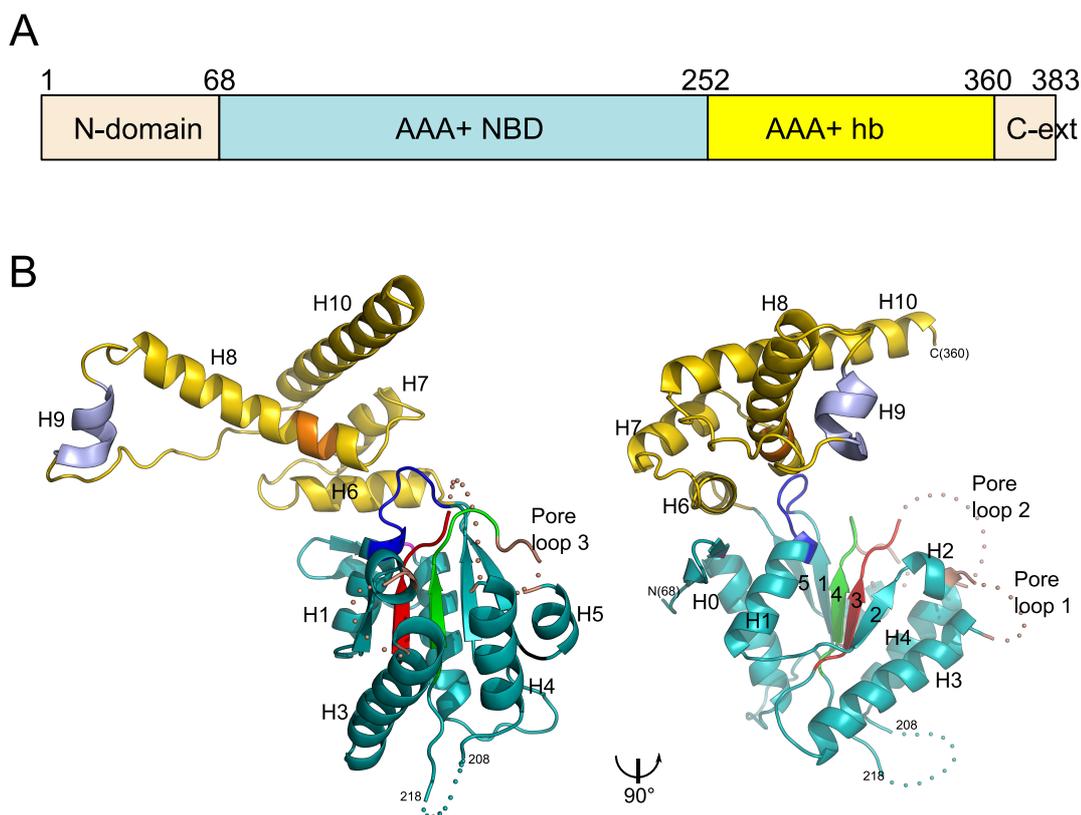
Figure 4.7: **Structure of Rubisco Activase from *Nicotiana tabacum*.**

A: Domain architecture of Rca. (N-domain: N-terminal domain. NBD: nucleotide binding α/β-subdomain. hb: α-helical subdomain. C-ext: C-terminal extension.)

B: Ribbon representation of the crystal structure of Rca from tobacco. Disordered loops are indicated by dotted lines. Two views related by 90 ° are shown. The α/β and the α-helical subdomains are indicated in teal and gold, respectively. The canonical AAA+ structural motifs are indicated as follows: Walker A (dark blue), Walker B (red), sensor I (green) and sensor II (orange). The disordered pore loops are indicated as light brown dots. The specificity helix (H9) is shown in violet. Secondary structure elements, pore loops and chain termini are indicated.

Overall, the solved structure exhibits a AAA+ fold topology, consisting of a nucleotide-binding α/β-subdomain (residues 68-252) and an α-helical subdomain (residues 253-360).

Helix H0 of the α/β-subdomain is preceded by a short N-linker, a β-hairpin, which stacks on top

of helix H1 and provides the connection with the N-domain. The α/β-subdomain is folded in a non-classical Rossmann-fold typical for the AAA$^+$ proteins and comprises the Walker A, Walker B and sensor 1 loops required for ATP binding and hydrolysis [53].

The α-helical subdomain of Rca is at its core a 4-helix bundle fold, which , however, differs further from the classical AAA$^+$ fold topology. While Helices H6, H7, H8 and H10 form the core of the typical 4-helix bundle, Helix H8 is extended by 10 residues and is followed by a short helical insertion, H9 (residues 315-319). H9 is connected to helix H10 via a 16 residue linker and has no contact with the main body of the α-helical subdomain (Figure 4.7 B). Interestingly, mutation of the residues aspartate 316 (to lysine) and leucine 319 (to valine) in helix H9 reversed the specificity of Rca from *Solanaceae* Rubisco to the Rubisco of non-solanaceous plants [184]. Thus this helix must play an important role in substrate recognition.

Four loop segments (residues 142-144, 177-190, 208-218 and 235-236) were not resolved in the crystal structures, most likely due to their flexibility.

4.2.4 Effects of N- and C-terminal truncations on the activity of Rca

In order to test the activity of the NtRca constructs used in crystallisation, ATPase activity as well as Rubisco activation assays were performed. Rubisco activation assays were performed with the decarbamylated, RuBP inhibited, enzymatically inactive form of Rubisco.The speed of reactivation in the presence of NtRca and ATP was measured by monitoring the incorporation of $^{14}CO_2$ into organic sugars by Rubisco (3.3.7) [225].

The effect of N- and C-terminal truncations on the ATPase and the Rubisco-activation activitiy has been investigated previously for Rca of spinach and tobacco [172] [173]. In agreement with these reports, the N-terminal truncation of 67 residues (ΔN68) resulted in unchanged ATPase activity, but loss of Rubisco activation (Figure 4.8). The mutation of tryptophan 16 to alanine (W16A) resulted in the same phenotype (Figure 4.8), indicating an essential role of this residue and the N-terminal domain in Rubisco activation. It is possible that the Rca N-terminal domain acts as the initial interaction site with Rubisco, similar to the N-terminal substrate adaptor domains in other AAA+ proteins (e.g. Spastin [57], proteasomal ATPases (ARC,PAN) [226], Vps4 [227]).

Surprisingly, the C-terminal truncation of 23 residues (ΔC360, ΔN68 ΔC360) resulted in a dramatically decreased ATPase rate and loss of Rubisco activation (Figure 4.8). In contrast truncation of the
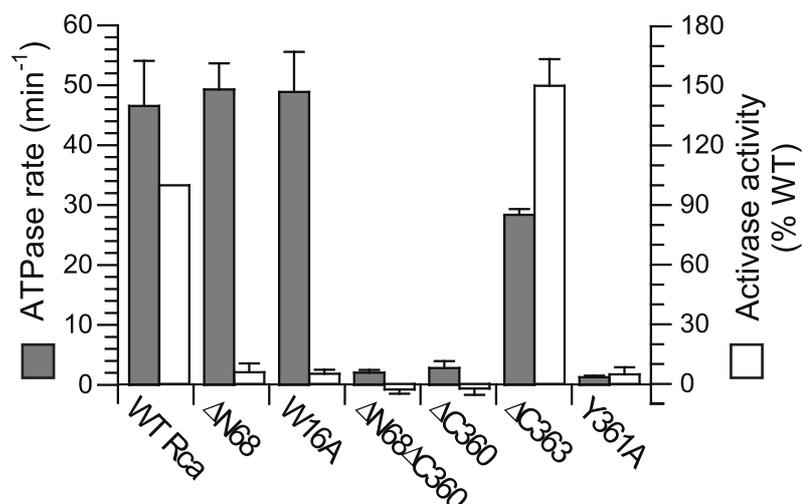
Figure 4.8: **Effects of N- and C-terminal truncations and mutations on the ATPase and activation activities of NtRca.**

Assays were performed at 20°C with 2.35 μM Rca (protomer) (ATPase) and 4.7 μM Rca and 11.9 μM RuBP-inhibited Rubisco active sites (activase activity). Error bars signify s.d. of mean values of at least three experiments.

C-terminal 20 residues (ΔC363) decreased the ATPase activity by ~40 % and stimulated Rubisco activation 1.5 fold, consistent with a corresponding C-terminal truncation in spinach [173]. When the conserved tyrosine residue 361 (see was mutated to alanine (Y361A), the ATPase activity was dramatically decreased again and Rubisco activation activity was lost (Figure 4.8). These results suggest a role for the C-terminal extension and specifically for tyrosine 361 in ATP binding or hydrolysis, similar to findings in the AAA$^+$ protein Spastin [57]. A direct or indirect involvement of the C-terminal extension in ATP binding or hydrolysis may also provide a mechanism for redox-regulation of the ATPase activity of long Rca isoforms in other species [228].

4.2.5 Structural analysis of the N-terminal domain

Since the N-terminal domain is critical for the activase function of Rca (see Figure 4.8), it was expressed in *E. coli* and purified. As expected, the N-terminal domain alone was neither able to hydrolyse ATP nor to activate Rubisco (data not shown). The secondary structure of this domain in the presence and absence of the NtRca AAA+ domain was characterized by CD-spectroscopy (see Figure 4.9). Table 4.2 summarizes the results of the CD-spectra fitted with the CONFIT algorithm on the SMP56 base dataset. While full length NtRca and ΔN68 NtRca showed features consistent with a predominantly α-helical secondary structure (44 % and 48 % respectively), the spectrum of the N-terminal domain alone

suggested an unfolded structure with a low content of β-sheets and almost no α-helical content. This is consistent with the observation that the CD-spectrum of ΔN68 NtRca is similar to the spectrum of full length NtRca.
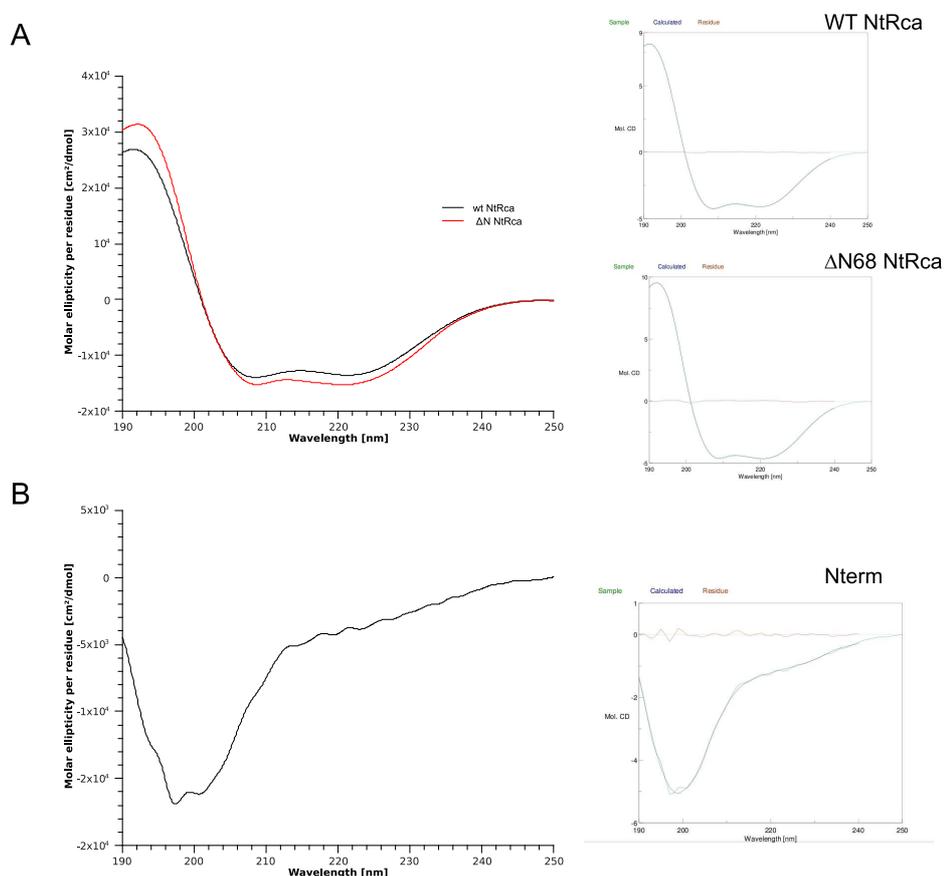
Figure 4.9: **CD spectroscopic analysis of the secondary structure of Rca domains.**
While full length NtRca and ΔN68 NtRca show spectra representative for a mostly ordered α-helical structure, the N-terminal domain alone appears to be largely unstructured.
A: Spectra (right) and CONFIT fit (left) of NtRca and ΔN68 NtRca. B: Spectrum (right) and CONFIT fit (left) of the N-terminal domain.

The intrinsically unstructured nature of the N-domain is also supported by the observation that this domain was cleaved in the initial crystallisation trials. Taken together with its critical role in Rubisco activation, it is tempting to speculate that the N-terminal domain is largely unstructured in the absence of Rubisco and upon binding of Rubisco becomes structured, thereby acting as an efficient substrate adaptor domain. This hypothesis is in accordance with current models for the function of intrinsically unstructured proteins or protein domains [229].

Table 4.2: Results of the CONFIT fit on the CD-spectra. reg.: regular dis.:distorted

sample	reg. α-helix	dist. α-helix	reg. β-sheet	dis. β-sheet	turn	disordered
NtRca	0.263	0.181	0.047	0.057	0.180	0.271
ΔN68 NtRca	0.293	0.186	0.049	0.051	0.168	0.253
Nterm	0.022	0.071	0.142	0.095	0.237	0.433

4.2.6 Arrangement of NtRca in the crystal lattice

Activase subunits in the crystal were arranged helically with six subunits per turn (Figure 4.10). In the crystal each Rubisco activase subunit is in contact to other subunits by three different protein-protein interfaces. These interfaces were analysed using the PISA webserver (*http://www.ebi.ac.uk/msd-srv/prot_int/pisart.html*) at the European Bioinformatics Institute (EMBL-EBI). Table 4.3 provides an overview of the results. Of the three occurring interfaces only the one between the subunits related by the hexameric symmetry operator (illustrated in Figure 4.10) displayed characteristics, that may point to a biological relevance for oligomerisation in solution. This interface involves a large number (>30) of different residues on each subunit and buries approximately 1140 Å^2 of surface area. In comparison, the other two interfaces involve only a small number of residues and bury less surface area, indicative of a contact that probably only occurs in the crystalline state.
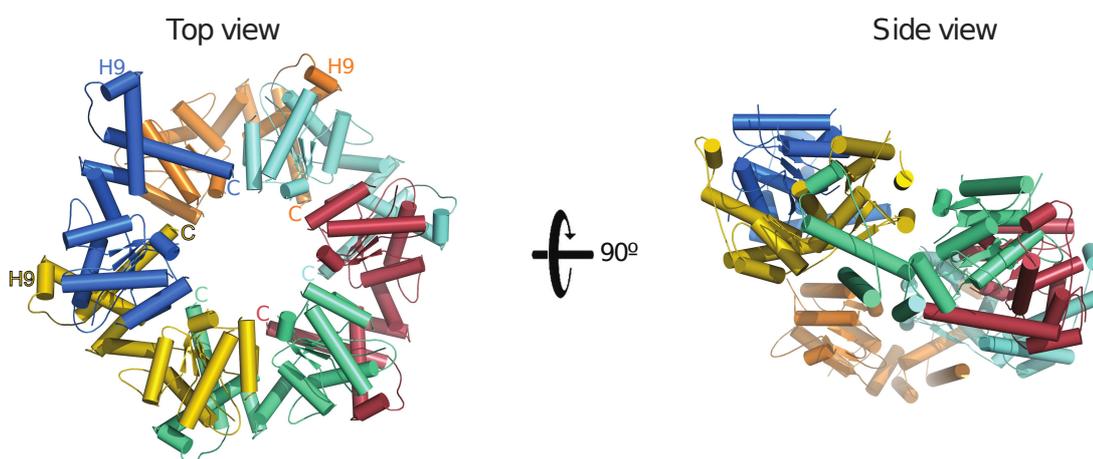
Figure 4.10: **Arrangement of Rubisco activase subunits in the crystal.**
Top view (left) and side view (right) illustrating the hexagonal symmetry and the screw axis are shown. Each Rca subunit is colored in a different color and helices are shown as cylinders. Positions of the specificity helix H9 as well as the C-termini are indicated in the top view.

Table 4.3: PISA analysis of the interface. Sym. Op.: Symmetry operator. # res A/A*: number of residues involved in the interface. Area: buried surface area. HB: number of hydrogen bonds. SB: number of putative salt bridges.

Interface	Sym. Op.	# res. A	# res. A*	Area (Å^2)	HB	SB
1	x-y,x,z-1/6	34	33	1145.6	8	3
2	-y+1,x-y,z+2/3	4	9	206.2	4	4
3	-x+1,-y+1,z-1/2	2	3	21.9	0	0

The interface between the subunits related by the hexameric symmetry operator was predominantly hydrophobic, with residues of helices H8 and H9 of the α-helical subdomain of one subunit (n) packing against a hydrophobic patch of the α/β-subdomain of the adjacent subunit (n+1) (Figure 4.11 A and C). Furthermore, residues participating in this interface were highly conserved, as illustrated in Figure 4.11 B. Conserved charged contacts are lining the periphery of this hydrophobic interface, such as a salt bridge between lysine 92 of the α/β-subdomain and aspartate 299 of the α-helical subdomain and a hydrogen bond of the carbonyl-oxygen of asparagine 99 and arginine 294 (indicated in Figure 4.11 A).
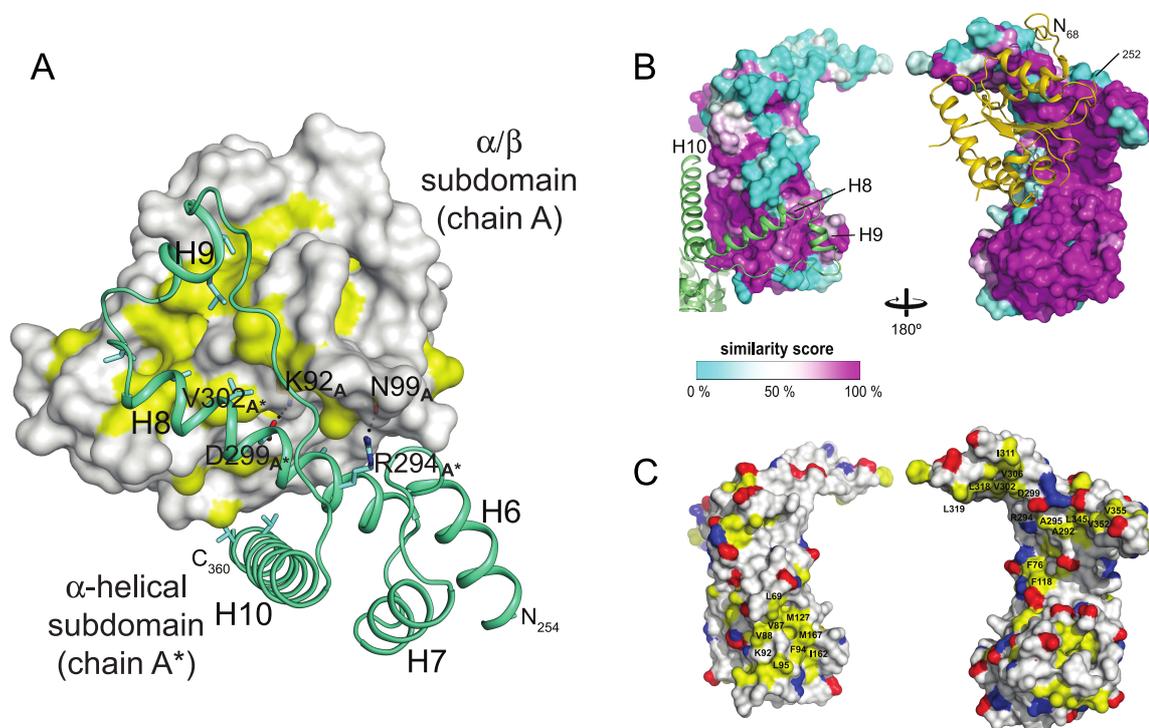
Figure 4.11: **Interface of NtRca subunits in the crystal.**

The interface between the subunits related by the hexameric symmetry operator shows features indicative for a biologically relevant oligomerization interface in solution. A large number of conserved residues (>30) in each subunit generate a largely hydrophobic binding interface between the α-helical subdomain of one chain and the α/β subdomain of the adjacent chain.

A: Interface between the α-helical subdomain of chain A* and the α/β subdomain of chain A. Chain A* shown as ribbon representation colored in teal, chain A shown as surface representation. Hydrophobic residues are colored yellow on chain A. Mutated contacts (see 4.2.7 and 4.4) are indicated. B: Two views of the respective interface related by 90° are shown. Rca subunit of chain A shown as a surface representation and colored by conservation. α-helical and α/β subdomain of the symmetry mates shown in ribbon representation in teal and gold respectively. C: NtRca subunit shown in a surface representation in the same orientation as in B. Hydrophobic residues colored in yellow, basic residues in blue and acidic residues in red. Residues participating in intersubunit interactions are labelled.

4.2.7 Mutational analysis of the Rca subunit interface

In solution Rubisco activase forms active heterogeneous homooligomers of up to 600 kDa in size [196]. Since the large buried surface area as well as the conservation of the subunit interface in the crystal indicated a possible biological relevance for the oligomerisation of Rca in solution, mutations were designed to disrupt this interface. When valine 302 (indicated in Figure 4.11 A) was mutated to lysine, Rubisco activase was highly aggregation prone and could not be purified in a soluble active state (data not shown), most likely due to aberrant hydrophobic contacts. Mutations that were designed to disrupt the charged peripheral contacts, however, were not aggregation-prone and could be purified in a soluble state.

Analytical size exclusion chromatography (Superdex200 PC3.2 on Ettan LC System) of the mutant proteins K92D, D299A and D299K (residues indicated in Figure 4.11 A), designed to disrupt the conserved salt bridge between lysine 92 and aspartate 299, revealed an apparent molecular weight of ~70 kDa (Figure 4.12), consistent with a Rca monomer (expected molecular weight 42 kDa). Interestingly, the double mutant K92D D299K, designed to restore the interface, also eluted at an apparent molecular weight of ~70 kDa, indicating that this salt bridge could not be maintained when the charges are switched. This might be due to the formation of favourable intramolecular salt bridges in the mutant protein (K299 to D300, D92 to K85 or K96). In contrast, wildtype Rca eluted at an apparent molecular weight of ~230 kDa when applied at a concentration of 4.7 μM protomer. The addition of 1 mM ATP did not substantially change the apparent molecular weight of these mutants or of the wildtype Rca (Figure 4.12). As expected, monomeric mutant proteins were neither able to hydrolyse ATP nor to activate Rubisco (Figure 4.13).

Disruption of the conserved hydrogen bond between arginine 294 and the carbonyl-oxygen of asparagine 99 by the mutation of arginine to alanine or valine (R294A, R294V) resulted in an apparent molecular weight of ~90 kDa in the absence of ATP (Figure 4.12), as judged from analytical gel filtration. This is consistent with monomeric or dimeric oligomers of the mutant proteins. Interestingly, the apparent molecular weight in the presence of 1 mM ATP was shifted to ~300 kDa and ~360 kDa for R294A and R294V, respectively, indicating a nucleotide-dependent oligomerisation of these mutants (Figure 4.12). In agreement with this oligomerisation, R294V was fully active in ATPase and Rubisco activation assays (Figure 4.13) whereas R294A showed a ~50 % reduced ATPase rate and only ~20 % of the Rubisco activation activity of wildtype, comparable to previous reports for this mutant [179]. The properties of these arginine mutants will be explored in more detail in chapter 4.4.
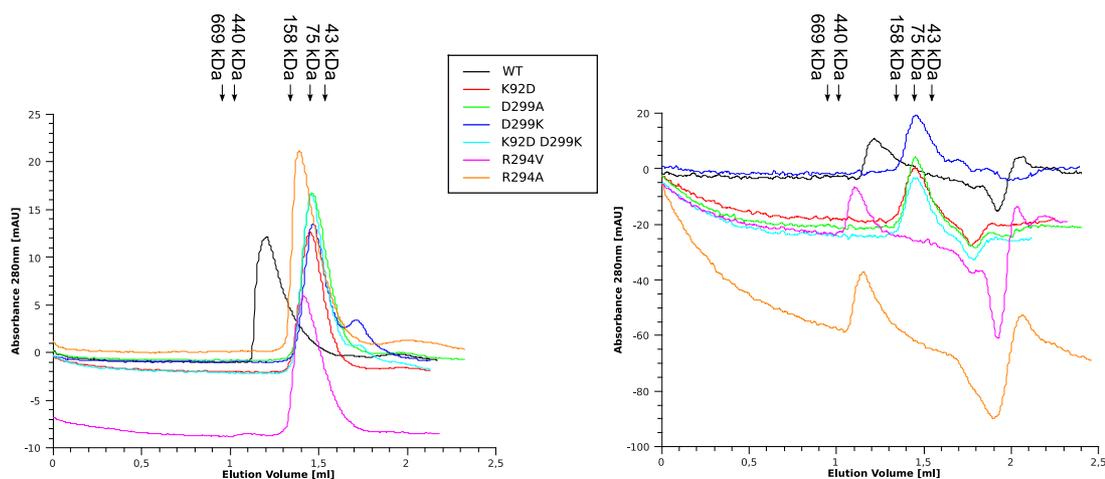
Figure 4.12: **Analytical gel filtration (Superdex200 PC3.2) of wildtype and mutant NtRca.**

Mutations of conserved, charged residues in the Rca-Rca subunit interface resulted in monomeric Rca species in the absence of ATP. NtRca R294A and R294V oligomerised in the presence of 1 mM ATP.

Chromatograms of 4.7 μM of Rca protomers (0.2 mg/ml) in the absence (left) and presence of 1 mM ATP (right) in buffer (20 mM Tris-HCl pH 8, 50 mM NaCl, 2 mM MgCl$_2$) are shown. Elution peaks of globular molecular weight standards are indicated.
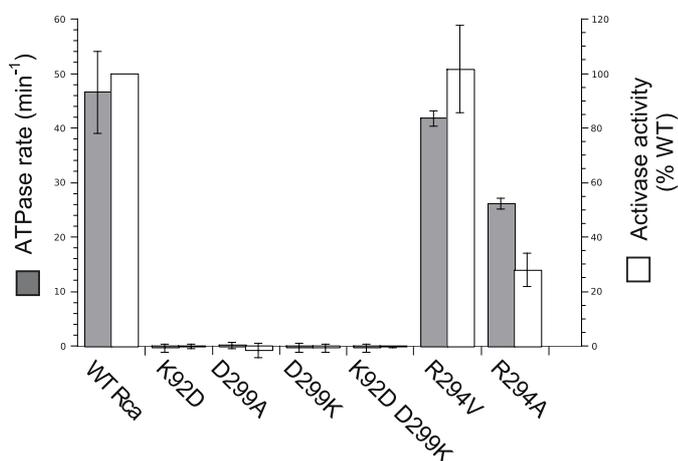
Figure 4.13: **Effects of interface residue mutations on the ATPase and activation activities of NtRca.**

While mutation of aspartate 299 and lysine 92 lead to a complete loss of activity and monomeric Rca, mutation arginine 294 lead to an ATP dependent rescue of activity and oligomerisation. Assays were performed at 20°C with 2.35 μM Rca (protomer) (ATPase) and 4.7 μM Rca and 11.9 μM RuBP-inhibited Rubisco active sites (activase activity). Error bars signify s.d. of mean values of at least three experiments.

4.3 Rubisco activase and the interaction with Rubisco

To further characterise the mechanism of Rubisco activation by Rca, characterization of the interaction of Rca with Rubisco was attempted. Pulldown assays, gel filtration and crosslinking experiments, however, did not provide evidence for a direct interaction of tobacco Rubisco activase with tobacco Rubisco, even with different nucleotide bound states of Rca (data not shown). This is most likely due to the transient nature of this complex. However, since many AAA+ proteins (e.g. Spastin [57], ClpX / ClpA [230]) interact with their respective substrates by recognition of a specific peptide motif, the interaction of C-terminally FLAG tagged Rca with a peptide array of tobacco Rubisco large subunit was investigated.

4.3.1 A possible interaction of activase with a peptide motif of tobacco Rubisco

The possibility of an interaction of Rca with a peptide motive of the Rubisco large subunit was investigated by a peptide array binding assay. In this assay 13-mer peptides covering the sequence of the large subunit of *N. tabacum* Rubisco were spotted on to a membrane redundantly, so that adjacent peptide spots overlapped in 11 residues. This allows for a redundant screening of recognized peptide motives. FLAG-tagged Rca constructs were incubated with this membrane and their binding was detected by indirect immunochemoluminescence after several washing steps. Since the array is redundant with regard to peptide sequence, only binding to a series of sequential spots was regarded as a potential interaction. Figure 4.14 illustrates the results.

N-terminally FLAG-tagged Rca and ΔN68 NtRca were ATPase active and still able to activate Rubisco (~100 % ATPase activity and ~75 % / ~15 % activation activity of wildtype Rca respectively), whereas the FLAG-tagged N-terminal domain alone was inactive as observed with the untagged domain.

Primary (anti-FLAG M2 (SIGMA)) and secondary antibodies (HRP coupled anti mouse Ig-G (Jackson ImmunoResearch)) used to detect bound FLAG tagged proteins showed strong background binding to an acidic peptide motif of *N. tabacum* Rubisco large subunit (Figure 4.14 Background). This might be due to a an affinity of of the FLAG-antibody itself for this specific sequence motif, since the FLAG epitope is also highly acidic. However, FLAG NtRca and FLAG ΔN68 NtRca but not the FLAG-tagged N-terminal domain alone bound additionally to a continuous peptide motif with the consensus sequence LHIHR (Figure 4.14). This peptide motif (residues 291-295 of RbcL) is highly conserved in Rubisco large subunits of different species and directly involved in coordination of a phosphate group of the Rubisco substrate RuBP and the active site Mg^{2+} ion (Figure 4.15).
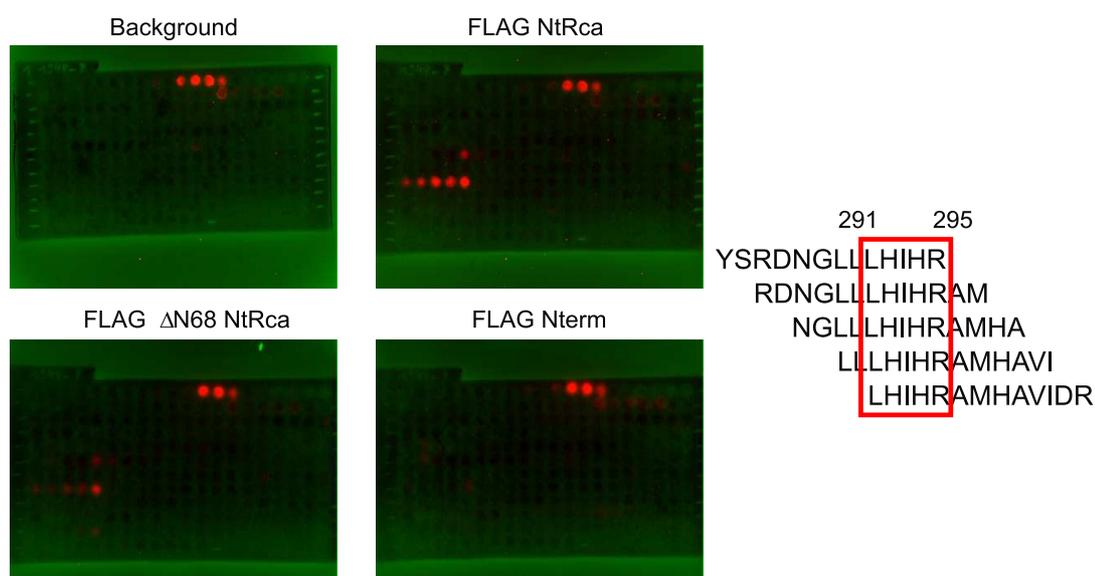
Figure 4.14: **Peptide array screen for the interaction of NtRca with *N. tabacum* Rubisco Large subunit.**
FLAG tagged NtRca and ΔN68 NtRca interacted with a continuous sequence motif of *N. tabacum* Rubisco.
Left: Peptide array screens incubated with different FLAG-tagged proteins and detected by immunochemolumines-
cence.Overlay of grey scale photograph false colored in green and chemoluminescence in red. Respective protein /
background indicated above the images. Right: Sequence of the detected spots for FLAG tagged NtRca and ΔN68
NtRca. Consensus sequence indicated by a red box. Residue numbers correspond to *N. tabacum* Rubisco large
subunit numbering.

Interestingly, this peptide is not accessible on the surface of fully-folded Rubisco (Figure 4.15). There-
fore it is highly unlikely, that binding of Rca to this peptide motif serves as an initial recognition step.
However, an interaction with this peptide motif might be involved in triggering the substrate release from
the inhibited ER complex at a late step in the reaction cycle after ATP dependent remodeling of Rubisco
by activase.
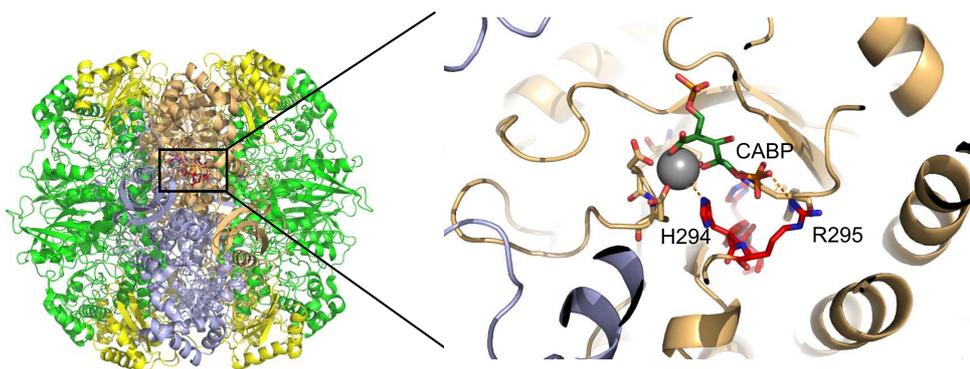
Figure 4.15: **Structure of *N. tabacum* Rubisco and the identified peptide (4RUB [121]).**
The identified sequence motive of Rubisco is not exposed on the surface and directly involved in coordination of
Mg^{2+} and the sugar phosphate substrate in the active site.
Left: Rubisco structure 4RUB in ribbon presentation. A L_2-Dimer shown in brown and blue, other large subunits
in green and small subunits in yellow. Right: View of the active site. Identified peptide motive indicated in red, the
transition state analogue CABP is shown in dark green and Mg^{2+} ion as grey sphere. Residues of the identified
sequence motive involved in coordinating the Mg^{2+} ion and a CABP-phosphate are indicated.

Subsequently, experiments designed to verify the interaction of Rca with this peptide motif in free so-
lution were performed. To this end, a N-terminally acetylated and C-terminally amidated tridecapeptide
corresponding to the sequence Ac-NGLLLHIHRAMHA-NH$_2$ was synthesized by the peptide service
facility of the MPI of Biochemistry and the interaction with activase was analysed by isothermal titration
calorimetry (ITC). Since the peptide was insoluble at pH 8 the interaction was analysed at pH 6, where
NtRca was still ATPase active (~65 % of the ATPase activity at pH 8).

2 mM peptide in buffer (100 mM MES pH 6, 10 mM KCl, 2 mM MgCl$_2$) was titrated to 128.66
μM NtRca in the same buffer in a VP-ITC MicroCalorimeter (Microcal) and the resulting enthalpy was
measured as the current necessary to keep the reference cell and the titration cell at the same temperature.

The measured energy change during the addition of peptide failed to reach saturation. The reaction
is endothermic and compared to usual peptide-protein interactions the measured enthalpy is low (Figure
4.16) , suggesting that the measured energy is a result of the solvation energy of the peptide itself. The
interaction in solution of Rca with the synthesized peptide, therefore, could not be shown using ITC.

Since this might be due to the low pH used in this experiment, the peptide sequence together with a
N-terminal five amino acid linker was cloned in the pHUE vector, expressed in *E. coli* and purified as
N-terminal hexa-histidine tagged Ubiquitin fusion. The resulting protein construct was soluble at pH 8
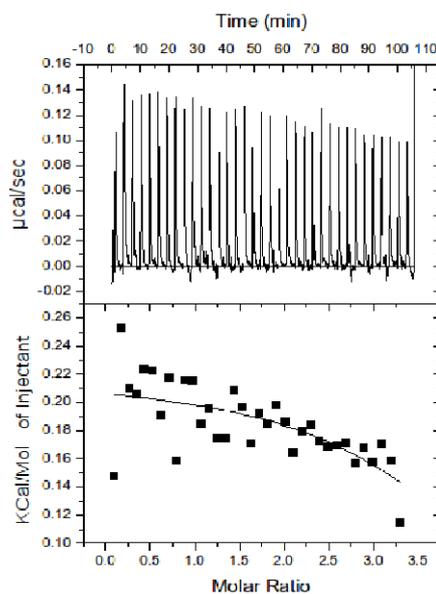
Figure 4.16: **Isothermal titration calorimetry of NtRca and the identified tridecapeptide.**

The identified sequence motive of Rubisco is not binding to Rca in free solution.

Titration of 2 mM peptide to 128.66 μM Rca in 100 mM MES pH 6, 10 mM KCl, 2 mM $MgCl_2$. Upper panel: Energy flow between cells as a result of injection of the peptide plotted against time. Lower panel: Apparent binding enthalpy in Kcal/mol plotted against the molar ratio of peptide / Rca.

and used as bait in NiNTA agarose pulldown experiments. Figure 4.17 summarises the results. NtRca was used as prey in various nucleotide states, but did not interact with the bait His_6Ub-peptide construct bound to NiNTA agarose beyond hexa-histidine tagged Ubiquitin background levels under all tested conditions.

   In summary an interacting peptide motif of the *N. tabacum* Rubisco large subunit was identified by screening for interactions on a peptide array. However, ITC as well as pulldown experiments suggest that Rca does not interact with this motif in free solution. Given these results, it seems more likely that the identified peptide is a false positive identification in the peptide array screen, arising from the effects of the artificial solid-liquid phase interaction between the peptide (coupled to the solid support membrane) and Rca (in solution) in the array.
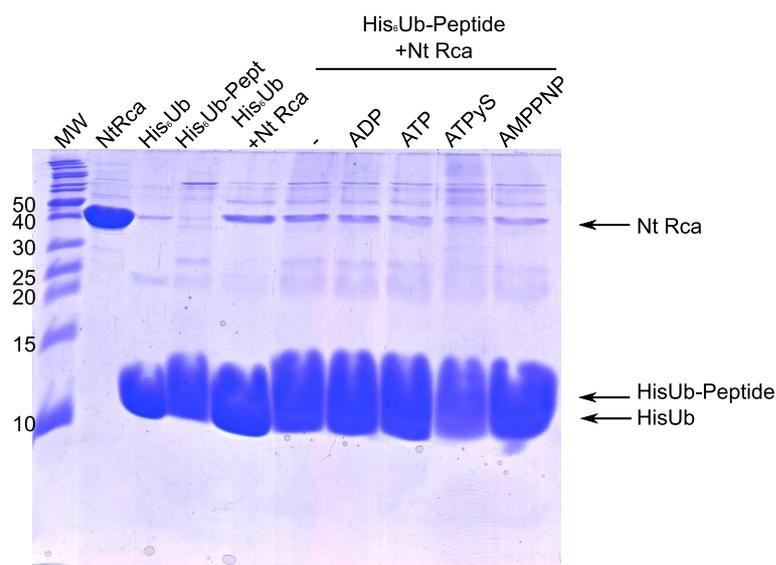
Figure 4.17: **Pulldown of NtRca by His$_6$Ub-Peptide coupled to NiNTA agarose.**

NtRca in various nucleotide states was not retained by the bait His$_6$Ub-peptide construct under all tested conditions. The detected trace amounts of Rca probably reflect the affinity of NtRca to the column material or His-Ubiquitin alone, since they can also be observed in a His-Ubiquitin control pulldown.

Protein input and nucleotide are indicated above the respective lanes. First three sample lanes represent pure protein input controls. Bands of the respective proteins are indicated by arrows.

4.4 Evidence for a hexamer as the active oligomeric form of Rubisco activase

In solution Rubisco activase forms concentration-dependent, active, heterogeneous oligomers of molecular weight ranging from 90 kDa up to 600 kDa (data not shown and [196]). Remarkably, the specific activity of Rca increases with increasing concentration. This has been interpreted as a large molecular weight complex being responsible for Rubisco activation. Together with preliminary EM studies [199], this behavior led to a model of Rca surrounding Rubisco in the activation process [171].

Recently, it has been reported that a previously characterized mutation of arginine 294 to alanine [179] forms hexameric complexes in the presence of ATPγS, as analysed by nano electron spray ionization mass spectrometry [197]. Therefore, the fully active arginine 294 to valine mutant was analysed by disuccinimidyl suberate (DSS) crosslinking and negative stain electron microscopy (in collaboration with Dr. Petra Wendler and Susanne Ciniawsky; LMU Munich Gene Center).

4.4.1 DSS crosslinking of NtRca R294V

To analyse the oligomerisation state of wildtype NtRca and NtRca R294V in the presence and absence of nucleotides, DSS-crosslinking was employed. DSS reacts with primary amines such as lysine side chains or the N-terminal amino group of a protein and crosslinks amines that are within ~12 Å. Protein samples (4.7 μM) were crosslinked with a ~40-fold excess of DSS in buffer (20 mM HEPES pH 8, 150 mM NaCl, 2 mM MgCl$_2$) in the presence or absence of 1 mM ATPγS for 1 min on ice. Samples were analysed by 6 % SDS-PAGE. Figure 4.18 shows the results. Since DSS reacts unspecifically with any primary amine, crosslinked bands on SDS-PAGE are expected to appear less focused than bands of non-crosslinked samples.
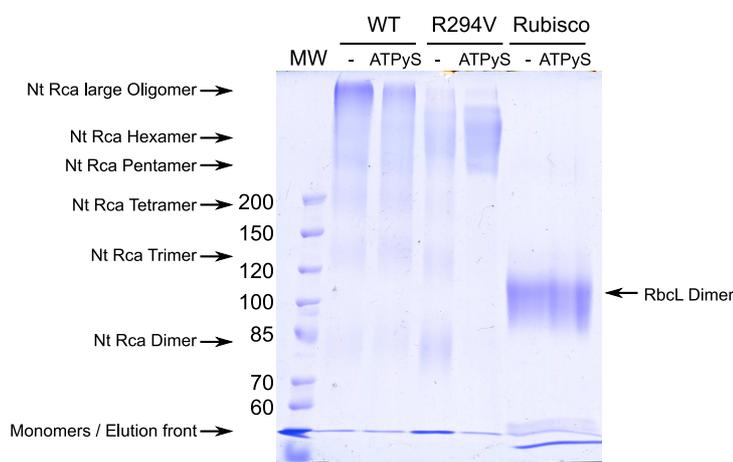
Figure 4.18: **DSS-crosslinks of wildtype NtRca and R294V analysed by 6 % SDS-PAGE.**
Wildtype NtRca crosslinks as heterogeneous, large complexes in presence and absence of nucleotides, while R294V seemed to form predominantly hexamers in the presence of 1 mM ATPγS.
Protein and crosslink conditions are indicated above the respective lanes, 0.2 mg/ml (4.7 μM NtRca, 3 μM *N. tabacum* Rubisco active sites) protein were crosslinked with a ~40-fold excess of DSS (200 μM). MW indicates molecular weight marker. Bands of the respective protein complexes are indicated by arrows.

While ATPγS had no effect on the crosslinking pattern of *N. tabacum* Rubisco, used as control, the band patterns of wildtype NtRca and mutant R294V were changed in the presence of 1 mM ATPγS. In the absence of nucleotides, wildtype Rca formed heterogeneous crosslinked complexes ranging from dimers to a predominant, large complex. In the presence of the nucleotide analogue, this large complex was less populated and bands corresponding to a smaller oligomer became more pronounced. Interestingly, the R294V mutant also formed heterogeneous crosslinked complexes in the absence of nucleotides, however,

bands probably corresponding to dimers, trimers and smaller oligomers were more pronounced than for wildtype Rca. In the presence of ATPγS, crosslinked NtRca R294V predominantly formed a smaller, defined oligomer, that judged by the overall crosslinking band patterns may correspond to a hexamer (Figure4.18).

These observations lead to the hypothesis, that mutation of arginine 294 disturbs the ability of Rca to form very large oligomeric assemblies in solution, but instead a stable hexamer is formed in the presence of nucleotide. To further confirm this, mutant R294V complexes were analysed by negative stain EM (in collaboration with Dr. Petra Wendler and Susanne Ciniawsky).

4.4.2 Negative stain electron microscopy

Negative stain EM grids of NtRca R294V were prepared, imaged and analysed by Dr. Petra Wendler and Susanne Ciniawsky (LMU Munich Gene Center). NtRca R294V forms ring-shaped particles in the presence of ATP and ATPγS but is amorphous in the presence of ADP or the absence of nucleotide (Figure 4.19 A). Particle classification by multivariate statistics and symmetry analysis confirmed a hexameric symmetry of these particles similar to most other AAA+ proteins [53] [52](Figure 4.19 B).

The final three dimensional reconstructions of Rca R294V in the presence of ATPγS were performed with imposed six-fold symmetry and resulted in a nominal resolution of ~24 Å as judged by a Fourier shell correlation cutoff criterion of 0.5 (Figure 4.19 C and D). The final density map has a toroidal shape with an approximate diameter of ~135 Å and a height of ~56 Å. The ring features a central pore with an inner diameter of ~36 Å (Figure 4.19 E).

These results suggest a well ordered hexameric arrangement with a central pore for the fully active mutant R294V. It was concluded, therefore, that wildtype NtRca is also likely to form hexamers, however, the propensity of forming heterogeneous oligomeric states impaired characterization of these hexamers by crosslinking or negative stain EM.
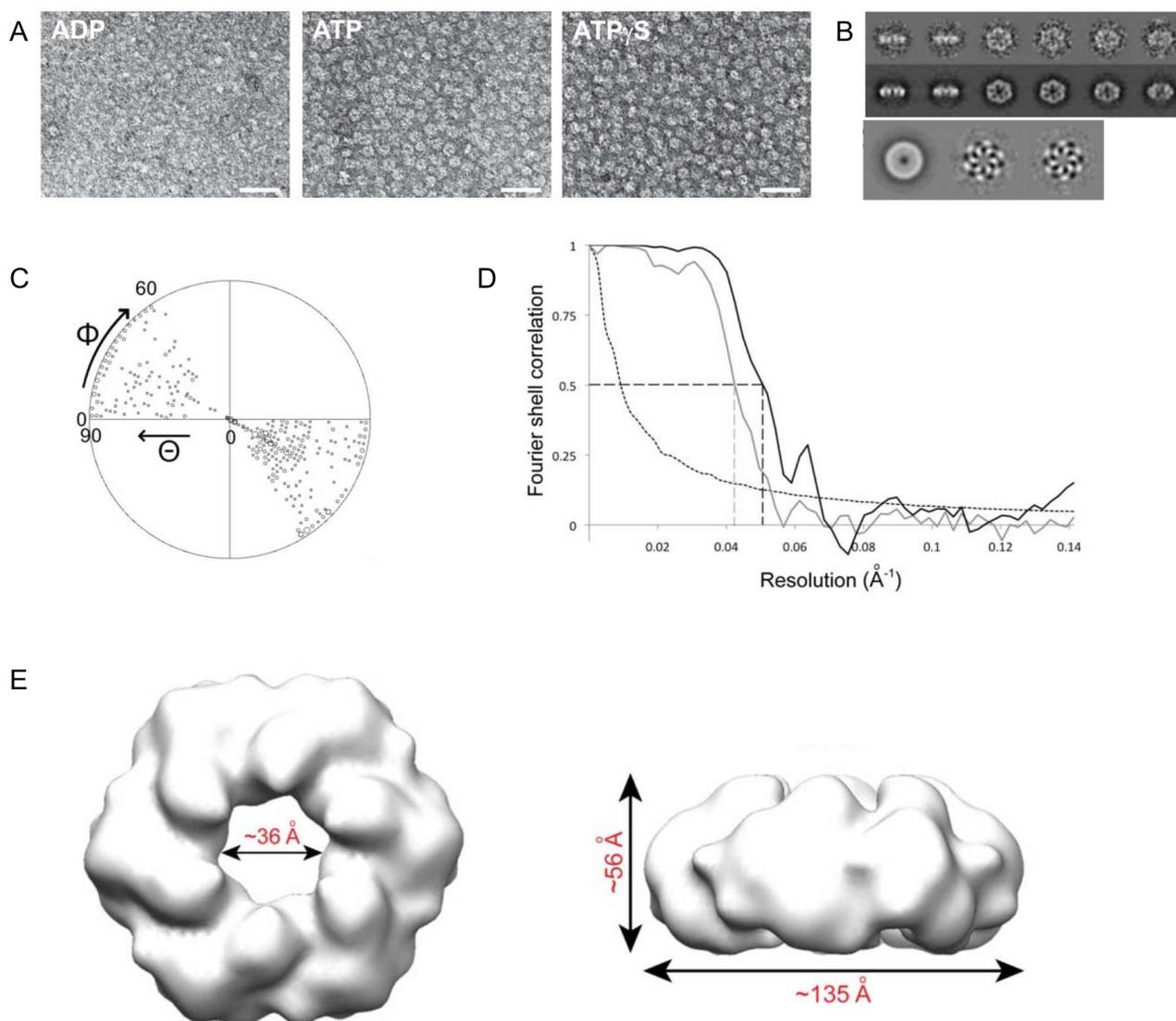
Figure 4.19: **Negative stain electron microscopy of NtRca R294V.**

NtRca R294V forms hexameric rings in the presence of ATP and ATPγS.

A: Negative stain EM micrographs of NtRca R294V (22-44 μg/ml) in the presence of 1 mM Mg-ADP, ATP or ATPγS. Scale bars of 50 nm are shown. B: Upper panel: Initial class averages of NtRca R294V in the presence of ATPγS (upper row, each class contains 5-10 particles) and reprojections of the final density map corresponding to the class averages (lower row). Lower panel: First two Eigen images derived from multivariate statistical analysis of randomly rotated top views. C: Euler angle distribution of the reconstruction with imposed 6-fold rotational symmetry. D: Fourier shell correlation curves of the final 3D reconstruction with (black line) and without (grey line) loose mask. Resolution measured at the 0.5 cut-off criterion is 20 Å and 24 Å, respectively, as indicated by dashed lines. E: Final three-dimensional reconstruction of NtRca R294V in the presence of ATPγS. Top and side view are shown.

4.4.3 Titration of wildtype Rca with monomeric mutants reveals the minimal functional unit

To further verify the active oligomeric state of wildtype Rca, the potential of the previously introduced monomeric mutants to inhibit the activity of wildtype Rca was utilised. The unidirectional mutants K92D, D299K and D299A were expected to interact with wildtype Rca via their remaining intact binding interface but not with the interface containing the mutation, whereas the bidirectional, double mutation K92D D299K was predicted to interact no longer with wildtype Rca (illustrated in Figure 4.20 C). Consistent with this hypothesis, the unidirectional mutants inhibited the ATPase as well as the activase activity of wildtype Rca when mixed stochiometrically. The bidirectional double mutant K92D D299K had no substantial effect on the activity of wildtype Rca when mixed in the same ratio (see Figure 4.20 A).

When equimolar concentrations (4.7 μM of each Rca) of FLAG tagged wildtype NtRca and the non-tagged monomeric mutants were crosslinked by DSS (1 min on ice, ~20 fold excess of DSS), the unidirectional mutants K92D, D299A and D299K but not the bidirectional mutant K92D D299K lowered the apparent molecular weight of crosslinked wildtype Rca oligomers, as judged by 6 % SDS-PAGE followed by immunoblotting against the FLAG tag.

FLAG tagged, crosslinked wildtype NtRca formed heterogeneous, large oligomers on its own or in the presence of the double mutant K92D D299K. The band corresponding to these large oligomeric states became even more intense in the presence of additional untagged wildtype Rca, since this effectively doubles the amount of oligomerisation competent subunits. In the presence of the unidirectional mutants, however, bands corresponding to dimeric or trimeric complexes became more prominent while bands corresponding to large oligomers became less prominent (Figure 4.20 B). This indicates that the unidirectional mutants probably inhibit the ATPase and activation activity of wildtype Rca by interfering with the proper oligomerisation of Rca (illustrated in Figure 4.20 C).
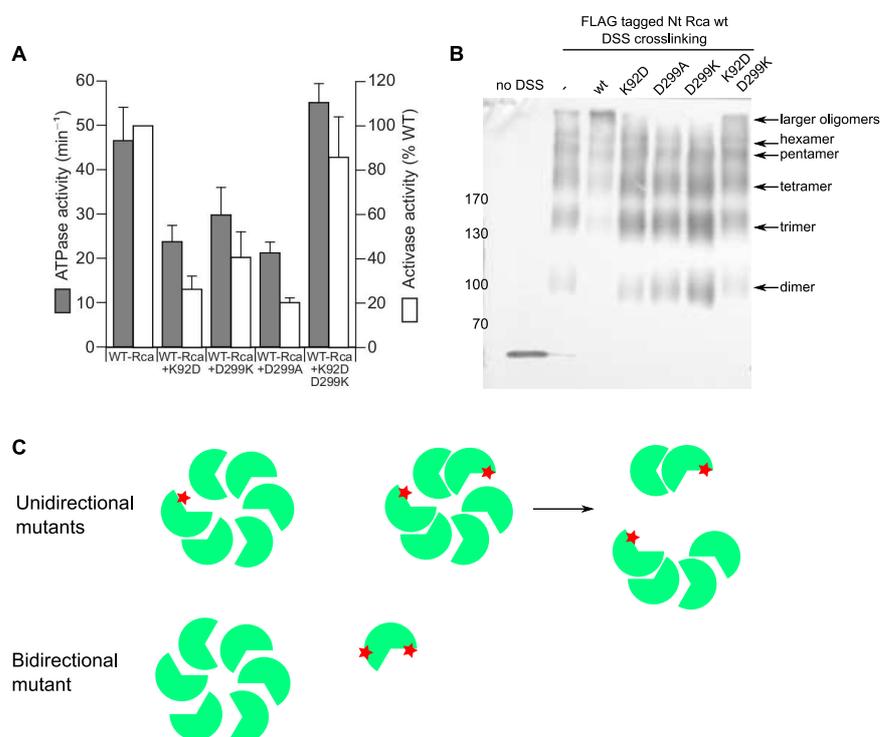
Figure 4.20: **Unidirectional, monomeric mutants inhibit wildtype NtRca by interfering with oligomerisation.**
While the unidirectional mutants K92D, D299A and D299K inhibit wildtype Rca ATPase and activation of Rubisco, the bidirectional mutant K92D D299K did not. Unidirectional mutants inhibit the oligomerisation of wildtype subunits as judged by DSS-crosslinking followed by 6 % SDS-PAGE and immunoblotting.

A: ATPase and activation activities of wildtype NtRca mixed with mutant proteins. Assays were done at 20°C with 2.35 μM of each Rca (protomer) (ATPase) and 4.7 μM Rca and 11.9 μM RuBP-inhibited Rubisco active sites (activase activity). Error bars signify s.d. of mean values of at least three experiments. B: DSS-crosslinking of wildtype NtRca mixed with mutant Rca. 4.7μM of each Rca subunit were crosslinked in the presence of a ~20-fold excess of DSS (100 μM) for 1 min on ice in 20 mM HEPES pH 8, 150 mM NaCl, 2 mM MgCl$_2$. Bands of the respective protein complexes are indicated by arrows. Positions of bands of the protein marker are indicated by their respective molecular weight C: Graphical illustration of the mechanism of inhibition by the unidirectional, monomeric mutants. Rca protomers shown in green, mutations in the binding interface are indicated by red stars. Unidirectional mutants can still be incorporated in wildtype Rca complexes and therefore interfere with oligomer formation (upper row). The bidirectional, double mutant has both binding interfaces mutated and therefore no longer incorporates into wildtype Rca complexes (lower row).

The addition of increasing concentrations of unidirectional mutants to a constant concentration of wildtype Rca lead to a concentration dependent decrease in the activity of Rca (Figure 4.21). This titration experiment can be used to estimate the minimal number of subunits needed for the activity of Rca using a mathematical model [231] [232]. The distribution of mutant and wildtype subunits in a given n-mer can be estimated by a binomial distribution:

$$P(k) = \binom{n}{k} p^k (1-p)^{(n-k)} \tag{4.1}$$

where P(k) is the probability that exactly k mutant subunits were incorporated into an oligomer with the size n and p is the probability to incorporate a mutant subunit. Since it can be assumed that wildtype and mutant subunits have the same probability to be incorporated into the complex, p is given by their ratios in solution

$$p = \frac{c(mutant)}{c(WT) + c(mutant)} \tag{4.2}$$

It is assumed that if an n-mer of any size incorporates more than one subunit with a mutated binding site on one interface, it will decay into smaller pieces (as illustrated in Figure 4.20) and therefore is assumed to be inactive. In other words, the active species can only be an n-mer that has either none or just one mutant subunit incorporated. Thus, the distribution of active n-mers in solution is given by the equation:

$$P(active) = P(0) + P(1) = (1-p)^n + np(1-p)^{(n-1)} \tag{4.3}$$

which can be fitted on data of the activity of wildtype Rca titrated with mutant subunits to result in parameters for n, the minimal size of the active complex.

Fitting of this mathematical model to concentration series with increasing concentrations of the unidirectional mutants suggested that a trimeric complex is the minimal ATPase-active assembly. By contrast, the functional wildtype Rca oligomer was found to require at least 3.5, 4.2 or 5.4 subunits for Rubisco activation, as calculated for D299K, K92D and D299A, respectively. This analysis is based on the assumption that the unidirectional mutants have the same likelihood of being incorporated into oligomers as does the wildtype subunit and that if an n-mer of any size incorporates more than one subunit with a mutated binding site on one interface, it will decay into smaller pieces. It would appear that this assumption is better fulfilled by the mutant D299A than by D299K or K92D.

Together with the finding that Rca R294V forms active hexameric rings, these titration experiments support the conclusion that the minimal oligomer of functional wildtype Rca is a hexamer.
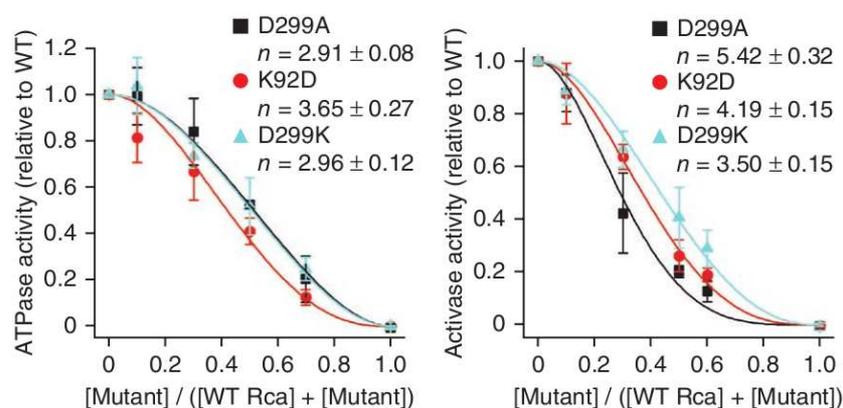
Figure 4.21: **Titration of wildtype Rca with increasing concentration of unidirectional mutant subunits.**

Fitting of the mathematical model (Equation 4.3) to the titration experiment results in parameters for the minimal functional oligomer of wildtype Rca .

Wildtype Rca concentration was constant at 2.35 μM (ATPase) and 4.7 μM (activase) protomer, except for the last data points, which represent the respective mutant alone. Rubisco was present as 11.9 μM RuBP-inhibited active sites. Error bars signify s.d. of mean values of at least three experiments. The minimal number of wildtype subunits per functional oligomer is indicated, as calculated for the different unidirectional assembly mutants.

4.4.4 A hexameric model for Rubisco activase

To generate a hexameric ring model of Rca, the module of the crystal structure consisting of the α/β subdomain of one chain bound to the α-helical subdomain of the adjacent chain (Figure 4.11) was fitted onto the ATPase-active D2 ring of the hexameric AAA$^+$ protein p97 ([55] (PDB 3CF3)). This module forms a building block that is invariant through the conformational changes of AAA$^+$ hexamers [62]. The transition from the helical ring arrangement in the Rca crystal to a flat hexamer involved only a minor inter-domain reorientation. The resulting hexamer is similar in dimensions to the helical subunit arrangement in the crystal lattice, with an outer diameter of ~125 Å, height of ~49 Å and a pore width of ~30 Å. It fits well into the EM density (diameter of ~135 Å, height of ~56 Åand a central pore of ~36 Å) obtained by three-dimensional reconstruction of the Rca R294V hexamer (Figure 4.19 and 4.22). The fit leaves unfilled density on top of the hexameric ring close to the central pore (Figure 4.22). This density could accommodate all or part of the 67-residue N-terminal domain, which is not present in the crystal structure. A similar arrangement of the N-terminal domain is observed in the hexamer model of spastin [57]. In the hexamer Rca model, the specificity helix H9 is located in proximity to the putative N-domain

density, consistent with a cooperative function in the initial recognition of Rubisco (Figure 4.22).
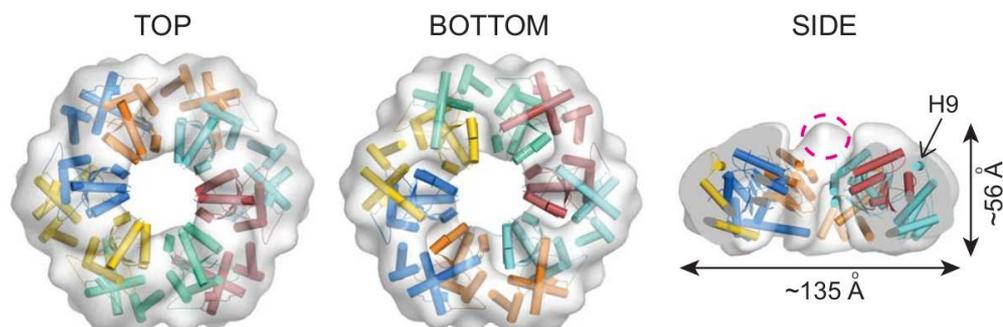
Figure 4.22: **A hexameric model of NtRca.**

A hexameric model based on the crystal structure of NtRca and negative stain EM of NtRca R294V. Additional density is visible in the EM derived density on top of the hexameric arrangement. This density might accommodate part of the N-terminal domain not contained in the crystal structure.

Top, bottom and cutaway side view of the overlay of a hexamer model based on six α A-α/β A* units (as shown in Figure 4.11) onto the EM reconstruction of the Rca hexamer obtained from negative-stain data (Figure 4.19). Helices are shown as cylinders. Each Rca subunit is shown in a different colour. The pink dashed circle indicates the unfilled density on the top of the hexamer. The position of the specificity helix H9 is indicated.

In order to verify the position of the N-terminal domain at the top of the Rca hexamer the mutant R294V lacking the first 67 residues (ΔN68 NtRca R294V) was analysed by negative stain EM in collaboration with Dr. Petra Wendler and her student, Susanne Ciniawsky. Like full length NtRca R294V, ΔN68 NtRca R294V formed hexameric rings in the presence of ATPγS (Figure 4.23 A). Three-dimensional reconstruction gave rise to a ring shaped density with a height of ~52 Å, a diameter of ~130 Å and central pore with a diameter of ~38 Å (Figure 4.23 B). The reconstruction had a resolution of ~25 Å as judged by a Fourier shell correlation analysis. However, unlike the full length protein, ΔN68 NtRca R294V did not orient randomly on the grid but showed a tendency to present fewer top and side but more intermediate views.
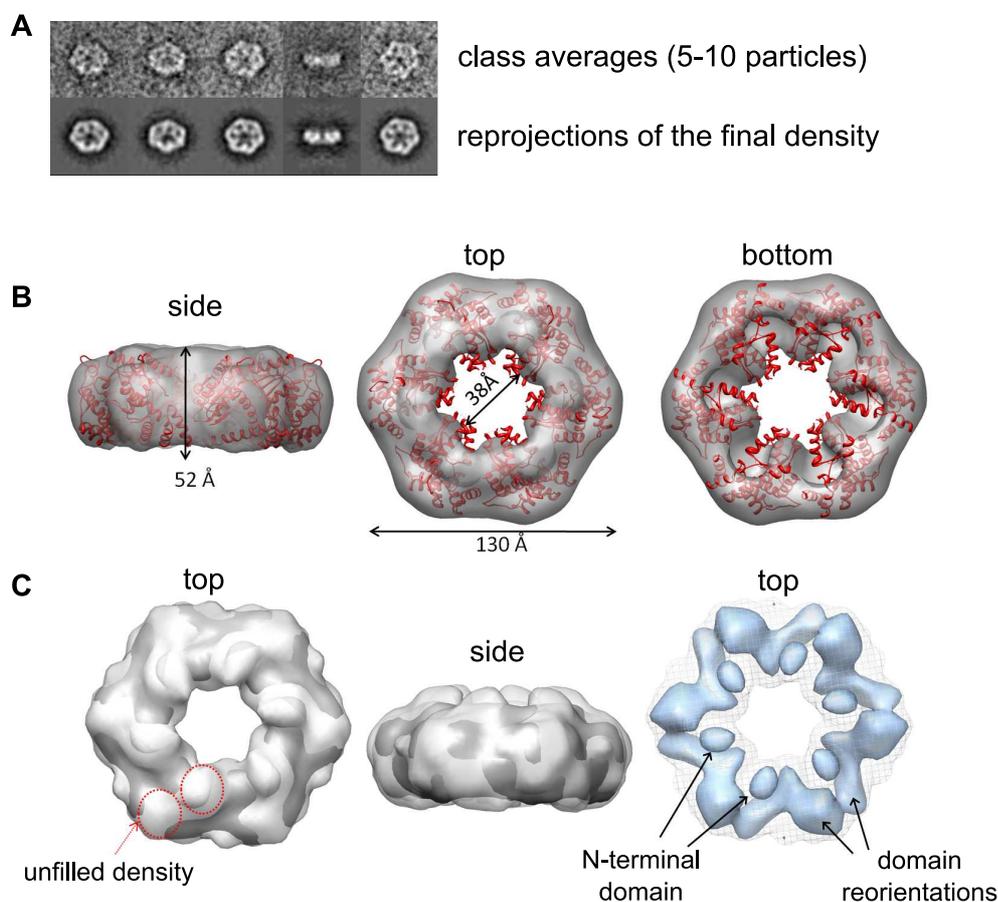
Figure 4.23: **Negative stain electron microscopy of ΔN68 NtRca R294V.**

Comparison of the three dimensional reconstruction of ΔN68 NtRca R294V with the full length protein gave rise to difference density at the top and the side of the hexamer, consistent with the presumed location of the N-terminal domain. Based on the difference density map, AAA+ domain reorientations in the absence of the N-terminal domain are likely.

A: Initial class averages of ΔN68 NtRca R294V (44 μg/ml in the presence of ATPγS (upper row, each class contains 5-10 particles) and reprojections of the final density map corresponding to the class averages (lower row). B: Final three dimensional reconstruction of ΔN68 NtRca R294V in the presence of ATPγS. Top, bottom and side view are shown. The hexameric model shown as ribbon representation in red. C: Comparison of the EM density maps of NtRca R294V (white) and ΔN68 NtRca R294V (grey). Overlay of both maps shown in top and side view. Difference density map shown in blue in top view. Unfilled density and putative location of the N-terminal domain indicated by arrows. Difference density corresponding to AAA$^+$ domain reorientations indicated.

When overlayed with the three-dimensional reconstruction of full length NtRca R294V, the reconstruction of ΔN68 NtRca R294V shows marked differences. The reconstruction of ΔN68 NtRca R294V lacks density at the top and the sides and is 4 Å shorter in height compared to full-length NtRca R294V, consistent with the proposal that the density on top represents the N-terminal domain, which is not present in the crystal structure. Additionally, large difference density peaks were observed for the core AAA+ domain, possibly indicating inter-domain reorientations of the complex in the absence of the N-terminal domain.

These results are all consistent with the proposed location of the N-terminal domain at the top of the hexameric ring, however, possible stain artefacts and the biased view of the ΔN68 NtRca R294V might hinder proper interpretation of the resulting maps. Therefore, the location of the N-terminal domain remains putative.

4.5 Mutational analysis of pore loop residues

Many AAA$^+$ proteins remodel the conformation of their substrates by (partial) threading of their binding partner through the central pore mediated by the so called pore loops [53]. In these AAA+ proteins pore loop 1 usually features a very conserved aromatic-hydrophobic-glycine motive, that is essential for the remodeling of the protein substrate but not for the ATPase activity. In Rubisco activase sequences this conserved motive does not exist in pore loop 1. However, the sequence in this region is still highly conserved within the Rca family (Figure 4.24 A). Pore loop sequences (residues 141-145 pore loop 1, 177-191 pore loop 2 and 232-237 pore loop 3) were not fully resolved in the crystal structure of ΔN ΔC NtRca most likely due to their flexibility, however, their approximate location in the hexameric model is consistent with an exposure to the central pore (Figure 4.24 B).
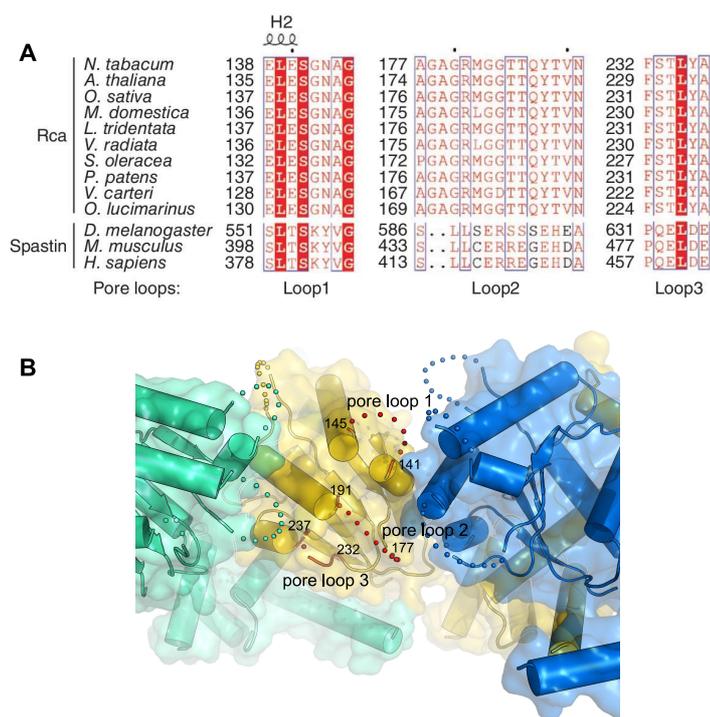
Figure 4.24: **Sequence alignment and location of pore loops in Rubisco activase.**

The highly conserved pore loop sequences of Rca do not match classical pore loop sequences in Spastin, but are exposed towards the central pore channel.

A: Alignment of the pore loop sequences in Rubisco activases and Spastins from different species. Similar residues are shown in red and identical residues in white on red background. B: Cutaway view of the central pore in the hexameric model. Individual Rca subunits are shown in green, gold and blue, respectively, with transparent surface representation. The pore loop residues are indicated by a red color for the subunit in gold. Residue numbers for the residues flanking the the pore loops are indicated. Residues not resolved in the crystal structure are modelled as dotted line.

It was therefore tested whether mutation of the conserved pore loop 1 residues in NtRca would affect ATPase or Rubisco activation activity. Pore loop 1 is preceded by a short helix (H2) as observed in other classical clade AAA+ proteins. Mutations in helix H2 (E138A and E140A) and the first residue of the pore loop (S141A) resulted in a 60-100 % reduction in ATPase rate and loss of Activase function (Figure 4.25), possibly due to the close proximity of the Walker B loop of Rca. Mutation of other residues in pore loop 1 (G142V, N143A and A144V) reduced the ATPase rate by ~25 % while the activase activity was reduced by ~40 % for the mutants G142V and N143A but completely abolished for the mutant A144V (Figure 4.25). This indicates an essential role for pore loop 1 in Rubisco activation. Since the

aromatic residue of the canonical pore loop 1 motive is completely absent in activase sequences, it was tested whether conserved aromatic residues in pore loop 2 and pore loop 3 might compensate for this.
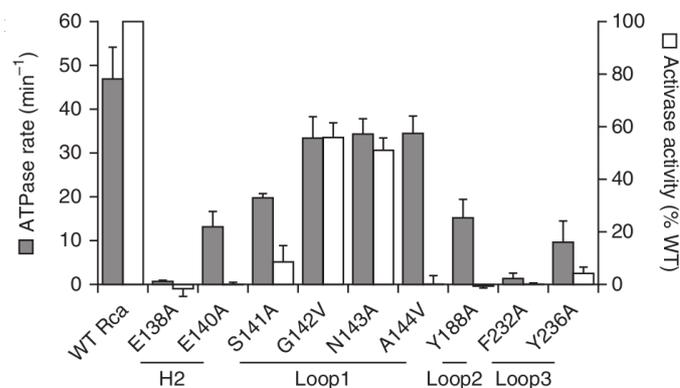
Figure 4.25: **Mutation of conserved pore loop residues.**

Pore loop residue A144 in pore loop 1 as well as aromatic residues in pore loop 2 (Y188) and pore loop 3 (F232,Y236) are essential for activity.

ATPase and Rubisco activase activities of wildtype NtRca , helix H2 mutants and loop mutants, as indicated. Assays were performed at 20°C with 2.35 μM Rca (protomer) (ATPase) and 4.7 μM Rca and 11.9 μM RuBP-inhibited Rubisco active sites (activase activity). Error bars signify s.d. of mean values of at least three experiments.

Mutations of the aromatic residues in pore loop 2 (Y188A) and pore loop 3 (F232A, Y236A) strongly reduced the ATPase activity and abolished activase function (Figure 4.25). Taken together these results argue for an active role of all three pore loops in Rubisco reactivation. Interestingly, an active role for all three pore loops has also been shown for the AAA+ protein Spastin [57].

5 Discussion

Rubisco activase is an essential enzyme for photosynthesis in plants, since it is required for the ATP-dependent, maintenance of Rubisco activity in vivo. Since Rca is heat labile, it has been proposed that Rca is a limiting factor for the photosynthetic potential of plants under moderate heat stress. This will probably become increasingly problematic in the future as the temperature of the Earth's atmosphere increases due to climate change [162] [163].

In this study, the elucidation of the AAA+ core domains of Rubisco activase lead to a model for the oligomeric state of the functional activase complex and together with structure guided mutations provided a framework for the mechanism of Rubisco activase. The presented work may serve as the foundation of more detailed mechanistic studies.

The following sections contrast and compare the result of this study with results in previous studies and present an integrated model for the Rca mechanism.

5.1 Comparison of the available crystal structures of the C-terminal α-helical subdomain of Rubisco activase

While this thesis was in preparation, the 1.8 Å crystal structure of the isolated C-terminal α-helical subdomain of Rubisco activase from Creosote bush (*Larrea tridentata*, a *non-Solanacae* species) was published (PDB 2THG [201]). While the main chain atoms for the central 4-helix bundle of both crystal structures (residues 250-297 and 325-346 in Creosote bush and residues 253-300 and 328-349 in tobacco) can be readily superposed with a rmsd of 0.929 Å, the backbones of the extended part of the C-terminal domain (298-324 in Creosote bush, 301-327 in tobacco) cannot be easily superposed as indicated by a high rmsd of 4.722 Å (structures were superposed with the programs superpose and lsqkab in CCP4i).

Indeed, the crystal structure of *Larrea tridentata* (Creosote bush) features an extended helix H8 with three additional turns and lacks the small helical insertion H9 present in *Nicotiana tabacum*. This extended conformation of helix H8 seems to be stabilised by a salt bridge between aspartate 309 and lysine
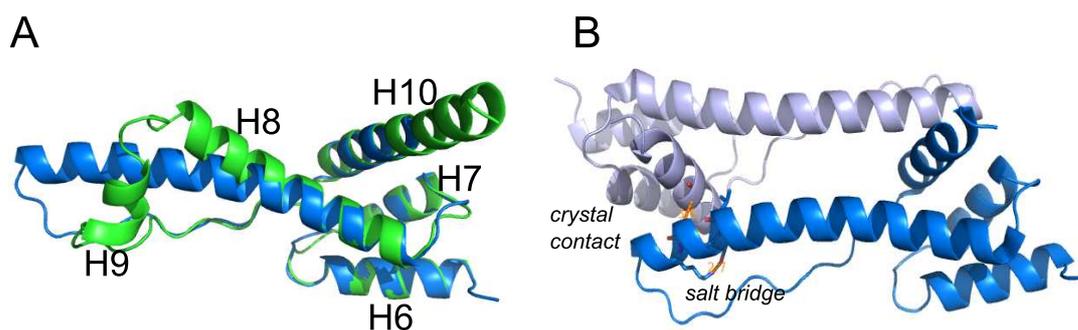
Figure 5.1: **Comparison of the crystal structures for the C-terminal α-helical subdomain of Rca.**
A: Ribbon representation of the overlay between the α-helical subdomain of *Larrea tridentata* Rca (blue) and *Nicotiana tabacum* Rca (green). While helix H8 is even more extended, the small helical insertion H9 is absent in Rca from *Larrea tridentata*. Nomenclature of helices is according to *Nicotiana tabacum* Rca. B: Crystal contact of the C-terminal domains of *Larrea tridentata* Rca . Rca subdomains are shown in ribbon representation in blue and light blue respectively. The interface between the two chains consists of a close contact of two glycine residues shown in stick representation (Gly 312 of chain A to Gly 337 in A*, distance 3.6 Å). The extended conformation of H8 is further stabilised by a salt bridge between Asp 309 and Lys 313 represented as sticks.

313 and a crystal contact to a symmetry related molecule (glycine 312 of chain A to glycine 337 in A* see Figure 5.1 B), while in the crystal structure of *Nicotiana tabacum* the conformation is mainly stabilised by the interface to the α/β nucleotide binding domain of the symmetry mate. Interestingly, the proposed Rubisco binding surface of *Larrea tridentata* Rca (region around tryptophan 302 [201]) is buried in the intersubunit interface in the structure of *Nicotiana tabacum* (see Figure 4.11).

However, since *Larrea tridentata* is a *non-Solanacae* species, the structural differences may also reflect the differences in specificity. Indeed, the stabilizing salt bridge to lysine 313 in *Larrea tridentata* Rca cannot form in the *Solanacaea* Rca as this residue is substituted by an aspartate, which defines the species specificity of the *Solanacea* family (see 2.5.1.3 and Figure 5.2). Whether the differences in the structures are a consequence of crystal packing or reflect differences in function or specificity remains to be established and cannot be decided in the absence of the structure of a Rubisco-Rubisco activase complex or mutational analyses (e.g. mutation of the region around tryptophan 302).

5.2 Oligomeric state

Rubisco activase forms polydisperse oligomers with a range of sizes from ~50 kDa to ~600 kDa in a concentration and nucleotide dependent manner [193] [194] [195] [196]. Intriguingly the specific ATPase-rate as well as the Rubisco activation activity and the oligomeric size increase with increasing concentrations and in the presence of crowding agents (e.g. 5 % PEG 3350) [194]. This has been interpreted to indicate that a large oligomer (>8 up to 16) is the active state of activase [195] [196], but is also consistent with a smaller, perhaps transient complex being the active species. Recent analyses of the oligomerisation state of an ATP hydrolysis deficient activase mutant (R294A) in the presence of Mg^{2+} and ATPγS by mass spectrometry showed that Rca formed a complex with a mass corresponding to a hexamer [197].

The structure of Rca determined in this study allowed the identification of the oligomerisation interface of Rca, which - in agreement with previous crosslinking studies [198] - consists of an interface between the C-terminal α-helical subdomain and the N-terminal nucleotide binding domain. Crosslinking and negative stain EM of the functionally active mutant R294V confirmed a hexameric arrangement in the ATPγS bound state. Since arginine 294 contributes to the intersubunit interface by forming a hydrogen bond to the carbonyl-oxygen of asparagine 99, the hypothesis is that arginine 294 mediates higher oligomerisation of Rca, whereas its mutation stabilises the active hexameric arrangement in the presence of ATPγS. Titration experiments with oligomerisation defective mutants (D299A, D299K, K92D) further support a hexameric arrangement of activase in the active state.

In summary the results presented in this study suggest, that activase forms a transient hexamer similar to other classical clade AAA+ proteins, such as Spastin or Katanin in the active state. This state is most likely dependent on interactions with the nucleotide or even the substrate Rubisco. In the absence of the proper conditions for hexamer formation Rubisco activase can oligomerise into higher order oligomers, which might have a protective or regulatory function.

5.3 Mechanism of Rubisco activation by Rubisco activase

Figure 5.2 shows an alignment of activase sequences and highlights residues for which mutation data are available. Taken together with the structure of the hexameric , active state a preliminary model for the Rubisco activation mechanism by Rubisco activase emerges.

Block 1:

```
N_tabacum       1   EQIDVDPKKQTDSDRWKGLVQDFS..DDQQDITRGKGMVDSLFQAPTGTG.THHAVLQSYEYVSQGLRQYN.
A_thaliana      1   ...AVKEDKQTDGDRWKGLAYDTS..DDQQDITRGKGMVDSVFQAPMGTG.THHAVLSSYEYVSQGLRQYN.
O_sativa        1   .AKELDEGKQTDGDRWKGLAYDIS..DDQQDITRGKGFVDSLFQAPTDG.THAVLSSYEYLSQGLRTYD.
M_domestica     1   .ASVDEDKQTDKDRWKGLAFDTS..DDQQDITRGKGMVDTLFQAPMQSG.THFAIMSSYDYISQGLRQYN.
L_tridentata    1   .AQEISEDQQTDKDRWKGLAYDIS..DDQQDITRGKGMVDTLFQAPMQSG.THYAVMSSYDYISQGLRQYN.
V_radiata       1   AEKEIEESQQTNKDRWKGLAYDIS..DDQQDITRGKGMVDPLFQAPMDAG.THYAVMSSYEYLSTGLRQ...
S_oleracea      1   .....AENEEKNTDKWAHLAKDFS..DDQQDITRGKGMVDSLFQAPADAG.THVPIQSSFEYESQGLRKYD.
P_patens        1   .VCEAEGEKKKKVDRWAGLGNDIS..DDQQDIRGKGMVDALFQGAVGIG.TQVVTMSSWDYVSTGQRTYN.
V_carteri       1   ..........QVGRWKNIDADVDTSRGRQMVDDLFQGGFGAGGTHNAVLSTEYLS...QARAS
O_lucimarinus   1   ..........EMSRKGMDMDIS..DDQQDIARGKNMVDSKFQGATGIGGTHNAVMSSQDYLSAGMKTYSA
consensus>50        .........q.d.drW.gl..D.s..DDQqDItRGkgmVD.lFQap.g.G.Th.avmsSy#Y.S.glr.yn.
```

Block 2:

(markers: H0, TT, β1, TT, H1, β2, H2)

```
N_tabacum      69   LDNKLDG.FYIAPAFMDKLVVBITKNFLKLPNIKVPLILGIWGGKGQGKSFQCELVFRKMGINPIMSAGEL
A_thaliana     66   LDNMMDG.FYIAPAFMDKLVVHITKNFLTLPNIKVPLILGIWGGKGQGKSFQCELVMAKMGINPIMSAGEL
O_sativa       68   FDNTMGG.FYIAPAFMDKLVVHITKNFTLPNMKVPLILGIWGGKGQGKSFQCELVFAKMGINPIMSAGEL
M_domestica    67   FDNNMDG.FYIAPAFMDKLVVHITKNFMTLPNMKVPLILGIWGGKGQGKSFQCELVFAKMRISPIMSAGEL
L_tridentata   68   LDNNMDG.FYIAPAFMDKLVVHITKNFLSLPNMKIPLILGIWGGKGQGKSFQCELVFAKMGINPIMSAGEL
V_radiata      67   LDNIKDG.FYIAPAFLDKLVVHITKNFMTLPNIKVPLILGIWGGKGQGKSFQCELVFAKMGINPIMSAGEL
S_oleracea     64   IDNMLGD.FYIAPAFMDKLVVHITKNFLNLPNIKVPLILGIWGGKGQGKSFQCELVFAKLGINPIMSAGEL
P_patens       68   FDNMKDG.LYIAPAFLDKLIHIAKNFMNLPGIKVPLILGIWGGKGQGKSFQCELVMAKMGINPIVMSAGEL
V_carteri      59   FNNIEDG.FYISPAFLDKMTIHIAKNFMDLPKIKVPLILGIWGGKGQGKTFQCALAYKKLGISPIVMSAGEL
O_lucimarinus  60   HDNITTENRYISPSYMDKVIVHVAKNFMDLPKIKVPLILGIWGGKGQGKSFQCALLIFKKLGISPILMSAGEL
consensus>50        .#N..dg.fYIaPa%$DKlv!H!.KNF$.LPniK!PlILG!WGGKGQGKsFQceLvfaK$gInPImMSAGEL
```

N-linker, Walker A

Block 3:

(markers: ℓ, H3, β3, H4)

```
N_tabacum     140   ESGNAGEPAKLIRQRYREAAEIIRKGNMCCLFINDLDAGAGRMGETTQYTVNNQMVNATLMNIADNPTNVQL
A_thaliana    137   ESGNAGEPAKLIRQRYREAADIIKKGKMCCLFINDLDAGAGRMGGTTQYTVNNQMVNATLMNIADNPTNVQL
O_sativa      139   ESGNAGEPAKLIRQRYREAADIIKKGKMCCLFINDLDAGAGRMGGTTQYTVNNQMVNATLMNIADNPTNVQL
M_domestica   138   ESGNAGEPAKLIRQRYREAADIIRKGKMCALFINDLDAGAGRLGGTTQYTVNNQMVNATLMNIADNPTNVQL
L_tridentata  139   ESGNAGEPAKLIRQRYREAADLIAKGKMCALFINDLDAGAGRLGGTTQYTVNNQMVNATLMNIADNPTNVQL
V_radiata     138   ESGNAGEPAKLIRQRYREAADLIAKGKMCALFINDLDAGAGRLGGTTQYTVNNQMVNATLMNIADNPTNVQL
S_oleracea    135   ESGNAGEPAKLIRQRYREAADLIAKGKMCALFINDLEPGAGRMGGTTQYTVNNQMVNATLMNIADNPTNVQL
P_patens      139   ESGNAGEPAKLIRQRYREAADVIKGKMCALFINDLDAGAGRMGDTTQYTVNNQMVNATLMNIADSPTNVQL
V_carteri     130   ESGNAGEPAKLIRQRYREASDIIKKGKMCSLFINDLDAGAGRMGDTTQYTVNNQMVNATLMNIADNPTNVQL
O_lucimarinus 132   ESGNAGEPAKLIRQRYRGMSTLFINDLDAGAGRGMSTLFINDLaGAGR$GgTTQYTVNNQMVNATLMNIADnPTNVQL
consensus>50        ESGNAGEPAKL!RqRYREAa#i!.KGkMc.LFINDL#aGAGR$GgTTQYTVNNQMVNATLMNIADnPTNVQL
```

pore loop 1, Walker B, pore loop 2

Block 4:

(markers: β4, H5, β5, H6, η1, H7)

```
N_tabacum     212   PGMYNKQENARVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCTGIFRTDNVPAEDVKIVDNFP
A_thaliana    209   PGMYNKEENARVPILCTGNDFSTLYAPLIRDGRMEKFYWAPTREDRVGVCKGIFRTDKIKDEDIVTLVDPFP
O_sativa      211   PGMYNKEDNPRVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRVGVCKGIFRTDNVPDEDIVKIVDSFP
M_domestica   210   PGMYNKEENPRVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCIGIFRSDNVAKEDIVKLVDTFP
L_tridentata  211   PGMYNKEENPRVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCKGIFRTDNVADDDIVKLVDTFP
V_radiata     210   PGMYNKEENARVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRVGVCKGIFRTDGVPEEDITKLVDTFP
S_oleracea    207   PGMYNKQDNARVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCTGIFKTDKVPAEHVVKLVDAFP
P_patens      211   PGVYNKEEIPRVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCKGIFRLDNVSDDNVTKLVDQFP
V_carteri     202   PGVYKNEEIPRVPIVCTGNDFSTLYAPLIRDGRMEKYYWAPTREDRIGVCMGIFQEDNVSRGDVERLVDTFP
O_lucimarinus 204   PGNYEVIEIPRVPIAATGNDFSTLYAPLVRDGRMEKFYWAPTREDRVGIANGIFMADGIEKEDVEVLVDTFE
consensus>50        PGmYnke#n.RVPI!vTGNDFSTLYAPL!RDGRM#K%YWaPTR#DR!G!c.GIFr.D.!..ed!.klVD.Fp
```

sensor 1, pore loop 3

Block 5:

(markers: H8, H9, H10)

```
N_tabacum     284   GQSIDFFGALRARVYDDEVRKWVSGTGIEKIGDKLLNSFDGPPTFEQPKMTIEKLLEYGNMLVQEQENVKR
A_thaliana    281   GQSIDFFGALRARVYDDEVRKFVESLGVEKIGKPLVNSREGPPVFEQPKMTIEKLMEYGNMLVMEQENVKR
O_sativa      283   GQSIDFFGALRARVYDDEVRKWVSDTGVENIGKKLVNSREGPPEFEQPKMTIEKLMEYGYMLVKEQENVKR
M_domestica   282   GQSIDFFGALRARVYDDEVRKWITGVGVDSIGKKLVNSKEGPPTFEQPKMTIEKLLEYGNMLVQEQENVKR
L_tridentata  283   GQSIDFFGALRARVYHDEVRKWVSEVGVDTIGKKLVNSKEGPPSFEQPKMTIDKLLGYGGMLVQEQENVKR
V_radiata     282   GQSIDFFGALRARVYDDEVRKWVSVGVDATGKKLVNSKEGPPTFDQPKMSLDKLLQYGNMLVQEQENVKR
S_oleracea    279   GQSIDFFGALRARVYHDEVRKWVNSVGVDNVGKKLVNSKDGPPVFEQPEMTLQKLMEYGNMLVQEQENVKR
P_patens      283   GQSIDFFGALRARVYDDEVRKWIAEVGVENIGKNLVMEYGEMLVQEQNVKR
V_carteri     274   GQSIDFFGALRARVYDDMVRKWIAEVGIEGISKLVNGR..QKVSFPKVSMSLDVLLKYGRALVDEQENVKR
O_lucimarinus 276   GQSIDFFGALRSRVYDDIVRDFILEVGYEALGPRLINPRKGEEVNFNPPKMTLEVLLAYGKELENEQENIKR
consensus>50        GQSIDFFGALRaRVYdDeVRkw!...Gv#.iGk.LvNs.egpp..Feqp.Mt.#kL$eYG.mLvqEQ#N!KR
```

sensor 2, species specifity

Block 6:

```
N_tabacum     355   VQLADDYLKEAALGDANADAINNGSFFAS.........AQQV.NLPVPEGCTDPVAENFDPTARSDDGTCVYNF
A_thaliana    352   VQLAETYLSQAALGDANADAIGRGTFYGKG.AQQV.NLPVPEGCTDPVAENFDPTARSDDGTCVYNF
O_sativa      354   VQLAEQYLSEAALGDANSDAMKTGSFYGQG.AQQAGNLPVPEGCTDPVAKNFDPTARSDDGSCLYTF
M_domestica   353   VQLADKYLSEAALGDANSDAMNTGTFYG.
L_tridentata  354   VQLADKYMSEAALGDANNDAIKRGTFYGGQAAQQVGNVPVPEGCTDPQATNYDPTARSDDGSCVYKF
V_radiata     353   VQLADKYLNEAALGNANEDAIKSGSFFK.
S_oleracea    350   VQLAEQYLSSAALGDANKDAIKRGTFGKAAQQVS..LPVAQGCTDPEAKNYDPTARSDDGSCTYNL
P_patens      354   VQLAEQYLSSAALGDANADAIKQGTFYAGGQAPASQ.
V_carteri     344   VQLADAYLSGAELAGREGSSLPEEYSRR.
O_lucimarinus 348   IQLADAYLDGAVLAGEGGSSNTEKSLNQ.
consensus>50        !QLA#.Y$s.AaLgdan.da...g.fy............................................
```

Legend:

● ATP binding and hydrolysis ► Rubisco binding ▲ pore loops
★ Arg finger ▼ Rca-Rca interface ◄ thermostability

Figure 5.2: **Sequence alignment of Rubisco activase of different species**

Amino acid sequences of selected Rubisco activase (Rca) homologues from plants and green algae were aligned using Clustal-W. The only Rca from a *Solanaceae* plant listed is Rca of *Nicotiana tabacum*. Most species have two or more Rca isoforms differing in the C-terminal sequence. Secondary structure elements and names for *Nicotiana tabacum* Rca are indicated above the sequences. The Rca subdomain structure is indicated by blue and teal colouring of secondary structure. In the alignment, similar residues are shown in red and identical residues in white using bold lettering on red background. Blue frames indicate homologous regions. The consensus sequence is shown at the bottom. Signature sequences of the AAA+ family are indicated below the sequences by coloured bars. Symbols below the sequence indicate residues for which mutation data are available as detailed in the legend. All sequence numbering is based on the mature proteins, i.e. after the chloroplast signal sequence is removed. The cleavage site was experimentally verified only for sequences 1, 2, and 8.

The Uniprot accession codes for the aligned Rca sequences are: Q40460, *Nicotiana tabacum*; P10896-2, *Arabidopsis thaliana*; P93431, *Oryza sativa subsp. japonica*; Q40281, *Malus domestica*; Q7X9A0, *Larrea tridentata*; O98997, *Vigna radiata var. radiata*; P10871, *Spinacia oleracea*; A9TBP0, *Physcomitrella patens subsp. patens*; D8TZU3, *Volvox carteri f. nagariensis*; A4RW20, *Ostreococcus lucimarinus* (strain CCE9901).

Figure 5.3 shows an overview of the mechanism for Rubisco reactivation. It is interesting to note that, in contrast to previously proposed models [171], structure and diameter of the hexameric arrangement exclude the possibility that a hexamer of Rca surrounds Rubisco.

Mutations in either the N-terminal domain or the specificity residues in helix H9 almost completely abolished Rubisco activation by Rca (this study and [172] [173] [184]). This indicates a mechanism in which the initial recognition of Rubisco is mediated by these domains. Indeed, helix H9 and the N-terminal domain seem to be juxtaposed on the "top" surface of the hexameric arrangement, suggesting an interaction of Rca and exposed surface residues or loop segments of Rubisco (e.g. R89, K94 in *Solanacaea*).

Mutations in loops lining the central pore severely inhibited activation activity and therefore must be actively involved in the remodeling process, suggesting a partial threading mechanism of the substrate in analogy to ClpX or Spastin and other AAA+ protein remodeling enzymes. In comparison the central pore of Rca is much wider (Rca pore diameter 36Å) than the central pore of other hexameric AAA+ proteins, such as ClpX (essentially fully closed [62]) or Spastin (20 Å at the widest point [57]), suggesting that Rca might engage a more complex secondary structure element of Rubisco for remodeling and not a single peptide motif.

ATP-dependent,
transient hexamerization

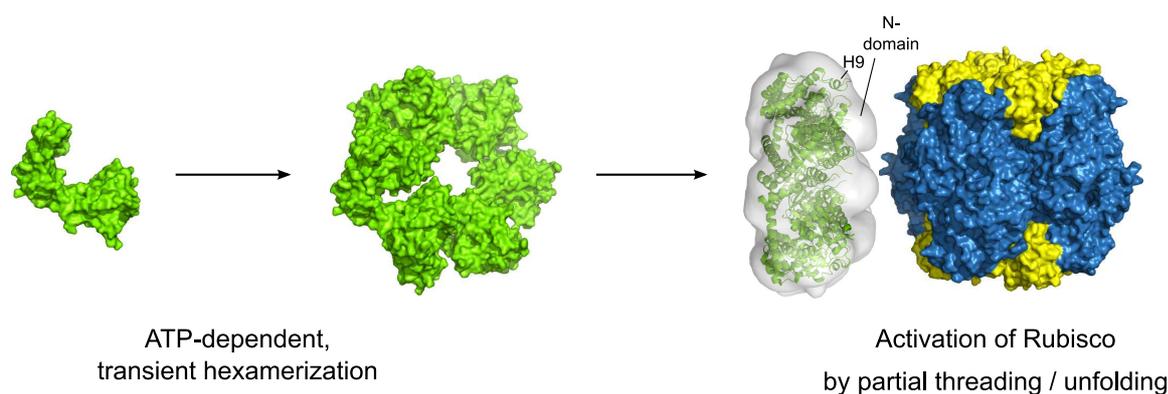Activation of Rubisco
by partial threading / unfolding

Figure 5.3: **Preliminary model of Rubisco activation by Rca.**

Rca forms ATP dependent, transient hexamers in the active state. The initial Rubisco interaction is mediated by helix H9 and the N-terminal domain, before larger secondary structure element of Rubisco is partially unfolded by threading through the central pore.

Surface representation of monomeric crystal structure and hexameric model of Rca shown in green, the N-terminal domain was not part of the crystallised construct. EM density shown in grey. Location of the specificity helix H9 and the potential location of the N-terminal domain are indicated. Surface representation of Rubisco (closed, PDB 4RUB [121]) from Nicotiana tabacum shown with large subunits coloured in blue and small subunits in yellow.

5.4 Possible mechanisms for the regulation of Rubisco activase

In many plants activase (e.g. *Arabidopsis thaliana*, *Spinacia oleracea*, *Oryza sativa*) has two isoforms, a short (40-42 kDa) and long isoform (~46 kDa) usually generated by alternative splicing, which vary in the length of the C-terminal extension . The oxidation of the long isoform of Rca results in a more pronounced ADP inhibition [188], believed to be involved in regulation of the activase activity in response to light irradiation (see also 2.5.2). Residues of the oxidized C-terminus of the long isoform can be crosslinked to residues near the nucleotide binding site where they are supposed to interfere with ATP (but not ADP) binding and hydrolysis [191]. The previous model for redox-regulation suggested that this interaction with the nucleotide binding pocket is only present in the oxidized state [191].

In this study a conserved tyrosine in the C-terminal extension (residue number 361) was identified, which is essential for the ATPase activity of Rca. Mutation of this tyrosine to alanine almost completely abolished ATPase activity and hence also the Rubisco activation. Interestingly, this tyrosine is conserved in both long and short isoforms of Rca (see Figure 5.2) and it can be inferred from the crystal structure that it is close to the nucleotide binding pocket of the next subunit in the oligomer.

The crystal structure as well as crosslinking of N and C-termini [198][191] suggest that the redox-

regulation mechanism is acting in trans in agreement with the positioning of tyrosine 361. It is therefore possible that the extended C-terminus of the long isoform is always near the nucleotide binding pocket of the next subunit, however, upon oxidation the conformation of the tyrosine 361 is altered to inhibit ATPase activity and preferentially bind ADP. One way to distinguish between these different models for the mechanism of redox-regulation would be high resolution crystal structures of the long isoform of Rca in the reduced and oxidized states, respectively.

5.5 Convergent evolution of activases for green- and red-type Rubisco

Recently an activase for red-type Rubisco (form I C-D) from *Rhodobacter sphaeroides*, called CbbX, was identified [233]. CbbX is also a hexameric AAA+ protein but shows only poor sequence homology to Rubisco activase (< 10 % sequence identity). Interestingly, the large subunits of red-type Rubiscos have an extended C-terminal sequence compared to green-type Rubiscos (form I A-B). The red-type activase CbbX engages the closed ER complex, partially threads the extended C-terminus of the large subunit into the central pore and thereby remodels Rubisco [233], mechanistically resembling ClpX or other AAA+ unfoldases. The formation of the active, hexameric complex for CbbX requires binding of ATP and RuBP to CbbX. Since ATPase activity depends on binding of RuBP to reach the hexameric active state and is even further stimulated by the presence of Rubisco, CbbX seems to have evolved a complex regulation of ATPase and activation activity in response to the Rubisco substrate RuBP. Such a regulation is absent in green-type Rubisco activase.

The crystal structures of both enzymes were solved at comparable resolution and are compared in Figure 5.4. While the NBDs of both AAA+ show the expected similarities, the C-terminal α-helical subdomain seems more structurally diverse. Notably, the helical insertion (H9) responsible for species specificity in green-type Rubisco activase is absent in CbbX.

Since negative stain EM reconstructions of both enzymes are available, the green- and red-type Rubisco activases can also be compared in the active, hexameric state (ATPγS bound NtRca R294V and RuBP ATPγS bound CbbX, see Figure 5.5). Most prominently the central pore in the hexamer in CbbX appears much narrower than in Rca (25 Å vs 36 Å). Additionally, the N-terminal domain of Rca is clearly protruding on "top" of the ring, whereas CbbX lacks this domain and appears to be almost flat (Rca height 56 Å vs CbbX height 45 Å, see Figure 5.5).
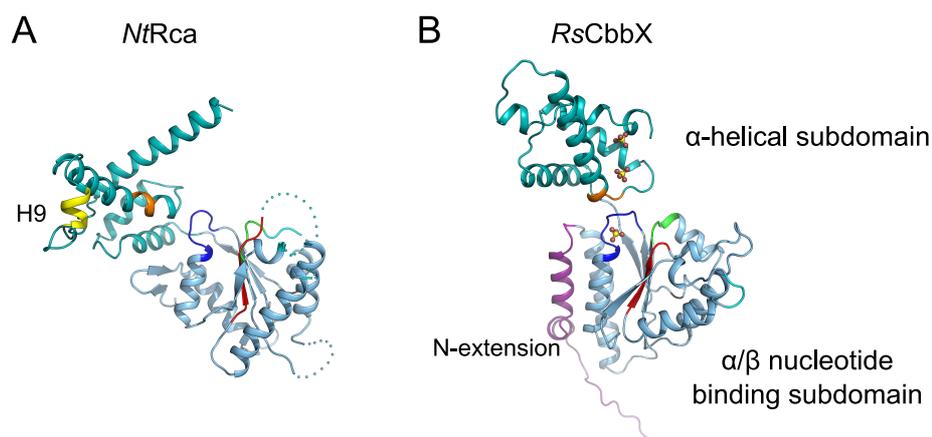
Figure 5.4: **Comparison of the monomeric structures of *Nicotiana tabacum* Rca and *Rhodobacter sphaeroides* CbbX.**
A: Ribbon representation of crystal structure of the green-type Rubisco activase *Nicotiana tabacum* Rca B: Ribbon representation of the crystal structure of the red-type Rubisco activase CbbX [PDB ID: 3SYL] in the same orientation of the nucleotide binding domain. α/β nucleotide binding subdomains are shown in light blue, the α-helical subdomains are shown in teal. The Walker A and B motifs are shown in dark blue and red, and the sensor I and II regions in green and orange, respectively. The pore loop(s) are indicated in cyan. The helical insertion in the C-terminal α-helical subdomain (H9) in Rca is shown in yellow. Bound sulphates are shown in ball- and stick-form for CbbX. The N-terminal extension of CbbX is shown in purple.

In summary, biochemical experiments as well as a direct comparison of the structures suggest that both activases are ATP dependent remodeling enzymes employing a different mechanism. While CbbX is able to engage a single extended C-terminal peptide of the large subunit of Rubisco and threads it through the narrow pore, the wider pore in Rca suggests that Rca engages a more complex secondary structure element for remodeling. Additionally, the initial interaction of Rca with Rubisco seems to be mediated by the N-terminal domain and the specificity helix H9, both of which have no counterpart in CbbX. Interestingly, different mechanisms of regulation have evolved in the green- and red-type Rubisco activases. The red-type activase, CbbX, is regulated by binding of RuBP, the substrate of its target protein and its ATPase activity is stimulated in the presence of Rubisco. The green-type Rubisco activase on the other hand is regulated by inhibition by ADP and redox-regulation of the long isoform.

The mechanistic, structural and regulatory differences of CbbX and Rca suggest that two different AAA+ proteins were initially recruited for the activation of Rubisco. Furthermore, Rca and CbbX show a rather poor sequence homology beyond the core AAA+ sequence motives, whereas their respective
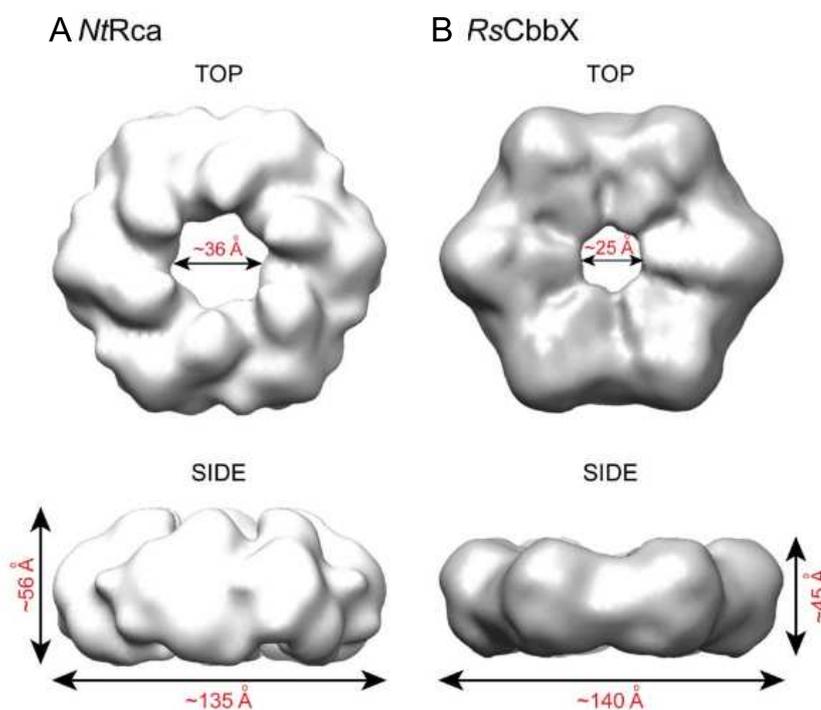
Figure 5.5: **Comparison of the negative stain electron microscopy reconstructions of hexameric Rca and CbbX.**

A: Top and side view of a negative stain EM reconstruction of Rca. B: Top and side view of a negative stain EM reconstruction of CbbX [233]. Both AAA+ show a hexameric complex with similar outer diameters, however, CbbX features a narrower pore and the density attributed to the N-terminal domain of Rca is expectedly absent in CbbX. Pore diameter and height of the complexes are indicated in red.

substrates form I A and B (green-type) and form I C and D (red-type) Rubisco are highly homologous [115]. It seems unlikely that while the Rubisco sequences diverged only little during evolution, the activating AAA+ diverged so far, therefore CbbX and Rca rather seem to have evolved convergently.

5.6 Implications and outlook

This work presents the structure of the AAA+ core of *Nicotiana tabacum* Rca in the nucleotide free state. Further structural studies either employing X-ray crystallography or cryo-EM with bound nucleotides or nucleotide analogues might contribute to a more detailed view of the conformational changes in response to the nucleotide state during Rubisco activation and also provide further insights into the general mechanism of AAA+ proteins.

An important future goal for understanding Rubisco activation by Rca will be the characterization of the Rubisco-Rca interaction. To this end, heterologous complexes of Rca (or Rca-like sequences from cyanobacteria) with Rubisco might present a strategy to produce stable, "trapped" complexes as observed for the assembly chaperone RbcX and Rubisco [139]. Furthermore, structural and biochemical studies of the N-terminal domain of Rca of other species might provide evidence for the proposed role of the N-terminal domain as substrate adaptor.

The identification of a conserved residue in the C-terminal extension, that is essential for the ATPase activity has lead to a new hypothesis regarding the redox-regulation of the long isoform of Rubisco activase. Since to date only a short isoform of Rca was identified in *Nicotiana tabacum*, structural studies of long isoforms of other species in the oxidized and reduced states, might provide more insight into this important regulation mechanism.

The biochemical and structural data presented in this work lead to a preliminary model for the mechanism of Rubisco activase and provides the structural framework for further studies on Rca. Additionally, structure guided mutations to improve the heat stability of Rca (e.g. optimizing the packing of the hydrophobic core or generation of intramolecular salt bridges) might be employed to improve the photosynthetic properties of genetically manipulated crops under moderate heat stress, ultimately leading to higher crop yields in climate conditions expected for the near future [163] [162].

6 References

[1] Anfinsen, C. B., Redfield, R. R., Choate, W. L., Page, J., and Carroll, W. R. Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease. *J Biol Chem* **207**(1), 201–210 (1954).

[2] Anfinsen, C. B. and Haber, E. Studies on the reduction and re-formation of protein disulfide bonds. *J Biol Chem* **236**, 1361–1363 (1961).

[3] Haber, E. and Anfinsen, C. B. Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. *J Biol Chem* **237**, 1839–1844 (1962).

[4] Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**(96), 223–230 (1973).

[5] Mayer, M. P. Gymnastics of molecular chaperones. *Mol Cell* **39**(3), 321–331 (2010).

[6] Levinthal, C. Are there pathways for protein folding? *J Chim Phys* , **65**, 44–45 (1968).

[7] Hartl, F. U., Bracher, A., and Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**(7356), 324–332 (2011).

[8] Ellis, R. J. and Minton, A. P. Cell biology: join the crowd. *Nature* **425**(6953), 27–28 (2003).

[9] Ellis, R. J. and Minton, A. P. Protein aggregation in crowded environments. *Biol Chem* **387**(5), 485–497 (2006).

[10] Hartl, F. U. and Hayer-Hartl, M. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**(5561), 1852–1858 (2002).

[11] Hartl, F. U. and Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* **16**(6), 574–581 (2009).

[12] Rutherford, S. L. and Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **396**(6709), 336–342 (1998).

[13] Tokuriki, N. and Tawfik, D. S. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* **459**(7247), 668–673 (2009).

[14] Kettern, N., Dreiseidler, M., Tawo, R., and Höhfeld, J. Chaperone-assisted degradation: multiple paths to destruction. *Biol Chem* **391**(5), 481–489 (2010).

[15] Morley, J. F. and Morimoto, R. I. Regulation of longevity in Caenorhabditis elegans by heat shock factor and molecular chaperones. *Mol Biol Cell* **15**(2), 657–664 (2004).

[16] Olzscha, H., Schermann, S. M., Woerner, A. C., Pinkert, S., Hecht, M. H., Tartaglia, G. G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F. U., and Vabulas, R. M. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144**(1), 67–78 (2011).

[17] Ellis, R. J. Proteins as molecular chaperones. *Nature* **328**, 378–379 (1987).

[18] Parsell, D. A., Kowal, A. S., Singer, M. A., and Lindquist, S. Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* **372**(6505), 475–478 (1994).

[19] Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. The crystal structure of the bacterial chaperonin GroEL at 2.8 A. *Nature* **371**(6498), 578–586 (1994).

[20] Braig, K., Adams, P. D., and BrÃ$\frac{1}{4}$nger, A. T. Conformational variability in the refined structure of the chaperonin GroEL at 2.8 A resolution. *Nat Struct Biol* **2**(12), 1083–1094 (1995).

[21] Hayer-Hartl, M. K., Weber, F., and Hartl, F. U. Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. *EMBO J* **15**(22), 6111–6121 (1996).

[22] Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **381**(6583), 571–579 (1996).

[23] Mayhew, M., da Silva, A. C., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* **379**(6564), 420–426 (1996).

[24] Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M., and Horwich, A. L. Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* **84**(3), 481–490 (1996).

[25] Xu, Z., Horwich, A. L., and Sigler, P. B. The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. *Nature* **388**(6644), 741–750 (1997).

[26] Saibil, H. R., Zheng, D., Roseman, A. M., Hunter, A. S., Watson, G. M., Chen, S., Mauer, A. A. D., O'Hara, B. P., Wood, S. P., Mann, N. H., Barnett, L. K., and Ellis, R. J. ATP induces large quaternary rearrangements in a cage-like chaperonin structure. *Curr Biol* **3**(5), 265–273 (1993).

[27] Ellis, R. J. Molecular chaperones. Opening and closing the Anfinsen cage. *Curr Biol* **4**(7), 633–635 (1994).

[28] Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U., and Hayer-Hartl, M. Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107**(2), 223–233 (2001).

[29] Tang, Y.-C., Chang, H.-C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., Hartl, F. U., and Hayer-Hartl, M. Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **125**(5), 903–914 (2006).

[30] Chakraborty, K., Chatila, M., Sinha, J., Shi, Q., Poschner, B. C., Sikor, M., Jiang, G., Lamb, D. C., Hartl, F. U., and Hayer-Hartl, M. Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* **142**(1), 112–122 (2010).

[31] 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., and Hartl, F. U. Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell* **122**(2), 209–220 (2005).

[32] Douglas, N. R., Reissmann, S., Zhang, J., Chen, B., Jakana, J., Kumar, R., Chiu, W., and Frydman, J. Dual action of ATP hydrolysis couples lid closure to substrate release into the group II chaperonin chamber. *Cell* **144**(2), 240–252 (2011).

[33] Ditzel, L., LÃ¶we, J., Stock, D., Stetter, K. O., Huber, H., Huber, R., and Steinbacher, S. Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell* **93**(1), 125–138 (1998).

[34] Knapp, S., Schmidt-Krey, I., Hebert, H., Bergman, T., Jörnvall, H., and Ladenstein, R. The molecular chaperonin TF55 from the Thermophilic archaeon Sulfolobus solfataricus. A biochemical and structural characterization. *J Mol Biol* **242**(4), 397–407 (1994).

[35] Munoz, I. G., Yébenes, H., Zhou, M., Mesa, P., Serna, M., Park, A. Y., Bragado-Nilsson, E., Beloso, A., de Cárcer, G., Malumbres, M., Robinson, C. V., Valpuesta, J. M., and Montoya, G. Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. *Nat Struct Mol Biol* **18**(1), 14–19 (2011).

[36] Gomez-Puertas, P., Martin-Benito, J., Carrascosa, J. L., Willison, K. R., and Valpuesta, J. M. The substrate recognition mechanisms in chaperonins. *J Mol Recognit* **17**(2), 85–94 (2004).

[37] Yam, A. Y., Xia, Y., Lin, H.-T. J., Burlingame, A., Gerstein, M., and Frydman, J. Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat Struct Mol Biol* **15**(12), 1255–1262 (2008).

[38] Reissmann, S., Parnot, C., Booth, C. R., Chiu, W., and Frydman, J. Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. *Nat Struct Mol Biol* **14**(5), 432–440 (2007).

[39] Rüdiger, S., Buchberger, A., and Bukau, B. Interaction of Hsp70 chaperones with substrates. *Nat Struct Biol* **4**(5), 342–349 (1997).

[40] Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* **346**(6285), 623–628 (1990).

[41] Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **272**(5268), 1606–1614 (1996).

[42] Liu, Q. and Hendrickson, W. A. Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **131**(1), 106–120 (2007).

[43] Polier, S., Dragovic, Z., Hartl, F. U., and Bracher, A. Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* **133**(6), 1068–1079 (2008).

[44] Zhuravleva, A. and Gierasch, L. M. Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. *Proc Natl Acad Sci U S A* **108**(17), 6987–6992 (2011).

[45] Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M., and Gierasch, L. M. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol Cell* **26**(1), 27–39 (2007).

[46] Bertelsen, E. B., Chang, L., Gestwicki, J. E., and Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc Natl Acad Sci U S A* **106**(21), 8471–8476 (2009).

[47] Erdmann, R., Wiebel, F. F., Flessau, A., Rytka, J., Beyer, A., Fröhlich, K. U., and Kunau, W. H. PAS1, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. *Cell* **64**(3), 499–510 (1991).

[48] Kunau, W. H., Beyer, A., Franken, T., Götte, K., Marzioch, M., Saidowsky, J., Skaletz-Rorowski, A., and Wiebel, F. F. Two complementary approaches to study peroxisome biogenesis in Saccharomyces cerevisiae: forward and reversed genetics. *Biochimie* **75**(3-4), 209–224 (1993).

[49] Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**(1), 27–43 (1999).

[50] Lenzen, C. U., Steinmann, D., Whiteheart, S. W., and Weis, W. I. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**(4), 525–536 (1998).

[51] Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C., and Stuckey, P. J. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics* **27**(23), 3315–3316 (2011).

[52] Vale, R. D. AAA proteins. Lords of the ring. *J Cell Biol* **150**(1), F13–F19 (2000).

[53] Hanson, P. I. and Whiteheart, S. W. AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6**(7), 519–529 (2005).

[54] Erzberger, J. P. and Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu Rev Biophys Biomol Struct* **35**, 93–114 (2006).

[55] Davies, J. M., Brunger, A. T., and Weis, W. I. Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* **16**(5), 715–726 (2008).

[56] Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., and Shi, Y. Structure and mechanism of the hexameric MecA-ClpC molecular machine. *Nature* **471**(7338), 331–335 (2011).

[57] Roll-Mecak, A. and Vale, R. D. Structural basis of microtubule severing by the hereditary spastic paraplegia protein spastin. *Nature* **451**(7176), 363–367 (2008).

[58] Wendler, P., Ciniawsky, S., Kock, M., and Kube, S. Structure and function of the AAA+ nucleotide binding pocket. *Biochim Biophys Acta* **1823**(1), 2–14 (2012).

[59] Zhang, X. and Wigley, D. B. The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. *Nat Struct Mol Biol* **11**, 1223–1227 (2008).

[60] Mogni, M. E., Costa, A., Ioannou, C., and Bell, S. D. The glutamate switch is present in all seven clades of AAA+ protein. *Biochemistry* **48**(37), 8774–8775 (2009).

[61] Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* **146**(1-2), 11–31 (2004).

[62] Glynn, S. E., Martin, A., Nager, A. R., Baker, T. A., and Sauer, R. T. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* **139**(4), 744–756 (2009).

[63] Martin, A., Baker, T. A., and Sauer, R. T. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature* **437**(7062), 1115–1120 (2005).

[64] Kondo, H., Rabouille, C., Newman, R., Levine, T. P., Pappin, D., Freemont, P., and Warren, G. p47 is a cofactor for p97-mediated membrane fusion. *Nature* **388**(6637), 75–78 (1997).

[65] Uchiyama, K., Totsukawa, G., Puhka, M., Kaneko, Y., Jokitalo, E., Dreveny, I., Beuron, F., Zhang, X., Freemont, P., and Kondo, H. p37 is a p97 adaptor required for Golgi and ER biogenesis in interphase and at the end of mitosis. *Dev Cell* **11**(6), 803–816 (2006).

[66] Meyer, H. H., Shorter, J. G., Seemann, J., Pappin, D., and Warren, G. A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways. *EMBO J* **19**(10), 2181–2192 (2000).

[67] Siddiqui, S. M., Sauer, R. T., and Baker, T. A. Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. *Genes Dev* **18**(4), 369–374 (2004).

[68] Martin, A., Baker, T. A., and Sauer, R. T. Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell* **29**(4), 441–450 (2008).

[69] Martin, A., Baker, T. A., and Sauer, R. T. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* **15**(11), 1147–1151 (2008).

[70] Goloubinoff, P., Mogk, A., Zvi, A. P., Tomoyasu, T., and Bukau, B. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Natl Acad Sci U S A* **96**(24), 13732–13737 (1999).

[71] Lee, S.-Y., Torre, A. D. L., Yan, D., Kustu, S., Nixon, B. T., and Wemmer, D. E. Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev* **17**(20), 2552–2563 (2003).

[72] Mogk, A., Schlieker, C., Strub, C., Rist, W., Weibezahn, J., and Bukau, B. Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. *J Biol Chem* **278**(20), 17615–17624 (2003).

[73] Schirmer, E. C., Homann, O. R., Kowal, A. S., and Lindquist, S. Dominant gain-of-function mutations in Hsp104p reveal crucial roles for the middle region. *Mol Biol Cell* **15**(5), 2061–2072 (2004).

[74] Haslberger, T., Weibezahn, J., Zahn, R., Lee, S., Tsai, F. T. F., Bukau, B., and Mogk, A. M domains couple the ClpB threading motor with the DnaK chaperone activity. *Mol Cell* **25**(2), 247–260 (2007).

[75] Wendler, P., Shorter, J., Plisson, C., Cashikar, A. G., Lindquist, S., and Saibil, H. R. Atypical AAA+ subunit packing creates an expanded cavity for disaggregation by the protein-remodeling factor Hsp104. *Cell* **131**(7), 1366–1377 (2007).

[76] Wendler, P., Shorter, J., Snead, D., Plisson, C., Clare, D. K., Lindquist, S., and Saibil, H. R. Motor mechanism for protein threading through Hsp104. *Mol Cell* **34**(1), 81–92 (2009).

[77] Lee, S., Choi, J.-M., and Tsai, F. T. F. Visualizing the ATPase cycle in a protein disaggregating machine: structural basis for substrate binding by ClpB. *Mol Cell* **25**(2), 261–271 (2007).

[78] Wendler, P. and Saibil, H. R. Cryo electron microscopy structures of Hsp100 proteins: crowbars in or out? *Biochem Cell Biol* **88**(1), 89–96 (2010).

[79] Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E. U., Dougan, D. A., Tsai, F. T. F., Mogk, A., and Bukau, B. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* **119**(5), 653–665 (2004).

[80] Chernoff, Y. O., Lindquist, S. L., Ono, B., Inge-Vechtomov, S. G., and Liebman, S. W. Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi+]. *Science* **268**(5212), 880–884 (1995).

[81] Halfmann, R., Jarosz, D. F., Jones, S. K., Chang, A., Lancaster, A. K., and Lindquist, S. Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature* **482**(7385), 363–368 (2012).

[82] Sagan, L. On the origin of mitosing cells. *J Theor Biol* **14**(3), 255–274 (1967).

[83] Soll, J. and Schleiff, E. Protein import into chloroplasts. *Nat Rev Mol Cell Biol* **5**(3), 198–208 (2004).

[84] Hohmann-Marriott, M. F. and Blankenship, R. E. Evolution of photosynthesis. *Annu Rev Plant Biol* **62**, 515–548 (2011).

[85] Merchant, S. and Sawaya, M. R. The light reactions: a guide to recent acquisitions for the picture gallery. *Plant Cell* **17**(3), 648–663 (2005).

[86] Cruz, J. A., Avenson, T. J., Kanazawa, A., Takizawa, K., Edwards, G. E., and Kramer, D. M. Plasticity in light reactions of photosynthesis for energy production and photoprotection. *J Exp Bot* **56**(411), 395–406 (2005).

[87] Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **411**(6840), 909–917 (2001).

[88] Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from Rhodopseudomonas viridis. *J Mol Biol* **180**(2), 385–398 (1984).

[89] Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., and Iwata, S. Architecture of the photosynthetic oxygen-evolving center. *Science* **303**(5665), 1831–1838 (2004).

[90] Kern, J., Loll, B., Zouni, A., Saenger, W., Irrgang, K.-D., and Biesiadka, J. Cyanobacterial photosystem II at 3.2 Å resolution - the plastoquinone binding pockets. *Photosynth Res* **84**(1-3), 153–159 (2005).

[91] Umena, Y., Kawakami, K., Shen, J.-R., and Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**(7345), 55–60 (2011).

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

[93] Campbell, N. A. and Reece, J. B. *Biology*. Benjamin Cummings, 6th edition, (2001).

[94] Cegelski, L. and Schaefer, J. NMR determination of photorespiration in intact leaves using in vivo 13CO2 labeling. *J Magn Reson* **178**(1), 1–10 (2006).

[95] Bauwe, H. Recent developments in photorespiration research. *Biochem Soc Trans* **38**(2), 677–682 (2010).

[96] Peterhansel, C. and Maurino, V. G. Photorespiration redesigned. *Plant Physiol* **155**(1), 49–55 (2011).

[97] Long, S. P., Ainsworth, E. A., Leakey, A. D. B., Nösberger, J., and Ort, D. R. Food for thought: lower-than-expected crop yield stimulation with rising CO2 concentrations. *Science* **312**(5782), 1918–1921 (2006).

[98] Taler, D., Galperin, M., Benjamin, I., Cohen, Y., and Kenigsbuch, D. Plant eR genes that encode photorespiratory enzymes confer resistance against disease. *Plant Cell* **16**(1), 172–184 (2004).

[99] Bloom, A. J., Burger, M., Asensio, J. S. R., and Cousins, A. B. Carbon dioxide enrichment inhibits nitrate assimilation in wheat and Arabidopsis. *Science* **328**(5980), 899–903 (2010).

[100] Price, G. D., Badger, M. R., Woodger, F. J., and Long, B. M. Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J Exp Bot* **59**(7), 1441–1461 (2008).

[101] Bonacci, W., Teng, P. K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P. A., and Savage, D. F. Modularity of a carbon-fixing protein organelle. *Proc Natl Acad Sci U S A* **109**(2), 478–483 (2012).

[102] Moroney, J. V. and Ynalvez, R. A. Proposed carbon dioxide concentrating mechanism in Chlamydomonas reinhardtii. *Eukaryot Cell* **6**(8), 1251–1259 (2007).

[103] Badger, M. R., Andrews, T. J., Whitney, S. M., Ludwig, M., Yellowlees, D. C., Leggat, W., and Price, G. D. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO2-concentrating mechanisms in algae. *Canadian Journal of Botany* **76**(6), 1052–1071 (1998).

[104] Edwards, E. J., Osborne, C. P., Strömberg, C. A. E., Smith, S. A., Consortium, C. G., Bond, W. J., Christin, P.-A., Cousins, A. B., Duvall, M. R., Fox, D. L., Freckleton, R. P., Ghannoum, O., Hartwell, J., Huang, Y., Janis, C. M., Keeley, J. E., Kellogg, E. A., Knapp, A. K., Leakey, A. D. B., Nelson, D. M., Saarela, J. M., Sage, R. F., Sala, O. E., Salamin, N., Still, C. J., and Tipple, B. The origins of C4 grasslands: integrating evolutionary and ecosystem science. *Science* **328**(5978), 587–591 (2010).

[105] Brown, N. J., Newell, C. A., Stanley, S., Chen, J. E., Perrin, A. J., Kajala, K., and Hibberd, J. M. Independent and parallel recruitment of preexisting mechanisms underlying C4 photosynthesis. *Science* **331**(6023), 1436–1439 (2011).

[106] Kajala, K., Covshoff, S., Karki, S., Woodfield, H., Tolley, B. J., Dionora, M. J. A., Mogul, R. T., Mabilangan, A. E., Danila, F. R., Hibberd, J. M., and Quick, W. P. Strategies for engineering a two-celled C(4) photosynthetic pathway into rice. *J Exp Bot* **62**(9), 3001–3010 (2011).

[107] Dodd, A. N., Borland, A. M., Haslam, R. P., Griffiths, H., and Maxwell, K. Crassulacean acid metabolism: plastic, fantastic. *J Exp Bot* **53**(369), 569–580 (2002).

[108] Ellis, R. The most abundant protein in the world. *Trends in Biochemical Sciences* **4**, 241–44 (1979).

[109] Spreitzer, R. J. and Salvucci, M. E. Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu Rev Plant Biol* **53**, 449–475 (2002).

[110] Mann, C. C. Genetic engineers aim to soup up crop photosynthesis. *Science* **283**(5400), 314–316 (1999).

[111] Mueller-Cajar, O. and Whitney, S. M. Directing the evolution of Rubisco and Rubisco activase: first impressions of a new tool for photosynthesis research. *Photosynth Res* **98**(1-3), 667–675 (2008).

[112] Uemura, K., Anwaruzzaman, Miyachi, S., and Yokota, A. Ribulose-1,5-bisphosphate carboxylase/oxygenase from thermophilic red algae with a strong specificity for CO2 fixation. *Biochem Biophys Res Commun* **233**(2), 568–571 (1997).

[113] Read, B. A. and Tabita, F. R. High substrate specificity factor ribulose bisphosphate carboxylase/oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial RubiSCO containing "algal" residue modifications. *Arch Biochem Biophys* **312**(1), 210–218 (1994).

[114] Whitney, S. M., Baldet, P., Hudson, G. S., and Andrews, T. J. Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J* **26**(5), 535–547 (2001).

[115] Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E., and Scott, S. S. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J Exp Bot* **59**(7), 1515–1524 (2008).

[116] Taylor, T. C. and Andersson, I. The structure of the complex between rubisco and its natural substrate ribulose 1,5-bisphosphate. *J Mol Biol* **265**(4), 432–444 (1997).

[117] Schneider, G., Lindqvist, Y., and Lundqvist, T. Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from Rhodospirillum rubrum at 1.7 Å resolution. *J Mol Biol* **211**(4), 989–1008 (1990).

[118] Kitano, K., Maeda, N., Fukui, T., Atomi, H., Imanaka, T., and Miki, K. Crystal structure of a novel-type archaeal rubisco with pentagonal symmetry. *Structure* **9**(6), 473–481 (2001).

[119] Li, H., Sawaya, M. R., Tabita, F. R., and Eisenberg, D. Crystal structure of a RuBisCO-like protein from the green sulfur bacterium Chlorobium tepidum. *Structure* **13**(5), 779–789 (2005).

[120] Andersson, I. and Backlund, A. Structure and function of Rubisco. *Plant Physiol Biochem* **46**(3), 275–291 (2008).

[121] Suh, S. W., Cascio, D., Chapman, M. S., and Eisenberg, D. A crystal form of ribulose-1,5-bisphosphate carboxylase/oxygenase from Nicotiana tabacum in the activated state. *J Mol Biol* **197**(2), 363–365 (1987).

[122] Newman, J., Branden, C. I., and Jones, T. A. Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301. *Acta Crystallogr D Biol Crystallogr* **49**(Pt 6), 548–560 (1993).

[123] Taylor, T. C., Backlund, A., Bjorhall, K., Spreitzer, R. J., and Andersson, I. First crystal structure of Rubisco from a green alga, Chlamydomonas reinhardtii. *J Biol Chem* **276**(51), 48159–48164 (2001).

[124] Sugawara, H., Yamamoto, H., Shibata, N., Inoue, T., Okada, S., Miyake, C., Yokota, A., and Kai, Y. Crystal structure of carboxylase reaction-oriented ribulose 1, 5-bisphosphate carboxylase/oxygenase from a thermophilic red alga, Galdieria partita. *J Biol Chem* **274**(22), 15655–15661 (1999).

[125] van Lun, M., van der Spoel, D., and Andersson, I. Subunit interface dynamics in hexadecameric rubisco. *J Mol Biol* **411**(5), 1083–1098 (2011).

[126] Duff, A. P., Andrews, T. J., and Curmi, P. M. 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**(5), 903–916 (2000).

[127] Spreitzer, R. J. Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch Biochem Biophys* **414**(2), 141–149 (2003).

[128] Spreitzer, R. J., Peddi, S. R., and Satagopan, S. Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proc Natl Acad Sci U S A* **102**(47), 17225–17230 (2005).

[129] Schneider, G., Lindqvist, Y., Branden, C. I., and Lorimer, G. Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum at 2.9 Å resolution. *EMBO J* **5**(13), 3409–3415 (1986).

[130] Watson, G. M., Yu, J. P., and Tabita, F. R. Unusual ribulose 1,5-bisphosphate carboxylase/oxygenase of anoxic Archaea. *J Bacteriol* **181**(5), 1569–1575 (1999).

[131] Maeda, N., Kitano, K., Fukui, T., Ezaki, S., Atomi, H., Miki, K., and Imanaka, T. Ribulose bisphosphate carboxylase/oxygenase from the hyperthermophilic archaeon Pyrococcus kodakaraensis KOD1 is composed solely of large subunits and forms a pentagonal structure. *J Mol Biol* **293**(1), 57–66 (1999).

[132] Ashida, H., Saito, Y., Kojima, C., Kobayashi, K., Ogasawara, N., and Yokota, A. A functional link between RuBisCO-like protein of Bacillus and photosynthetic RuBisCO. *Science* **302**(5643), 286–290 (2003).

[133] Imker, H. J., Fedorov, A. A., Fedorov, E. V., Almo, S. C., and Gerlt, J. A. Mechanistic diversity in the RuBisCO superfamily: the "enolase" in the methionine salvage pathway in Geobacillus kaustophilus. *Biochemistry* **46**(13), 4077–4089 (2007).

[134] Goloubinoff, P., Gatenby, A. A., and Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli. *Nature* **337**(6202), 44–47 (1989).

[135] Goloubinoff, P., Christeller, J. T., Gatenby, A. A., and Lorimer, G. H. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfoleded state depends on two chaperonin proteins and Mg-ATP. *Nature* **342**(6252), 884–889 (1989).

[136] Li, L. A. and Tabita, F. R. 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**(11), 3793–3796 (1997).

[137] Saschenbrecker, S., Bracher, A., Rao, K. V., Rao, B. V., Hartl, F. U., and Hayer-Hartl, M. Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. *Cell* **129**(6), 1189–1200 (2007).

[138] Liu, C., Young, A. L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B. V., Rao, K. V., Berninghausen, O., Mielke, T., Hartl, F. U., Beckmann, R., and Hayer-Hartl, M. Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* **463**(7278), 197–202 (2010).

[139] Bracher, A., Starling-Windhof, A., Hartl, F. U., and Hayer-Hartl, M. Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nat Struct Mol Biol* **18**(8), 875–880 (2011).

[140] Pon, N. G., Rabin, B. R., and Calvin, M. Mechanism of Carboxydismutase reaction. 1. Effect of preliminary incubation of substrates, metal ion and enzyme on activity. *Biochemische Zeitschrift* **338**, 7– (1963).

[141] Lorimer, G. H., Badger, M. R., and Andrews, T. J. The activation of ribulose-1,5-bisphosphate carboxylase by carbon dioxide and magnesium ions. Equilibria, kinetics, a suggested mechanism, and physiological implications. *Biochemistry* **15**(3), 529–536 (1976).

[142] Lorimer, G. H. and Miziorko, H. M. Carbamate formation on the epsilon-amino group of a lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by CO_2 and Mg^{2+}. *Biochemistry* **19**(23), 5321–5328 (1980).

[143] Andersson, I. Large structure at high resolution: spinach ribulose 1,5-bisphosphate carboxylase/oxygenase at 1.6 Å resolution. *Journal of Molecular Biology* **259**, 160–174 (1996).

[144] King, W. A., Gready, J. E., and Andrews, T. J. Quantum chemical analysis of the enolization of ribulose bisphosphate: the first hurdle in the fixation of CO2 by Rubisco. *Biochemistry* **37**(44), 15414–15422 (1998).

[145] Andrews, T. J. and Lorimer, G. H. *Rubisco: structure, mechanisms, and prospects for improvement.*, volume 10 of *In The Biochemistry of Plants: A Comprehensive Treatise*. New York: Academic Press, (1987).

[146] Chen, Z. X. and Spreitzer, R. J. Chloroplast intragenic suppression enhances the low CO2/O2 specificity of mutant ribulose-bisphosphate carboxylase/oxygenase. *J Biol Chem* **264**(6), 3051–3053 (1989).

[147] Karkehabadi, S., Satagopan, S., Taylor, T. C., Spreitzer, R. J., and Andersson, I. Structural analysis of altered large-subunit loop-6/carboxy-terminus interactions that influence catalytic efficiency and CO2/O2 specificity of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Biochemistry* **46**(39), 11080–11089 (2007).

[148] Satagopan, S. and Spreitzer, R. J. Substitutions at the Asp-473 latch residue of chlamydomonas ribulosebisphosphate carboxylase/oxygenase cause decreases in carboxylation efficiency and CO(2)/O(2) specificity. *J Biol Chem* **279**(14), 14240–14244 (2004).

[149] Edmondson, D. L., Kane, H. J., and Andrews, T. J. Substrate isomerization inhibits ribulosebisphospate oxygenase during catalysis. *FEBS Letters* **260(1)**, 62–66 (1990).

[150] Moreno, J., Garcia-Murria, M. J., and Marin-Navarro, J. Redox modulation of Rubisco conformation and activity through its cysteine residues. *J Exp Bot* **59**(7), 1605–1614 (2008).

[151] Parry, M. A. J., Keys, A. J., Madgwick, P. J., Carmo-Silva, A. E., and Andralojc, P. J. Rubisco regulation: a role for inhibitors. *J Exp Bot* **59**(7), 1569–1580 (2008).

[152] Marcus, Y., Altman-Gueta, H., Finkler, A., and Gurevitz, M. Mutagenesis at two distinct phosphate-binding sites unravels their differential roles in regulation of Rubisco activation and catalysis. *J Bacteriol* **187**(12), 4222–4228 (2005).

[153] Ishijima, S., Uchibori, A., Takagi, H., Maki, R., and Ohnishi, M. Light-induced increase in free Mg2+ concentration in spinach chloroplasts: measurement of free Mg2+ by using a fluorescent probe and necessity of stromal alkalinization. *Arch Biochem Biophys* **412**(1), 126–132 (2003).

[154] Salvucci, M. E., Portis, A. R., and Ogren, W. L. A soluble chloroplast protein catalyzes ribulose-bisphosphate carboxylase/oxygenase activation in vivo. *Photosynthesis Research* **60**, 193 (1985).

[155] Somerville, C. R., Portis, A. R., and Ogren, W. L. A Mutant of Arabidopsis thaliana Which Lacks Activation of RuBP Carboxylase In Vivo. *Plant Physiol* **70**(2), 381–387 (1982).

[156] Jr, A. R. P. and Salvucci, M. E. The discovery of Rubisco activase - yet another story of serendipity. *Photosynth Res* **73**(1-3), 257–264 (2002).

[157] Streusand, V. J. and Portis, A. R. Rubisco Activase Mediates ATP-Dependent Activation of Ribulose Bisphosphate Carboxylase. *Plant Physiol* **85**(1), 152–154 (1987).

[158] Robinson, S. P., Streusand, V. J., Chatfield, J. M., and Portis, A. R. Purification and Assay of Rubisco Activase from Leaves. *Plant Physiol* **88**(4), 1008–1014 (1988). Rubisco Activase assay.

[159] Robinson, S. P. and Portis, A. R. Adenosine triphosphate hydrolysis by purified rubisco activase. *Arch Biochem Biophys* **268**(1), 93–99 (1989).

[160] Salvucci, M. E., Osteryoung, K. W., Crafts-Brandner, S. J., and Vierling, E. Exceptional sensitivity of Rubisco activase to thermal denaturation in vitro and in vivo. *Plant Physiol* **127**(3), 1053–1064 (2001).

[161] Salvucci, M. E. and Crafts-Brandner, S. J. Relationship between the heat tolerance of photosynthesis and the thermal stability of rubisco activase in plants from contrasting thermal environments. *Plant Physiol* **134**(4), 1460–1470 (2004).

[162] Kurek, I., Chang, T. K., Bertain, S. M., Madrigal, A., Liu, L., Lassner, M. W., and Zhu, G. Enhanced Thermostability of Arabidopsis Rubisco activase improves photosynthesis and growth rates under moderate heat stress. *Plant Cell* **19**(10), 3230–3241 (2007).

[163] Sage, R. F., Way, D. A., and Kubien, D. S. Rubisco, Rubisco activase, and global climate change. *J Exp Bot* **59**(7), 1581–1595 (2008).

[164] Rokka, A., Zhang, L., and Aro, E. M. Rubisco activase: an enzyme with a temperature-dependent dual function? *Plant J* **25**(4), 463–471 (2001).

[165] Salvucci, M. E. Association of Rubisco activase with chaperonin-60beta: a possible mechanism for protecting photosynthesis during heat stress. *J Exp Bot* **59**(7), 1923–1933 (2008).

[166] Chen, J., Wang, P., Mi, H.-L., Chen, G.-Y., and Xu, D.-Q. Reversible association of ribulose-1, 5-bisphosphate carboxylase/oxygenase activase with the thylakoid membrane depends upon the ATP level and pH in rice without heat stress. *J Exp Bot* **61**(11), 2939–2950 (2010).

[167] Shan, X., Wang, J., Chua, L., Jiang, D., Peng, W., and Xie, D. The role of Arabidopsis Rubisco activase in jasmonate-induced leaf senescence. *Plant Physiol* **155**(2), 751–764 (2011).

[168] Li, L. A., Gibson, J. L., and Tabita, F. R. The Rubisco activase (rca) gene is located downstream from rbcS in Anabaena sp. strain CA and is detected in other Anabaena/Nostoc strains. *Plant Mol Biol* **21**(5), 753–764 (1993).

[169] Li, L. A. and Tabita, F. R. Transcription control of ribulose bisphosphate carboxylase/oxygenase activase and adjacent genes in Anabaena species. *J Bacteriol* **176**(21), 6697–6706 (1994).

[170] Li, L. A., Zianni, M. R., and Tabita, F. R. Inactivation of the monocistronic rca gene in Anabaena variabilis suggests a physiological ribulose bisphosphate carboxylase/oxygenase activase-like function in heterocystous cyanobacteria. *Plant Mol Biol* **40**(3), 467–478 (1999).

[171] Portis, A. R. Rubisco activase - Rubisco's catalytic chaperone. *Photosynth Res* **75**(1), 11–27 (2003).

[172] van de Loo, F. J. and Salvucci, M. E. Activation of ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco) involves Rubisco activase Trp16. *Biochemistry* **35**(25), 8143–8148 (1996).

[173] Esau, B. D., Snyder, G. W., and Portis, A. R. Differential effects of N- and C-terminal deletions on the two activities of rubisco activase. *Arch Biochem Biophys* **326**(1), 100–105 (1996).

[174] Salvucci, M. E., Rajagopalan, K., Sievert, G., Haley, B. E., and Watt, D. S. Photoaffinity labeling of ribulose-1,5-bisphosphate carboxylase/oxygenase activase with ATP gamma-benzophenone. Identification of the ATP gamma-phosphate binding domain. *J Biol Chem* **268**(19), 14239–14244 (1993).

[175] Salvucci, M. E. and Klein, R. R. Site-directed mutagenesis of a reactive lysyl residue (Lys-247) of Rubisco activase. *Arch Biochem Biophys* **314**(1), 178–185 (1994).

[176] Shen, J. B. and Ogren, W. L. Alteration of Spinach Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Activase Activities by Site-Directed Mutagenesis. *Plant Physiol* **99**(3), 1201–1207 (1992).

[177] Shen, J. B., Orozco, E. M., and Ogren, W. L. Expression of the two isoforms of spinach ribulose 1,5-bisphosphate carboxylase activase and essentiality of the conserved lysine in the consensus nucleotide-binding domain. *J Biol Chem* **266**(14), 8963–8968 (1991).

[178] van de Loo, F. J. and Salvucci, M. E. Involvement of two aspartate residues of Rubisco activase in coordination of the ATP gamma-phosphate and subunit cooperativity. *Biochemistry* **37**(13), 4621–4625 (1998).

[179] Li, C., Wang, D., and Portis, A. R. Identification of critical arginine residues in the functioning of Rubisco activase. *Arch Biochem Biophys* **450**(2), 176–182 (2006).

[180] Kallis, R. P., Ewy, R. G., and Portis, A. R. Alteration of the adenine nucleotide response and increased Rubisco activation activity of Arabidopsis rubisco activase by site-directed mutagenesis. *Plant Physiol* **123**(3), 1077–1086 (2000).

[181] Wang, Z.-Y., Snyder, G. W., Esau, B. D., Portis, A. R., and Ogren, W. L. Species-Dependent Variation in the Interaction of Substrate-Bound Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco) and Rubisco Activase. *Plant Physiol* **100**(4), 1858–1862 (1992).

[182] Larson, E. M., O'Brien, C. M., Zhu, G., Spreitzer, R. J., and Portis, A. R. Specificity for activase is changed by a Pro-89 to Arg substitution in the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *J Biol Chem* **272**(27), 17033–17037 (1997).

[183] Ott, C. M., Smith, B. D., Portis, A. R., and Spreitzer, R. J. Activase region on chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase. Nonconservative substitution in the large subunit alters species specificity of protein interaction. *J Biol Chem* **275**(34), 26241–26244 (2000).

[184] Li, C., Salvucci, M. E., and Portis, A. R. Two residues of rubisco activase involved in recognition of the Rubisco substrate. *J Biol Chem* **280**(26), 24864–24869 (2005).

[185] Werneke, J. M., Zielinski, R. E., and Ogren, W. L. Structure and expression of spinach leaf cDNA encoding ribulosebisphosphate carboxylase/oxygenase activase. *Proc Natl Acad Sci U S A* **85**(3), 787–791 (1988).

[186] Werneke, J. M., Chatfield, J. M., and Ogren, W. L. Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and Arabidopsis. *Plant Cell* **1**(8), 815–825 (1989).

[187] Brooks, A., Portis, A. R., and Sharkey, T. D. Effects of Irradiance and Methyl Viologen Treatment on ATP, ADP, and Activation of Ribulose Bisphosphate Carboxylase in Spinach Leaves. *Plant Physiol* **88**(3), 850–853 (1988).

[188] Zhang, N. and Portis, A. R. Mechanism of light regulation of Rubisco: a specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f. *Proc Natl Acad Sci U S A* **96**(16), 9438–9443 (1999).

[189] Zhang, N., Schrmann, P., and Portis, A. R. Characterization of the regulatory function of the 46-kDa isoform of Rubisco activase from Arabidopsis. *Photosynth Res* **68**(1), 29–37 (2001).

[190] Zhang, N., Kallis, R. P., Ewy, R. G., and Portis, A. R. Light modulation of Rubisco in Arabidopsis requires a capacity for redox regulation of the larger Rubisco activase isoform. *Proc Natl Acad Sci U S A* **99**(5), 3330–3334 (2002).

[191] Wang, D. and Portis, A. R. Increased sensitivity of oxidized large isoform of ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) activase to ADP inhibition is due to an interaction between its carboxyl extension and nucleotide-binding pocket. *J Biol Chem* **281**(35), 25241–25249 (2006).

[192] Rundle, S. J. and Zielinski, R. E. Alterations in barley ribulose-1,5-bisphosphate carboxylase/oxygenase activase gene expression during development and in response to illumination. *J Biol Chem* **266**(22), 14802–14807 (1991).

[193] Portis, A. R. Rubisco activase. *Biochim Biophys Acta* **1015**(1), 15–28 (1990).

[194] Salvucci, M. E. Subunit interactions of Rubisco activase: polyethylene glycol promotes self-association, stimulates ATPase and activation activities, and enhances interactions with Rubisco. *Arch Biochem Biophys* **298**(2), 688–696 (1992).

[195] Wang, Z. Y., Ramage, R. T., and Portis, A. R. Mg2+ and ATP or adenosine 5'-[gamma-thio]-triphosphate (ATP gamma S) enhances intrinsic fluorescence and induces aggregation which increases the activity of spinach Rubisco activase. *Biochim Biophys Acta* **1202**(1), 47–55 (1993).

[196] Barta, C., Dunkle, A. M., Wachter, R. M., and Salvucci, M. E. Structural changes associated with the acute thermal instability of Rubisco activase. *Arch Biochem Biophys* **499**(1-2), 17–25 (2010).

[197] Blayney, M. J., Whitney, S. M., and Beck, J. L. NanoESI Mass Spectrometry of Rubisco and Rubisco Activase Structures and Their Interactions with Nucleotides and Sugar Phosphates. *J Am Soc Mass Spectrom* **22**(9), 1588–1601 (2011).

[198] Salvucci, M. E. Potential for interactions between the carboxy- and amino-termini of Rubisco activase subunits. *FEBS Lett* **560**(1-3), 205–209 (2004).

[199] Buechen-Osmond C, Portis AR, A. T. Rubisco Activase modifies the appearance of Rubisco in the electron microscope. In *Murata N (ed) Research in Photosynthesis, IXth International Congress on Photosynthesis*, (1992).

[200] Bakkouri, M. E., Gutsche, I., Kanjee, U., Zhao, B., Yu, M., Goret, G., Schoehn, G., Burmeister, W. P., and Houry, W. A. Structure of RavA MoxR AAA+ protein reveals the design principles of a molecular cage modulating the inducible lysine decarboxylase activity. *Proc Natl Acad Sci U S A* **107**(52), 22499–22504 (2010).

[201] Henderson, J. N., Kuriata, A. M., Fromme, R., Salvucci, M. E., and Wachter, R. M. Atomic resolution x-ray structure of the substrate recognition domain of higher plant ribulose-bisphosphate carboxylase/oxygenase (Rubisco) activase. *J Biol Chem* **286**(41), 35683–35688 (2011).

[202] Baker, R. T., Catanzariti, A.-M., Karunasekara, Y., Soboleva, T. A., Sharwood, R., Whitney, S., and Board, P. G. Using deubiquitylating enzymes as research tools. *Methods Enzymol* **398**, 540–554 (2005).

[203] Sambrook, J., Fritsch, E., and Maniatis, T. *Molecular cloning - a laboratory manual*. Cold Spring Harbour laboratory press, 2nd edition, (1989).

[204] Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**(13), 3497–3500 (2003).

[205] Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**(22), 10881–10890 (1988).

[206] Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**(4), 305–308 (1999).

[207] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. *The Proteomics Protocols Handbook, Humana Press (2005).*, chapter Protein Identification and Analysis Tools on the ExPASy Server, 571–607. Humana Press (2005).

[208] Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259), 680–685 (1970).

[209] Towbin, H., Staehelin, T., and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**(9), 4350–4354 (1979).

[210] Servaites, J. C. Crystalline ribulose bisphosphate carboxylase/oxygenase of high integrity and catalytic activity from Nicotiana tabacum. *Arch Biochem Biophys* **238**(1), 154–160 (1985).

[211] Evans, G. and Pettifer, R. F. CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J Appl Cryst* **34**, 82–86 (2001).

[212] Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACBM Newsletter* **26** (1992).

[213] Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

[214] computational project Number 4, C. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).

[215] Evans, P. R. SCALA. *Joint CCP4 and ESF-EACBM Newsletter* **33**, 22–24 (1997).

[216] French, G. and Wilson, K. On the treatment of negative intensity observations. *Acta Crystallogr. A* **34**, 517–525 (1978).

[217] Schneider, T. R. and Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crysallogr. D Biol. Crystallogr.* **58**, 1772–1779 (2002).

[218] de la Fortelle, E. and Bricogne, G. Maximum likelihood heavy atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).

[219] Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972 (2000).

[220] Emsley, P. and Cowtan, K. Coot: model-building tools for molecular graphics. *Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132. (2004).

[221] Murshudov, G. N., Vagin, A. A., and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).

[222] Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26**, 283–291 (1993).

[223] Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.1r2pre. (2010).

[224] Kreuzer, K. N. and Jongeneel, C. V. Escherichia coli phage T4 topoisomerase. *Methods Enzymol* **100**, 144–160 (1983).

[225] Barta, C., Carmo-Silva, A. E., and Salvucci, M. E. Rubisco activase activity assays. *Methods Mol Biol* **684**, 375–382 (2011).

[226] Djuranovic, S., Hartmann, M. D., Habeck, M., Ursinus, A., Zwickl, P., Martin, J., Lupas, A. N., and Zeth, K. Structure and Activity of the N-Terminal Substrate Recognition Domains in Proteasomal ATPases. *Mol Cell* **34**, 580–590 (2009).

[227] Scott, A., Gaspar, J., Stuchell-Brereton, M. D., Alam, S. L., Skalicky, J. J., and Sundquist, W. I. Structure and ESCRT-III protein interactions of the MIT domain of human VPS4A. *Proc Natl Acad Sci U S A* **102**(39), 13813–13818 (2005).

[228] Portis, A. R., Li, C., Wang, D., and Salvucci, M. E. Regulation of Rubisco activase and its interaction with Rubisco. *J Exp Bot* **59**(7), 1597–1604 (2008).

[229] Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* **18**(6), 756–764 (2008).

[230] Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**(9), 1338–1347 (1998).

[231] Werbeck, N. D., Schlee, S., and Reinstein, J. Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. *J Mol Biol* **378**(1), 178–190 (2008).

[232] Franzmann, T. M., Czekalla, A., and Walter, S. G. Regulatory circuits of the AAA+ disaggregase Hsp104. *J Biol Chem* **286**(20), 17992–18001 (2011).

[233] Mueller-Cajar, O., Stotz, M., Wendler, P., Hartl, F. U., Bracher, A., and Hayer-Hartl, M. Structure and function of the AAA+ protein CbbX, a red-type Rubisco activase. *Nature* **479**(7372), 194–199 (2011).

7 Appendices

Lebenslauf

Mathias Michael Stotz

geboren am 28.12.1981 in Tübingen, Deutschland

Juni 2001	Allgemeine Hochschulreife
	Wildermuth Gymnasium Tübingen
Oktober 2002- Februar 2008	Studium der Biochemie
	Eberhard-Karls-Universität Tübingen
	Abschluss Diplom-Biochemiker
August 2005- August 2006	Auslandsstudium
	University of Massachusetts Amherst, USA
Juli 2008 - 2012	Experimentelle Doktorarbeit
	in der Abteilung Zelluläre Biochemie von Prof. Dr. F. U. Hartl
	am Max-Planck-Institut für Biochemie in Martinsried
	"Structure of green type Rubisco activase from Nicotiana tabacum"

Publikationen

Mol Cell. 2007 Apr 13;26(1):27-39.

Swain JF, Dinler G, Sivendran R, Montgomery DL, Stotz M, Gierasch LM

Hsp70 chaperone ligands control domain association

via an allosteric mechanism mediated by the interdomain linker.

Nature 2011 Nov 2;479:194-199.

Mueller-Cajar O, Stotz M, Wendler P, Hartl FU, Bracher A, Hayer-Hartl M

Structure and function of the AAA+ CbbX a red-type Rubisco activase

NSMB 2011 Nov 6;18:1366-1370

Stotz M, Mueller-Cajar O, Ciniawsky S, Wendler P,

Hartl FU, Bracher A, Hayer-Hartl M

Structure of green-type Rubisco activase from tobacco

München, den 20.04.2012