Self-Assembly Mechanism Of Rubisco Activase

by

Agnieszka Kuriata

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2014 by the
Graduate Supervisory Committee:

Rebekka Wachter, Chair
Giovanna Ghirlanda
Kevin Redding

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

The AAA+ ATPase Rubisco activase (Rca) regulates the activity of Rubisco, the photosynthetic enzyme responsible for catalyzing biological carbon fixation. However, the detailed mechanism by which Rca self-association controls Rubisco reactivation activity remains poorly understood. In this work, we are using fluorescence correlation spectroscopy (FCS) to better characterize the thermodynamics of the assembly process of cotton Rca.

We present FCS data for Rca in the presence of Mg•ATPgS and Mg•ADP and for the D173N Walker B motif mutant in the presence of Mg•ATP. Our data are consistent with promotion and stabilization of hexamers by Mg•ATPgS and Mg•ATP, whereas Mg•ADP facilitates continuous assembly.

We find that in the presence of Mg·ADP, Rca self-associates in a step-wise fashion to form oligomeric and higher order forms, with a strong size dependence on subunit concentration. The monomer is the dominant species below 0.5 micromolar, whereas the hexamer appears to be most populated in the 10-30 micromolar range. Large assemblies containing on the order of 24 subunits become dominant above 40 micromolar, with continued assembly at even higher concentrations. Our data are consistent with a highly dynamic exchange of subunits among oligomeric species of diverse sizes. The most likely ADP-mediated assembly mechanism seems to involve the formation of spiral supra-molecular structures that grow along the helical axis by the step-wise addition of dimeric units.

To examine the effect of Mg·ATP on oligomerization, we have generated the D173N mutant of Rca, which binds but does not hydrolyze ATP. In range of 8 and 70

micromolar, 60-80% of Rca is predicted to form hexamers in the presence of Mg•ATP compared to just 30-40% with Mg•ADP. We see a clear trend at which hexamerization occurs at high ATP:ADP ratios and in addition, at increasing concentrations of free magnesium ions to 5 milimolar that results in formation of six subunits. We present an assembly model where Mg•ATP promotes and stabilizes hexamerization at low micromolar Rca concentrations relative to Mg•ADP, and suggest that this results from closed ring hexamer formation in Mg•ATP and open hexameric spiral formation in Mg•ADP.

ACKNOWLEDGMENTS

The results of the research reported in this dissertation has been achieved through several collaborative efforts. I would like to thank my supervisor Dr. Rebekka Wachter for giving me the opportunity to join her research group and all the support through all the five years. The Fluorescence Correlation Spectroscopy (FCS) at Biodesign Institute at Arizona State University (ASU) under the supervision of Dr. Marcia Levitus (her student Manas Chakraborty) has been a great resource to study oligomerization of protein and to successfully publish articles in high impact journals. Additionally, use of Analytical Ultracentrifugation (AUC) instrument at University of Arizona (UofA) was provided by Dr. Chad Park, who thought me the finer points of AUC data processing. Finally, the principles of x-ray crystallography study of Rca protein was made possible through efforts with Dr. J. Nathan Henderson at ASU. Thank you all for sharing your knowledge, help, and resources with me.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Rubisco in photosynthesis

Global climate models suggest that elevated greenhouse gas concentrations are linked to rising average temperatures [1]. By the end of this century, temperate latitudes are predicted to experience temperatures higher than ever before, whereas tropical temperatures are projected to increase even more rapidly [2]. The identification of heat-related limitations on photosynthetic carbon assimilation is critical to our ability to forecast biomass production and crop yield at higher temperatures [3]. The photosynthetic apparatus and associated metabolic pathways may be subjected to temperatures at which inhibition of photosynthesis remains reversible or irreversible due to damage or denaturation of molecules by exceeded thermal tolerance limits.

In the process of photosynthesis, plants convert light into chemical energy. The energy produced by photosynthesis is then used to synthesize sugars and other foodstuffs. Photosynthesis is one of the most heat-sensitive processes in plants: heat stress inhibits photosynthesis, reducing the overall yield of the plant [4]. This sensitivity can have gigantic consequences for agricultural production, predominantly in grain and oilseed crops, which need a continuous source of recently fixed carbon to support the growing harvest.

In photosynthesis, the crucial element is RuBisCO (ribulose-1, 5-bisphosphate carboxylase/oxygenase), an enzyme that catalyzes the first major step of carbon fixation in order to create sucrose and similar molecules [5]. Many research groups have demonstrated that CO_2-fixing enzyme Rubisco represents a primary limiting factor of

photosynthesis, and this had generated significant interest in studying the enzyme. As the most abundant enzyme on Earth, substantial research effort has gone into its characterization and study.

Rubisco: active and inactive

For Rubisco to be catalytically active, it must undergo carbamylation of Lys201 by CO_2 .Its active site should not be occupied by any of the inhibitory sugars, but instead by Rubisco's substrate, ribulose-1,5-bisphosphate (RuBP) [6]. Rubisco can also be found in its catalytically inactive state, for example: when uncarbamylated active sites form dead-end complexes due to the very tight binding of RuBP [7]. As another example, carbamylated sites can form dead-end complexes with other sugar phosphates that are developed by misprotonation of an intermediate[8]. Discovery of protein rubisco activase in the early 1980s was an important event that helped researchers understand how Rubisco is able to be reactivated by the removal of inhibitory sugar phosphates from the active sites.

Structural classification of Rubisco activase

Studies from 1982 of the *rca* mutant of Arabidopsis revealed that a previously unknown protein was needed for achieving and maintaining high catalytic activity of Rubisco in plants [9]. The rubisco activase was named for its ability to restore the activity of Rubisco.

Unfortunately, rubisco activase is a very thermolabile enzyme and inhibition of activase leads to a decrease in the activation state of Rubisco [10] [11] [5]. Rubisco activase has a

relatively low optimum temperature, approximately 20-35°, which is the same as the optimum temperature for net photosynthesis [12] [5]. Reconstitution experiments have shown that activation of Rubisco by Rubisco activase decreases with increasing temperature, even when the pH, reductant, and Mg^{2+} ions are constant and an ATP regenerating system is used to maintain a high ratio of ATP to ADP [11, 12].

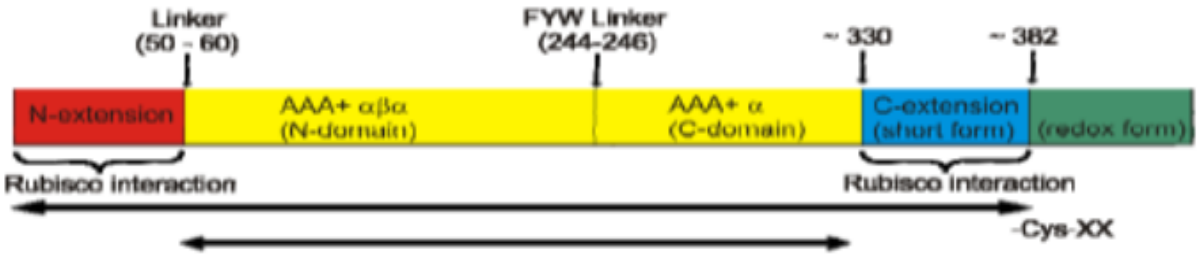


Figure 1. Domain organization of Rubisco activase non-redox (short β) and redox (long α) isoforms.

Activase is a member of the AAA^+ family, comprised of ATPases associated with diverse cellular activities [13]. It is a mechano-chemical motor protein that utilizes the energy released by ATP to carry out mechanical functions such as the disaggregation, unfolding, and structural remodeling of macromolecules [14]. Common features of these proteins are that they contain one or more copies of the AAA^+ motif, and they typically form a ring structure. The 200-250 residue AAA^+ fold contains two domains, an amino terminal domain (consisting of an alpha/beta fold and nucleotide binding site), and a carboxy terminal domain (consisting of alpha helices) (Fig. 1). Key sequence motifs in α/β domain are the P-loop Walker A (Lys and Thr/Ser residues bind the γ and β phosphate of ATP, and Mg^{2+}), and Walker B (conserved Glu, Asp in activase, likely serves as the catalytic base that coordinates the γ and β phosphates and play a role in

metal-ion coordination); the sensor-1, (conserved asparagine, Asn226 in activase) interacting with the γ-phosphate of ATP through a bound water. A conserved arginine (the Arg finger) in Box VII is essential in the nucleotide hydrolysis located in the adjacent subunit of the ring. The α domain consists of a Sensor 2 motif, which bears another arginine residue (Arg292 in activase) usually located on the third helix of the α-domain. This residue also interacts with the γ-phosphate of ATP, and is thought to serve as a mediator of movements of the N- and C-domains with respect to each other upon ATP hydrolysis [15]. The Sensor 2 motif in the C-domain is a major determinant of Rubisco specificity [16]. The degree of sequence conservation of the C-domain is considerably lower than that of the N-domain.

Rubisco activase isoforms

The main regulation of Rca in the chloroplast stroma has been proposed to involve the ambient ATP/ADP ratio. In most plants, two isoforms of Rubisco activase exist that usually arise from alternative splicing events, the α (46 kD) form and the β (42 kD) form. The β form is not sensitive to environmental redox conditions [17], whereas the α form is highly redox-sensitive due to a C-terminal extension of approximately 36 residues with two conserved cysteine residues, as well as several charged residues with regulatory roles [18].

Published X-ray and EM structures of bacterial and higher plant Rca.

Available crystal structures of AAA^+ proteins can give us an idea of the possible shapes of activase. A predicted model of *Arabidopsis* was made using a MODELLER program

[19] where all the motifs could be seen. Based on available x-ray structures and modeling of ring assemblies from monomer structures, it is now thought that AAA^+ modules in the classic clade form hexamers in physiologically relevant states [10].

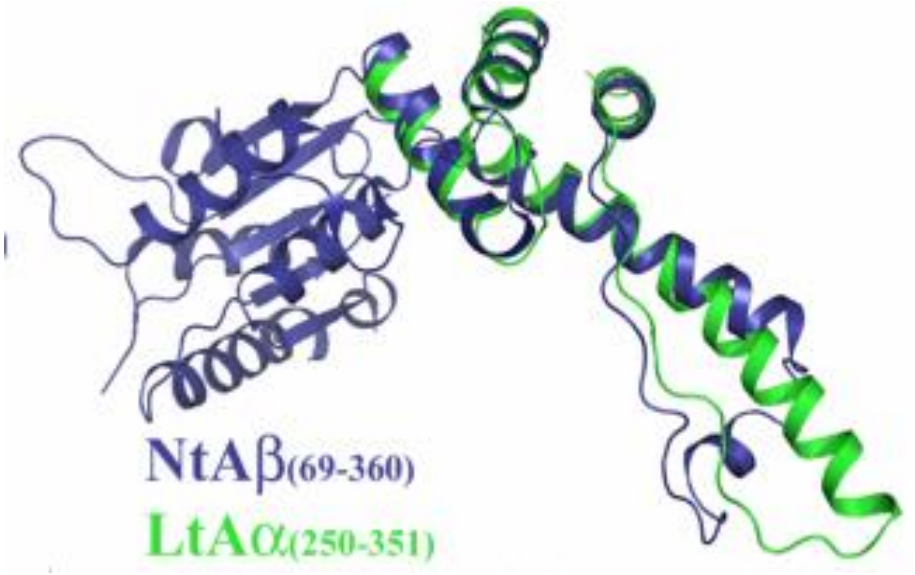


Figure 2. Overlay of the x-ray crystal structures of tobacco AAA^+ domain (blue) and creosote recognition domain (green).

The first crystallographic information on higher plant Rca was published in the year 2011, the 1.88 Å x-ray structure of a 97-residue C-terminal creosote fragment [20], and the 2.95 Å x- ray structure of the AAA+ domain (residues 68 - 360 out of 383 total) of tobacco Rca [21] (Fig.2). The tobacco apoprotein crystallized in a pseudo- hexameric spiral, and closed-ring hexameric models were assembled by fitting the density into negative stain electron microscopy (EM) maps. The 3.0 A structure of the bacterial Cbbx protein was reported, a red-type Rca from the proteobacterium *Rhodobacter sphaeroides*. Cbbx functions as an ATPase only in the presence of RuBP and Rubisco [22].

Although much is known about Rubisco and intensive studies have been performed on Rca, the mechanism of interaction between these two enzymes is still undefined.

Self-assembly studies on Rubisco activase.

In plants, the concentration of the Rubisco holoenzyme is about 500 μM, the monomer Rca concentration is about 170-500 μM, and the kinetically estimated dissociation constants for this complex range from 0.34 to 5 μM [23] [24]. To understand the interaction of Rubisco with Rca, it is crucial to study the self-assembly pathway of activase. Currently, there are plenty of techniques that are used for oligomerization studies. However, the work might be challenging if there are low protein yields, a high likelihood of protein aggregation, a high percentage of polydispersity, or simply the limitation of the instrument. Therefore, experiments and methods should be chosen carefully, and only those that provide researcher with high quality results should be performed.

The studied size of activase varies, depending on the method used or the conditions that were used for preparation or measurement. For example, large oligomerization (over 550,000 Da) was observed when using gel filtration chromatography at room temperature in the presence of ATP or ATP-γ-S and around 340,000 Da by adding ADP [25]. The distribution of the species is highly protein concentration dependent.

Another method, dynamic light scattering (DLS), measures the size of a particle undergoing Brownian motion. The scattered light of different particles in solution shifts and interferes, causing a fluctuation in the intensities of the scattered light. This is calculated by a mathematical algorithm called an autocorrelation function that can be related to the diffusion constant (D). Using the Stokes-Einstein relation, the

hydrodynamic radius (R_h) of the particle can be extracted and therefore the oligomeric state of the particle can be determined [26].

Another useful method is small angle x-ray Scattering (SAXS). This method can be used to provide information about the structure of a sample, such as particle size, shape, distribution, and surface-to-volume ratio. In agreement to the results mentioned above, the SAXS data depicted an average of multiple Rca species found in solution, mostly an equilibrium of monomer-hexamer complexes [27].

Structural information on Rubisco

Currently, a large number of Rubsico x-ray structures are available for Spinach, Tobacco and *Chlamydomonas*. The most common form is the Type I holoenzyme with 8 large and 8 small subunits (L8S8), where four dimers of the large subunits are clustered around a four-fold axis, forming a cylinder bearing two rings of four active sites each [28]. The size of the large subunit varies between 52,000-55,000 Da, whereas the small subunit's size is 13,000 Da [29].

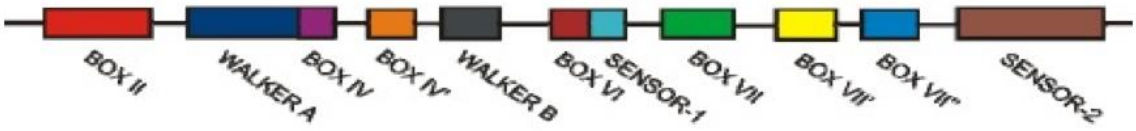


Figure 3. The AAA+ module of activase based on sequence alignment with the HslU protease.

The top and bottom parts of this 2 × L4 arrangement are capped by four small subunits each (2 × S4).

Plausible interaction models between Rubisco and activase

Protein-protein interactions, where the proteins bind together to carry out some biological function, are critical for function. It is still not known how Rubisco interacts with Rubisco activase, however there are some models presented. One of the more interesting and plausible models presents an activase ring stacked onto Rubisco L8S8, forming an extended cylinder-like shape. Similar interactions have been found in different ring-like AAA^+ structures and proteases. For example, the hexameric HslU ATPase, which binds to the dodecameric HslV peptidase, then forms an ATP-dependent HslVU protease [30]. Another possible geometry for Rca binding with Rubisco is a side-on interaction in which the recognition domains of both enzymes would be in a closer proximity (Fig. 4).

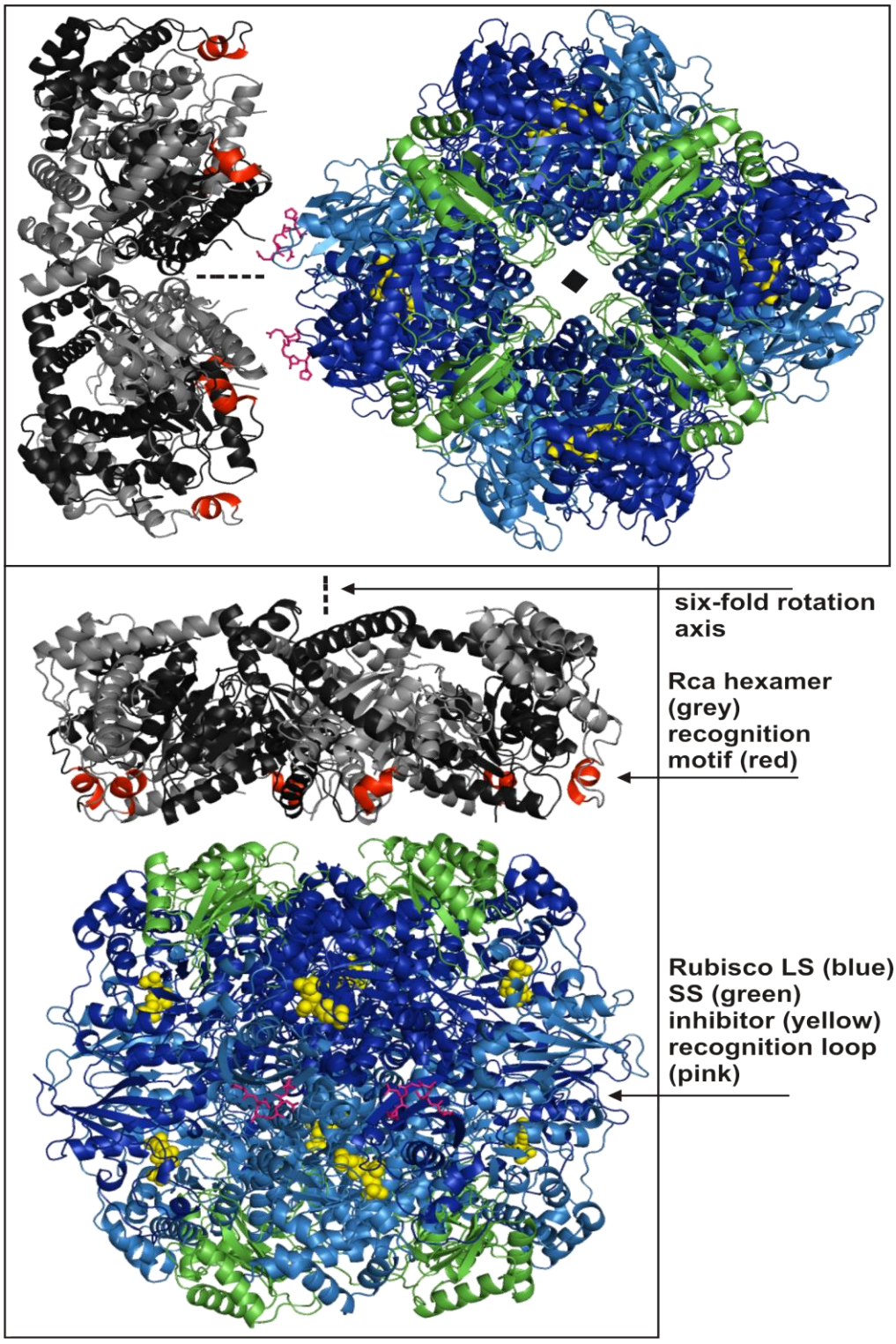


Figure 4. Plausible model for Rca and Rubisco interaction.

Protein oligomerization

It has already been shown that activase may form mostly hexameric or higher order species. Some studies have shown that dimers are active assemblies [27], therefore it is important to fully understand the oligomerization of this enzyme at most physiological conditions.

The average oligomeric state of cellular proteins is tetrameric [31], and 35% or more of proteins in a cell are oligomeric [32].

Before getting to the point of Rca assembly, it is important to first understand why the oligomerization is vital to study. Subunit association can differ in strength and time. When proteins are found only, or mostly, in the oligomeric state, they generally have dissociation constants in the nanomolar range. Proteins with higher K_D values in the micromolar or even millimolar range usually tend to have a weak tendency to associate, which is concentration, temperature, or pH dependent [33].

There are a several advantages to proteins forming assembly complexes. Oligomeric proteins may function more efficiently since they contain more active sites [34] and an additional level of control is present when proteins are allosterically regulated. If a large protein is composed of multiple, short subunits, it is less prone to errors during a transcription than a single, large subunit [35]. Another important advantage is the resistance to degradation and denaturation, seen in thermophilic organisms where the proteins tend to increase in oligomerization [36].

Nucleotide effect on activase oligomerization

The AAA^+ superfamily of proteins uses energy from ATP to drive macromolecular rearrangements. Here, Rca reactivates inhibited Rubisco. This group of proteins has a unique ability to form oligomeric rings. For instance: ClpX [37],HslU [38] ClpY, and Hsp104 [39] form stable hexamers in a presence of ATP. ClpA [40] [41] [42] was found as a monomer and dimer in the absence of nucleotides and formed single hexameric rings in the presence of ATP or its non-hydrolyzable analogs.

Many ATPases (ClpB) [43] undergo reversible nucleotide-dependent self-association, and this is one of the most intriguing characteristics of Rubisco activase.

Summary

Rubisco activase is definitely not an easy protein to work with: it is very thermolabile, prone to aggregation in the absence of nucleotide in solution, and it is very polydisperse.

It took almost 30 years to solve the first x-ray crystal structure of Rca, and another large effort has been undertaken trying to study the self-assembly and the mechanism of Rubisco and Rca interaction.

Without a doubt, Rca belongs to the group of proteins that forms oligomers and undergoes reversible nucleotide-dependent self-association. This dissertation contains the studies that were done on Rca, specifically the nucleotide and magnesium effects on the oligomerization.

CHAPTER 2

PROTEIN OLIGOMERIZATION MONITORED BY FLUORESCENCE FLUCTUATION SPECTROSCOPY: SELF-ASSEMBLY OF RUBISCO ACTIVASE[1]

[1] Accepted for publication as Chakraborty, M., Kuriata, A. M., Henderson, J., N., Salvucci, M. E., Wachter, R. M., Levitus, M. (2012). Protein Oligomerization Monitored by Fluorescence Fluctuation Spectroscopy: Self-Assembly of Rubisco Activase. Biophysical Journal, 103 949–958.

ABSTRACT

A methodology is presented to characterize complex protein assembly pathways by fluorescence correlation spectroscopy. We have derived the total autocorrelation function describing the behavior of mixtures of labeled and unlabeled protein under equilibrium conditions. Our modeling approach allows us to quantitatively consider the relevance of any proposed intermediate form, and K_d values can be estimated even when several oligomeric species coexist. We have tested this method on the AAA+ ATPase Rubisco activase (Rca). Rca self-association regulates the CO_2 fixing activity of the enzyme Rubisco, directly affecting biomass accumulation in higher plants. However, the elucidation of its assembly pathway has remained challenging, precluding a detailed mechanistic investigation. Here, we present the first thermodynamic characterization of oligomeric states of cotton beta-Rca complexed with Mg·ADP. We find that the monomer is the dominating species below 0.5 micromolar. The most plausible model supports dissociation constants of about 4, 1 and 1 micromolar for the monomer-dimer, dimer-tetramer and tetramer-hexamer equilibria, in line with the coexistence of four different oligomeric forms under typical assay conditions. Large aggregates become dominant above 40 micromolar, with continued assembly at even higher concentrations.

We propose that under some conditions, ADP-bound Rca self-associates by forming spiral arrangements that grow along the helical axis. Other models such as the stacking of closed hexameric rings are also discussed.

INTRODUCTION

Knowledge of the oligomeric state of a protein is essential for understanding its function. It has been estimated that more than 80% of proteins in *E. coli* are composed of two or more subunits, and among these, about 80% are homooligomers [32, 44]. Oligomerization is believed to be evolutionarily advantageous, as it provides opportunities for error control and regulation [32, 45]. However, the oligomerization state of a protein is often difficult to determine. Although several *in vitro* methods such as size exclusion chromatography, analytical ultracentrifugation and chemical cross-linking can be employed these techniques are not always adequate to identify transient intermediates.

For some proteins, the oligomerization state can be modulated by environmental factors such as concentration, temperature, pH, phosphorylation, or ligand binding [44, 45]. Nucleotide binding partakes in the assembly of Rubisco activase (Rca), a chemo-mechanical motor protein that plays an essential role in regulating Rubisco (ribulose 1,5-bisphosphate caboxylase/oxygenase) activity in higher plants [17, 46-48]. The enzyme Rubisco catalyzes the incorporation of CO_2 into simple carbohydrates [5], however, its activity is diminished by the generation of competitive inhibitors that produce dead-end complexes [49, 50]. In many photosynthetic organisms, the ring-forming ATPase Rca catalyzes conformational changes to reactivate Rubisco for carbon fixation [48]. The

activity of Rca is up-regulated by high ATP and down-regulated by high ADP levels, and moderate heat stress causes activity loss [46, 51-53].

Rca (~383 residues) belongs to the AAA+ superfamily, a ubiquitous group of proteins that utilizes ATP hydrolysis to carry out mechanical work on macromolecular substrates [54]. The AAA+ module consists of the N-terminal ring-forming domain bearing the Walker A and B motifs, and a less conserved C-terminal domain positioned around the periphery of toroidal assemblies [54]. Adenine nucleotides bind at the domain interfaces, with hydrolysis thought to be coupled to large-scale C-domain motions. Recently, the first X-ray models of some fragments of higher plant Rca have become available, the 1.9 Å structure of the creosote C-domain [20] and the 2.9 Å structure of the tobacco AAA+ module comprising residues 68 - 360 [21]. However, the Rca-specific N- and C-terminal regions flanking the AAA+ domain remain structurally uncharacterized. The nucleotide-free tobacco protein crystallized in a pseudo-hexameric spiral, and closed-ring hexameric models were generated by fitting the coordinates into negative stain electron microscopy (EM) maps (Fig. 7 in the Supporting Material of this chapter) [21]. Notably, the 3.0 Å crystal structure of a bacterial Rca was recently reported in combination with EM images that depict symmetric hexameric rings in the presence of ATP and ribulose-bisphosphate (RuBP) [22].

Rca is a member of the extended classic AAA+ clade known for the formation of closed-ring hexamers [54]. Some members of this clade require nucleotides and macromolecular binding partners to generate functional oligomers [55], whereas others require nucleotides to control their interactions with partner proteins [56]. Several families consist of tandem arrays of AAA+ modules that stack on top of each other [57]. In this

group, a well-studied example is ClpA, where hexamerization is ATP- and Mg^{2+}-dependent [42]. However, for Rca, the assembly pathway has been difficult to determine due to the high degree of size polydispersity observed in all protein preparations. Based on size-exclusion HPLC (SE-HPLC), molecular weight estimates have ranged from 58 to well over 550 kDa, with strong dependence on protein concentration and other assay conditions (ATP, ATP-γ-S, Mg^{2+}, ADP, apo-form, polyethylene glycol, ionic strength, glycerol, temperature) [25, 47, 52]. Typically, SE-HPLC elution profiles consist of broad, asymmetric bands indicative of a large number of co-eluting oligomeric species, with nucleotides facilitating self-association [47]. Although an inter-molecular salt bridge appears to be disrupted in the tobacco Rca-R294A and R294V variants [58], ATP-dependent hexamers have been identified by mass spectrometry and EM [21, 59]. The allosteric regulation of Rca is of intense current interest, as subunit assembly modulates the extent of biological carbon fixation. However, due to the difficulties inherent in characterizing coexisting oligomeric species, the regulation of Rca activity remains poorly understood. To elucidate bioenergetic parameters of Rubisco reactivation, a complete thermodynamic understanding of the step-wise assembly pathway is essential.

Here, we utilized fluorescence fluctuation methods to examine the cotton β-Rca (non-redox) assembly process in the presence of Mg•ADP (adenosine diphosphate complexed with Mg^{2+}). These techniques rely on the measurement of fluorescence intensity fluctuations in a small number of molecules contained in an optically restricted volume [60, 61]. In fluorescence correlation spectroscopy (FCS), temporal fluorescence fluctuations are analyzed by means of the autocorrelation function ($g(\tau)$). Results are then interpreted in terms of the dynamics of the different processes underlying these

fluctuations, such as Brownian motion [62, 63]. FCS has been widely used to study molecular transport *in vitro* and *in* vivo [64-68] to study chromosome dynamics [69, 70] and protein oligomerization [71, 72]. Our results show that Rca self-associates in a step-wise fashion to higher-order forms that exceed the molecular weight of a hexamer. Assembly models are discussed in relation to experimental data.

MATERIALS AND METHODS

Cloning, expression and purification of 6His-tagged and untagged cotton β-Rca

This work was carried out on cotton β-Rca-378AC, a short-form Rca with a C-terminal Ala-Cys insert [73]. A blunt-ended PCR product coding for β-Rca-378AC was amplified from a pET23a+ expression plasmid and cloned into a linear pET151/D-TOPO vector, which was transformed into One Shot TOP10 competent cells according to the manufacturer's instructions (Invitrogen). The correct insert was verified by DNA sequencing of individual transformants.

The pET151/D-TOPO vectors with correct insert were transformed into E. coli BL21*(DE3) (Invitrogen). Single colonies were cultured overnight in 25 mL LB media plus 100 μg/mL carbenicillin, and used to inoculate 1 L LB plus 100 μg/mL carbenicillin. Growth was continued at 250 rpm and 37°C until the OD_{600} reached ~ 0.6. The cultures were cooled to ~ 25°C, 100 mg of isopropyl-β-D-thio-galactoside (IPTG) was added to each flask, and growth was continued at 200 rpm and 25°C. After 8 h, the cells were harvested by centrifugation and pellets frozen at -80°C. Thawed cell paste was suspended

in 50 mL of 25 mM HEPES pH 8.0, 20 mM imidazole pH 8.0, 300 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM phenylmethanesulfonylfluoride (PMSF), 1 mM DTT, 0.5 mM ADP and 0.1% Triton X-100, then disrupted by sonication. The lysate was pelleted by centrifugation, and the supernatant was passed through a 0.8 μm syringe filter and loaded onto a Ni-NTA column (Qiagen). Rca was purified using an imidazole buffer step gradient (25 mM HEPES pH 8.0, 300 mM NaCl, 1 mM DTT and 0.5 mM ADP, plus imidazole). After washing with 50 mM and 100 mM imidazole buffer, His-tagged Rca was eluted with 25 mL of 500 mM imidazole buffer. Protein fractions were pooled, 1.2 mg of tobacco etch virus protease (TEV) was added to cleave the 6His-tag, and the sample was dialyzed overnight at 10°C against 1 L of 25 mM Tris pH 8.0, 300 mM NaCl, 0.1 mM ADP, 0.1 mM DTT and 1 mM EDTA. The dialysate was reapplied to a Ni-NTA column, and TEV-cleaved Rca was collected in the flow-through and 20 mM imidazole wash. Protein was concentrated in Centriprep concentrators (Millipore) to a volume of ~ 2.5 mL. Using a PD-10 column (GE Healthcare), the buffer was exchanged into 25 mM HEPES pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 2 mM ADP and 10% glycerol. Protein concentration was determined by the Bradford method, with typical yields of 3-4 mg/L cell culture. Recombinant cotton β-Rca-378AC was also expressed in E. coli without an affinity tag, and purified by classical procedures as described previously (39). All Rca preparations were flash-frozen and stored at -80°C.

Dye conjugation methodology

Cotton β-Rca-378AC was covalently labeled with the ALEXA 546 C_5- maleimide fluorophore (Invitrogen) using published procedures [74]. A typical reaction consisted of 236 µl of 28 µM Rca, 3 µl of 500 mM ADP, buffer stock, and 54 µl of 1.93 mM Alexa dye stock (prepared in 50 mM phosphate pH 7.2), to provide a final volume of 300 µl in 50 mM phosphate pH 7.2. Control reactions were supplemented with 50 mM dithiothreitol (DTT) prior to the addition of Alexa. All reaction tubes were incubated overnight at 4°C. To remove excess label, 300 µl saturated ammonium sulfate was added to each tube followed by incubation for 30 min at 4°C. The pellets were collected in a micro-centrifuge (12,000 rpm), suspended in 170 µl 25 mM HEPES pH 7.5, 250 mM KCl, 1 mM ADP and 10% glycerol, and desalted by two passes through a gel filtration spin column (Sephadex G50-fine, 2 ml bed volume, 400 *g* centrifugation for 2 min). 6 µl aliquots (30 – 60 µM Rca) were flash-frozen in 1X sample buffer and stored at -80°C. 1X sample buffer contained 25 mM HEPES pH 7.6, 250 mM KCl, 2 mM ADP, 5 mM $MgCl_2$ and 10% glycerol.

HPLC, spectrophotometric and mass spectrometric analysis of labeled protein

Alexa-labeled Rca samples were analyzed by reverse-phase HPLC (Waters 600 HPLC system, Waters 996 photodiode array detector) on a C18 analytical column (Vydac) using a linear water/acetonitrile gradient with 0.1% trifluoroacetic acid. Protein was monitored by O.D.280 nm, and Alexa by O.D.550 nm. All protein eluted at 53 min, whereas free ALEXA eluted at 47.5 min. The protein fraction was collected and its absorbance

spectrum determined (650 - 250 nm, UV2401-PC spectrophotometer, Schimadzu) using 50 % acetonitrile/TFA as a blank. The spectrum was utilized to calculate the molar Alexa:protein ratio. The concentration of Alexa was determined from the O.D.556 nm (ε_{556} = 104,000 $M^{-1}cm^{-1}$) and the concentration of Rca from the O.D.280 nm after correction for dye contributions. To this end, the O.D.556 was multiplied by the correction factor 0.146, and the resulting value subtracted from the O.D.280. The Rca extinction coefficient was calculated from the protein's sequence (ε_{280} = 44,350 $M^{-1}cm^{-1}$). To verify the correct molecular weight of labeled protein, MALDI spectra were collected on a Voyager DE STR mass spectrometer (see Supporting Material).

Sample preparation for fluorescence fluctuation measurements

In a typical experiment, 57 μl 2X sample buffer (50 mM HEPES pH 7.6, 500 mM KCl, 4 mM ADP, 10 mM $MgCl_2$ and 20% glycerol) and 57 μl nanopure water were added to 6 μl labeled Rca (52 μM) to provide a 2.6 μM stock of labeled protein. A 30 μM stock of unlabeled Rca was prepared in a similar manner. All solutions were kept at 22°C while performing FCS experiments (no more than 2.5 hours). For each experiment, appropriate amounts of labeled and unlabeled Rca were combined in a tube containing 1X buffer to give the desired final protein concentration. For example, to prepare a mix of 50 nM labeled and 5.0 μM unlabeled Rca, 3.8 μl of 2.6 μM labeled Rca stock were combined with 33.3 μl of 30 μM unlabeled Rca stock, and 81.5 μl each of 2X buffer and nanopure water were added (200 μl final volume). To allow for subunit equilibration, each sample

was incubated for 10 min at 22°C prior to analysis. All experiments were performed on 50 nM labeled and 0 – 125 µM unlabeled Rca.

Fluorescence Fluctuation experiments

FCS and PCH measurements were carried out using a custom made confocal microscope. A 532 nm CW laser (Compass 215m-10 Coherent GmbH, Germany) was focused ~4 µm above the cover slip using an oil immersion objective (Olympus PlanApo 100X/1.4NA Oil). Laser power was attenuated to about 50 µW using a neutral density filter. Emitted fluorescence was collected using the same objective and then passed through a 50µm pinhole to reject the out-of-focus light. The signal was detected using a silicon avalanche photodiode (SPCM-AQR-14; Perkin-Elmer, Fremont, CA). A band pass filter (Omega 3RD560-620) in front of the detector was employed to further reduce the background signal. An ALV correlator card (ALV 5000/EPP, ALV-GmbH, Germany) was used to correlate the detected fluorescence signal. Typical acquisition times were on the order of 10 min. Data for PCH analysis were acquired with 10 µs resolution using a PCI-6602 acquisition card (National Instruments, Austin,TX). Samples were measured in perfusion chambers (Grace BioLabs) pretreated with 0.1 mg/ml BSA (New England BioLabs) to minimize Rca adsorption onto the cover glass.

FCS data analysis

In FCS, fluctuations are analyzed in terms of their temporal behavior by means of the autocorrelation function, $g(\tau)$ [62, 63]

$$g(\tau) = \frac{\langle \delta I(t) \delta I(t+\tau) \rangle}{\langle I(t) \rangle^2} \qquad \text{(Eq.1)}$$

Here, I(t) is the fluorescence intensity at time t, the angular brackets denote averaging over the data accumulation time, and $\delta I(t) = I(t) - \langle I(t) \rangle$ represents fluctuations from the mean [62, 63]. For the simplest case in which a single species diffuses freely in solution and the point spread function of the instrument is adequately described by a three-dimensional Gaussian, the autocorrelation function can be expressed in terms of the diffusion coefficient of the fluorescent molecules (D) as

$$g(\tau) = \frac{2^{-3/2}}{\langle N \rangle} \left(1 + \frac{4D\tau}{r_0^2}\right)^{-1} \left(1 + \frac{4D\tau}{z_0^2}\right)^{-1/2} \qquad \text{(Eq.2)}$$

where r_0 and z_0 are the radial and axial semi-axes of the Gaussian confocal volume, and $\langle N \rangle$ is the mean number of molecules in an effective observation volume $V_{eff} = (\pi/2)^{3/2} r_0^2 z_0$ [62, 63]. The instrument was calibrated daily by measuring the FCS decays of free TAMRA dye (D = 414 $\mu m^2 s^{-1}$, [74, 75]) and/or free Alexa-546 dye (D = 341 $\mu m^2 s^{-1}$, [76]) in buffer, and fitting the experimental decays to Eq. 2 using r_0 and z_0 as fitting

parameters. The diffusion coefficient of Rca was then determined by fitting the experimental FCS decays using the values of r_0 and z_0 measured on the same day.

PCH data analysis

In PCH, the same raw data used to calculate the FCS decay are converted to a count rate probability distribution by determining the relative frequency with which 0, 1…k photons are measured in a given sampling time. This experimentally determined frequency histogram is then analyzed in terms of theoretical models of the probability distribution function that describe the expected distribution of photon counts for the system [77, 78]. In contrast to FCS, the description of the PCH of a single species diffusing in a 3D Gaussian volume cannot be expressed analytically. However, the histogram be calculated numerically, and it is uniquely characterized by two parameters: the mean number of molecules that occupy the observation volume, $\langle N \rangle$, and the molecular brightness defined as the average photon counts detected per diffusing particle per second, ε [79, 80]. ε is expected to be directly proportional to the number of labeled subunits present in the particle (monomer, dimer, etc), which is a direct measure of the oligomerization state of the protein. The value $\langle N \rangle$ was obtained from the amplitude of the FCS decay (Eq. 2), and the experimental PCH was then analyzed using ε as the only fitting parameter.

Diffusion coefficients of monomeric and oligomeric species

To interpret FCS data, we estimated the relative diffusion coefficients of different oligomers in two different ways. In the first approach, we used a value for D measured at 50 nM Rca as the diffusion coefficient of the monomer (D_1). According to the Stokes-Einstein equation, D of a spherical particle is inversely proportional to the cubic root of its volume. If all oligomers are approximately spherical with constant specific volume, $D_k = D_1/k^{1/3}$ for an oligomer with k subunits. The influence of molecular shape (i.e. non-spherical particles) was found to be minor in the interpretation of the experimental results (see Supporting Material). Alternatively, the radius of gyration R_g of different oligomeric states was calculated from the coordinates of the reconstructed closed-ring hexamer of tobacco Rca (pdb ID code 3ZW6) [21] as [77].

$$R_g^2 = \frac{1}{N}\sum_{k=0}^{n}(\vec{r} - \vec{r}_{mean})^2 \qquad \text{(Eq.3)}$$

The models for the monomer, dimer, trimer and tetramer consisted of the coordinates of one, two, three or four adjacent subunits respectively. D was calculated from Stokes-Einstein's equation as

$$D = \frac{k_B T}{6\pi\eta}\left(\left(\frac{5}{3}\right)^{3/2} R_g^3 \Big/ 0.7654 + \frac{3M_w}{4\pi N_{av}\delta_{water}}h\right)^{-1/3} \qquad \text{(Eq.4)}$$

where T is taken as 293 K, η is the solution viscosity (taken as pure water) and k_B is Boltzmann's constant. The first summand in the parenthesis takes into account that the

structural models comprise 76.54 % of the mass of the full length protein, and that the physical radius of a sphere is $(5/3)^{1/2}R_g$ [77]. The second summand represents the contribution of hydration. Here, M_w is the MW of the protein, N_{av} is Avogadro's number, δ_{water} is the density of water, and h is the mass of hydration water (typically 0.2-0.6 g water/g protein [78]). In our calculations, we arbitrarily used $h = 0.4$. This parameter has only a minimal effect on the calculated ratios of D and does not affect data interpretation (see Supporting Material).

Results

We have developed PCH and FCS methods to investigate Rca assembly at pH 7.6 in the presence of 2 mM ADP and 5 mM Mg^{2+}. To this end, we have labeled cotton β-Rca-378AC with a maleimide-functionalized Alexa 546 fluorophore by chemical modification of an engineered cysteine residue [73]. In this variant, Ala-Cys was inserted after residue 378 of the 380-residue protein [73], positioning the label 20 residues downstream of the C-terminus of the AAA+ X-ray model (Fig. 7) [21]. Compared to wild-type, the insertion mutant was reported to exhibit 81% of ATPase and 53% of Rubisco reactivation activity, although derivatization was shown to increase these activities to 150% and 131% respectively [73]. In line with previous reports [73], MALDI spectra were consistent with a single molecular mass of appropriate MW (see Supporting Material). Absorbance scans collected on HPLC-purified denatured protein provided a molar ratio of 1.04 ± 0.08 (n = 4) label attached per protein chain.

Both FCS and PCH rely on the analysis of fluorescence fluctuations under equilibrium conditions. The average number of molecules present in the observation volume, $\langle N \rangle$, is required to be small for the amplitude of intensity fluctuations to be large relative to its mean, imposing an upper limit of about 300 nM on the concentration of labeled Rca. Therefore, we performed experiments at higher Rca concentrations by mixing labeled and unlabeled protein. This strategy has been used to investigate oligomerization or aggregation in a large variety of proteins such as α-Synuclein [81], tubulin [72], barnase [82] and the tumor suppressor p53 [71], and it assumes that the labeled and unlabeled subunits equilibrate producing oligomers that bear a variable number of fluorescent subunits. Although mixing greatly complicates PCH analysis, a substantial amount of information can be extracted by FCS analysis.

Rca is a monomer at concentrations below 300 nM

Experiments on labeled Rca were performed in the 50-300 nM range. At each subunit concentration, the diffusion coefficient was calculated by fitting the experimental FCS decays to Eq. 2. At 50 nM labeled Rca, a value of $D_{50nM} = 64.7 \pm 3.7\ \mu m^2 s^{-1}$ was obtained by fitting ten independent experiments. At higher concentrations, the diffusion coefficients were obtained by averaging four independent determinations, and were found to be indistinguishable from D_{50nM} within error (Fig. 8). These results suggest that the Rca oligomerization state does not change between 50 and 300 nM. In addition, the mean number of particles obtained from the inverse of the FCS amplitude (Eq. 2) scales linearly with Rca concentration as expected for monodisperse samples (Fig. 8). Self-

association would decrease the particle number, resulting in a downward curvature of this plot.

The experimentally determined D_{50nM} can be compared to the diffusion coefficient calculated from the R_g value of a single subunit. Based on the crystallographic model [21], we estimate D_1 for an Rca monomer to be 59.1 $\mu m^2 s^{-1}$ at 20°C (see Materials and Methods). Since the experimental value is somewhat larger than the calculated one, whereas aggregation would reduce *D*, these results provide further support for monomeric Rca in the 50-300 nM range. Further evidence supporting a monomer can be obtained from PCH analysis, which is based on the molecular brightness (ε) defined as the average photon count detected per particle per second. Because each Rca monomer is labeled with a single fluorophore, ε of a dimer is expected to be twice that of a monomer. Although the PCH of a single species diffusing in a three-dimensional volume cannot be expressed analytically, the histogram can be calculated numerically in terms of the parameters $\langle N \rangle$ and ε [79]. At each Rca concentration, we determined ε from a non-linear fit of the experimental data, fixing $\langle N \rangle$ at the value obtained from the amplitude of the FCS decay (Eq. 2). To obtain the molecular brightness of individual labels, analogous experiments were performed with free dye in solution. Assuming that the fluorescence quantum yield and extinction coefficient of the dye are not affected by the protein environment, the ratio of the molecular brightness of labeled Rca (ε) and the free dye (ε_{FD}) is a measure of the oligomerization state of the protein. The ratios $\varepsilon/\varepsilon_{FD}$ were obtained by PCH analysis of experiments on labeled Rca and on free dye from 50 to 300 nM, and are plotted in Fig. 8. Over this concentration range, the ratios remain constant

within error at an average value of 0.8. The small reduction in brightness suggests that the product of the extinction coefficient and fluorescence quantum yield of the dye is diminished upon attachment to the protein. Regardless, the observed ratio is consistent with monomeric Rca, as self-association would provide a ratio greater than 1.

Rca forms oligomers in the 1-10 μM range, and large complexes at higher concentrations

To investigate Rca assembly above 300 nM concentration, we mixed labeled and unlabeled protein, and assumed that subunit equilibration will lead to a random distribution of labels over all oligomeric forms. In support of this notion, control experiments with varying incubation times (0 - 60 min) performed on 8 μM Rca demonstrated that the results do not depend on the time elapsed between sample preparation and data acquisition. Above ~ 550 nM, the FCS decays of solutions containing 50 nM labeled and increasing amounts of unlabeled Rca show a continuous shift to longer timescales (Fig. 5A), suggesting the formation of oligomers whose size and relative concentrations depend on total protein concentration (labeled and unlabeled). Control experiments carried out on free Alexa 546 in buffers containing up to 127 μM unlabeled Rca showed no measurable changes in the FCS decay of the dye, ruling out the possibility that the shifts observed with labeled Rca are due to changes in solution viscosity or index of refraction.

Consistent with previous reports [47, 52], the FCS data indicate that at subunit concentrations above 10 μM, Rca assembles to form supramolecular complexes

significantly larger than hexamers. In spite of the complexity of the assembly mechanism, all FCS decays can be fit with a simple one-component model (Eq. 2), such that an apparent diffusion coefficient (D_{app}) can be extracted (see below). Therefore, the ratio D_{app}/D_1 can be calculated for each experiment regardless of the number of contributing species (Fig. 5B-D). To provide a scale, the expected ratios $D_{app}/D_1 = k^{-1/3}$ for $k = 1\ldots20$ for the pure oligomeric forms are shown as red horizontal bars on the right side of the graphs. At 100 μM Rca, we observe a ratio of 0.38, substantially smaller than the value of 0.55 expected for $k = 6$, suggesting large aggregates bearing on the order of ~24 subunits. The variations in relative diffusion coefficients that would result from non-spherical molecular shapes and different amounts of bound water cannot account for this difference (see Supporting Material).

Rajagopalan et al. used FCS to investigate the self-association of the tumor suppressor protein p53, which was known to form tetramers from dimeric intermediates [71]. Because in this case the dissociation constants are separated by about two orders of magnitude, no more than two oligomeric forms co-exist at any protein concentration. As a consequence, a graph of diffusion time vs. protein concentration exhibits two clear inflection points at concentrations around the monomer/dimer and dimer/tetramer transitions, which allows the estimation of K_d values directly from this plot with minimal mathematical modeling (36). In contrast, we observe a smooth shift in the diffusion constant with concentration, indicating that Rca assembly involves a number of intermediates in a wide range of concentrations. This feature is supported by published SE-HPLC data [25, 47, 52, 58]. Therefore, the interpretation of FCS data in terms of Rca assembly requires modeling the diffusion coefficients of individual component species,

and a mathematical model that describes the total autocorrelation function in terms of the contributions of all oligomeric species present at each protein concentration.

Derivation of the total autocorrelation function to describe mixed-species FCS

The autocorrelation function for a polydisperse sample containing *M* components of well-defined diffusion coefficient (D_i) and brightness (ε_i) is [63]

$$G(\tau) = \frac{2^{-3/2}\sum_{i=1}^{M} N_i \varepsilon_i^2 g_i(\tau)}{\left(\sum_{i=1}^{M} N_i \varepsilon_i\right)^2} \qquad \text{(Eq. 5)}$$

where

$$g_i(\tau) = \left(1+\frac{4D_i\tau}{r_0^2}\right)^{-1}\left(1+\frac{4D_i\tau}{z_0^2}\right)^{-1/2} \qquad \text{(Eq. 6)}$$

is the normalized autocorrelation function for species *i*, r_0 and z_0 are the radial and axial semi-axis of the Gaussian confocal volume (see Materials and Methods), and N_i is the mean number of molecules of the *i*th species in the observation volume. In general, Eq. 5 will depend on the diffusion coefficients of all oligomers present in solution, and the dissociation constants of the different assembly equilibria. Despite the complexity of Eq. 5 we could fit all FCS decays with the one-component model of Eq. 2 to obtain an apparent diffusion coefficient, D_{app} (see Supplemental Material for a more in-depth discussion and representative examples). This indicates that the linear combination of several one-component decays gives a total autocorrelation decay that cannot be distinguished experimentally from that obtained for a monodisperse sample. It should be

noted that D_{app} is not the average of individual diffusion coefficients, since each species contributes to the autocorrelation function as the square of its brightness. For instance, dimers bearing two labels contribute four times as much to $G(\tau)$ than monomers bearing one label.

In Figs. 5B-D, experimental FCS results are represented as the ratio between D_{app} obtained from the fit to Eq. 2 and the diffusion coefficient for the monomer D_1 measured at 50 nM Rca. Error bars represent the standard deviation of at least four determinations. To obtain assembly models consistent with the experimental D_{app}/D_1 ratios, the brightness in Eq. 5 must be expressed in terms of the assembly state of the protein and the number of labeled subunits present in each oligomeric species. The brightness of each species ε_i is proportional to the number of fluorophores n (i.e. the number of labeled subunits present in the oligomeric particle). Assuming that the protein assembles into oligomers of k subunits, Eq. 5 can be expressed as

$$G(\tau) = \frac{2^{-3/2} \sum_{k} \sum_{n=0}^{k} n^2 N_{k,n} g_k(\tau)}{\left(\sum_{k} \sum_{n=0}^{k} n N_{k,n} \right)^2} \qquad \text{(Eq.7)}$$

where $N_{k,n}$ is the number of particles in the observation volume containing k Rca subunits and n fluorophores ($0 \leq n \leq k$), and $g_k(\tau)$ is the normalized autocorrelation function defined in Eq. 6. Eq. 7 is similar to the model discussed by Yu et al. to analyze the FCS decays of micelles containing a variable number of fluorophores, although in this work the authors examined only the amplitude of the decay and use a single diffusion coefficient for all micelles [83]. Due to the mixing of labeled and unlabeled subunits, the

number of fluorophores in a given oligomeric particle is a random variable distributed according to the binomial distribution. The probability that an oligomer of size *k* contains *n* labeled monomers is therefore given by

$$p_{k,n} = \binom{k}{n} f^n (1-f)^{k-n} = \frac{k!}{n!(k-n)!} f^n (1-f)^{k-n} \qquad \text{(Eq. 8)}$$

where *f*, the fraction of fluorescently-labeled protein monomers, is determined by the relative concentrations of labeled (C_L) and unlabeled (C_U) subunits present in the mixture: $f = C_L/(C_L + C_U)$. The value of $N_{k,n}$ is then given by $p_{k,n}N_k$, where $N_k = \sum_{n=0}^{k} N_{k,m}$

$\sum_{n=0}^{k} N_{k,n}$ is the total number of particles in the observation volume containing *k* subunits and any number of fluorophores. N_k can be expressed in terms of molar concentrations as $C_k V_{eff} N_{AV}$, where C_k is the concentration of the oligomer of size *k*, V_{eff} is the effective observation volume in liters ($V_{eff} = (\pi/2)^{3/2} r_0^2 z_0$) (26), and N_{AV} is Avogadro's number. The values of C_k, in turn, are determined by the dissociation constants of the different oligomerization equilibria and the total protein concentration ($C_L + C_U$).

The total autocorrelation function can then be expressed in terms of the concentrations of labeled and unlabeled Rca (C_L, C_U), the optical parameters of the setup (r_0, z_0), the diffusion coefficient of each oligomeric state (D_k), and the concentrations of the different oligomers in the solution (C_k).

$$G(\tau) = \frac{1}{\pi^{3/2} r_0^2 z_0 N_{AV}} \frac{\sum_k \sum_{n=0}^{k} n^2 \binom{k}{n} f^n (1-f)^{k-n} C_k \left(1+\frac{4D_k\tau}{r_0^2}\right)^{-1} \left(1+\frac{4D_k\tau}{z_0^2}\right)^{-1/2}}{\left(\sum_k \sum_{n=0}^{k} n \binom{k}{n} f^n (1-f)^{k-n} C_k\right)^2} \qquad \text{(Eq. 9)}$$

Model	**Equilibria**	K_d **values**
1	$6Rca \rightleftharpoons Rca_6$	50 μM^5
	$4Rca_6 \rightleftharpoons Rca_{24}$	10^3 μM^3
2	$3Rca \rightleftharpoons Rca_3$	5 μM^2
	$2Rca_3 \rightleftharpoons Rca_6$	5 μM
	$4Rca_6 \rightleftharpoons Rca_{24}$	35 μM^3
3	$2Rca \rightleftharpoons Rca_2$	3.5 μM
	$2Rca_2 \rightleftharpoons Rca_4$	1 μM
	$Rca_2 + Rca_4 \rightleftharpoons Rca_6$	1 μM
	$4Rca_6 \rightleftharpoons Rca_{24}$	25 μM^3

Table 1. Optimized dissociation constants. The dissociation constants (K_d) for each assembly step were obtained by modeling according to Eq. 9. Model 1 is judged to be inconsistent with experiment, whereas Models 2 and 3 provide good fits to the data.

We were able to estimate D_k values from the X-ray structure of tobacco Rca (see Materials and Methods). Therefore, C_k is the only unknown parameter in Eq. 9, with values depending on the K_ds and total Rca concentration. Therefore, the analysis of FCS decays requires previous knowledge of the number of different oligomeric species present and the mechanism by which each one is formed. Although this information is limited for Rca, different assembly mechanisms can be tested for their consistency with experimental data (Table 1). To this end, we first assume a set of values of *k* to represent the types of oligomers present in equilibrium, and a mechanism for their formation. A set of K_d values is then assumed, from which the concentration of each oligomer (C_k) is calculated as a function of Rca concentration ($C_U + C_L$). Eq. 9 is then used to predict the theoretical total autocorrelation function at each Rca concentration. As with the

experimental data, each of the theoretical decays is fitted to Eq. 2, from which a calculated value for D_{app} is obtained that describes the FCS decay of a complex mixture of species. The relative apparent diffusion coefficients D_{app}/D_1 expected for the assumed mechanism are plotted as a function of concentration (Fig. 5B-D, solid lines), and compared to the experimental data. The equilibrium constants are then modified iteratively until the calculated set of D_{app}/D_1 values is consistent with observations within error bars.

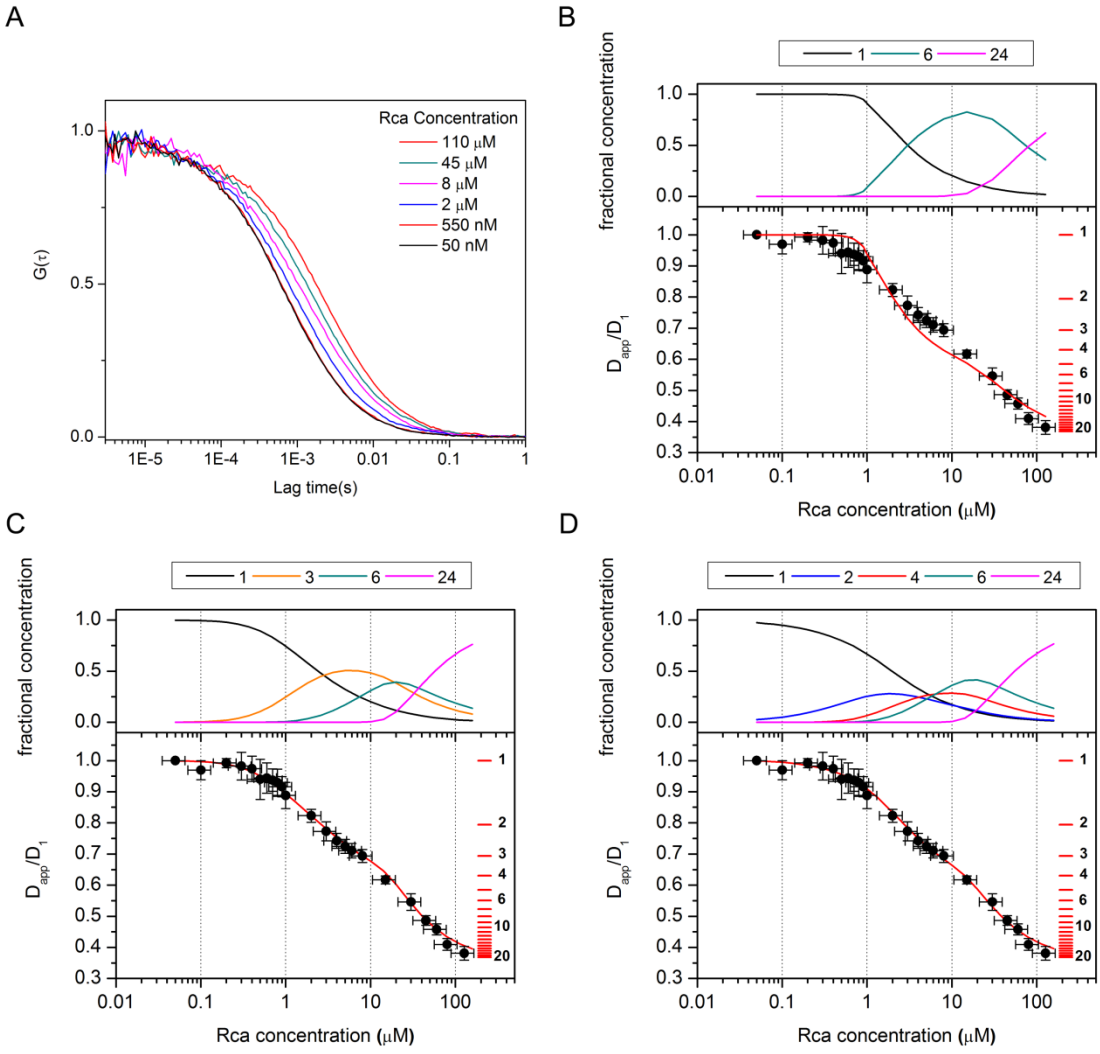


Figure 5. Rca assembly mechanisms. A. Representative FCS decays obtained with 50 nM labeled and increasing concentrations of unlabeled Rca. The concentrations in the inset represent total protein (labeled and unlabeled). **B-D**: Comparison of experimental results and modeling. The solid circles are experimental values of the ratio of the apparent diffusion coefficient at each concentration (D_{app}) and the diffusion coefficient obtained at 50 nM (D_1). The horizontal red lines are placed at values of $D_{app}/D_1 = k^{-1/3}$ for k = 1, 2… 20, and represent the expected D_{app}/D_1 values for monomers, dimers, etc. The solid curves were calculated according to Model 1 (panel **B**), Model 2 (panel **C**), and Model 3 (panel **D**) with the K_d values from Table 1. The concentration profiles above each graph (panels **B-D)** represent the fractional subunit concentrations of the different oligomeric forms assumed to contribute to each model, as calculated from the total Rca concentration and the K_d values in Table 1.

Interpretation of FCS data in terms of alternative Rca assembly mechanisms

Although to-date, the Rca assembly mechanism remains poorly characterized, some information on subunit stoichiometries is available. In our laboratory, we routinely observe dimeric forms by SE-HPLC (Wachter, unpublished results). Recently, hexameric forms have been observed by mass spectrometry and EM [21, 22, 59, 83]. In combination with the high propensity of classic AAA+ domains to hexamerize, these data provide strong support for a six-subunit intermediate state. Therefore, we analyzed the experimental FCS decays in terms of Rca self-association from monomers to hexamers to larger aggregates. Within this framework, we considered three mechanisms, Model 1 (monomer-hexamer-24mer), Model 2 (monomer-trimer-hexamer-24mer), and Model 3 (monomer-dimer-tetramer-hexamer-24mer) (Table 1). In Fig.5 B-D, we show predicted and experimental D_{app}/D_1 values, as well as the fractional concentration $kC_k / \Sigma_k kC_k$ of each oligomer as calculated from the predicted K_d values (panels above each graph).

Using Eq. 9, our calculations predict that Model 1 would cause a sharp decrease in D_{app}, as illustrated by the red curve in Fig. 5B, which shows the predicted autocorrelation function using K_{d1} = 50 μM^5 and K_{d2} = 10^3 μM^3. However, the observed relative D_{app} values decrease rather smoothly in the low-μM range (Fig. 5B, solid circles), suggesting the formation of intermediates. Although a reduction in K_{d1} would provide a better fit to the experimental data at the very low concentration range, such an adjustment would drop the D_{app}/D_1 ratio to about 0.6 around 1 μM Rca, whereas the experimental values remain high at about 0.9. Therefore, Model 1 is judged to be inconsistent with the data.

At least one intermediate is needed to predict the more moderate decrease in diffusion coefficients observed in the low-μM range. Therefore, in Model 2, we introduce a trimer as an intermediate. The best match between experimental data and modeling was obtained using K_{d1} = 5 μM^2, K_{d2} = 5 μM and K_{d3} = 35 μM^3 (Table 1 and Fig.5C). In order to estimate the uncertainties in dissociation constants, we evaluated a range of K_{d1} values that produce results consistent with experiment. These tests demonstrate that K_{d1} values in the range of 3.5-10 μM^2 provide a good fit to the data within experimental error (Fig. 15A). The values of K_{d2} and K_{d3} used to generate these curves were optimized for the lowest and highest K_{d1} values in this range (K_{d2} = 15 μM, K_{d3} = 2 μM^3 for K_{d1} = 3.5 μM^2; K_{d2} = 1.3 μM, K_{d3} = 70 μM^3 for K_{d1} = 10 μM^2). Clearly, a single trimeric intermediate is sufficient to describe the FCS results adequately.

However, based on the observation of dimers (see above), we considered an additional mechanism that proceeds through dimeric and tetrameric forms (Model 3, Fig. 5D). For this assembly path, the best agreement with experimental data was obtained with K_{d1} = 3.5 μM, K_{d2} = 1 μM K_{d3} = 1 μM and K_{d4} = 25 μM^3 (Fig. 5D). Because Model 3 has an additional adjustable parameter compared to Model 2, a wider range of K_d values provide a good fit, such as K_{d1} values in the range of 2.0-20 μM (Fig. 15B). The values of K_{d2} - K_{d4} were optimized for the values of K_{d1} on each extreme of the range (for K_{d1} = 2 μM, K_{d2} = 2 μM, K_{d3} = 1 μM, and K_{d4} = 20 μM^3; for K_{d1} = 20 μM, K_{d2} = 0.03 μM, K_{d3} = 2 μM, and K_{d4} = 8 μM^3). The concentration profiles predicted by this set of dissociation constants show more variations than the profiles obtained for Model 2 (Fig. 15B). This is

a consequence of the fact that our experimental results can be adequately described with one less species (Model 2).

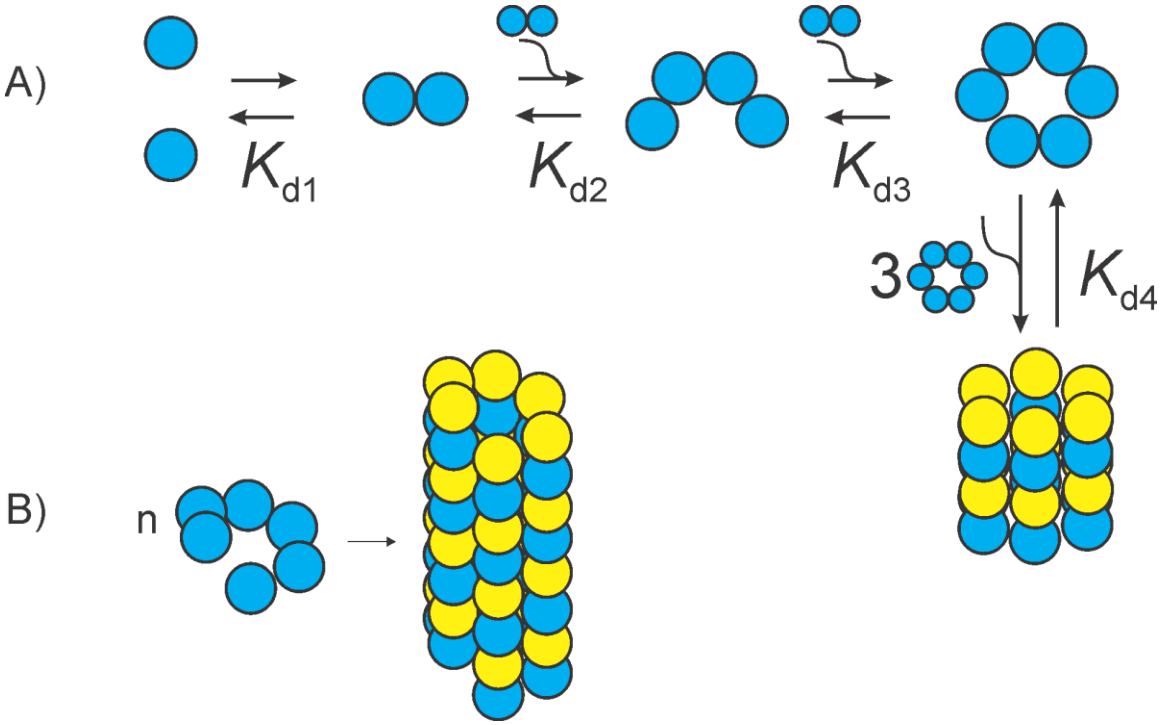


FIGURE 6. Rca assembly pathways consistent with experimental data. Schematic representation of Model 3 involving dimeric and tetrameric intermediates on the pathway of hexamer formation. Higher-order assemblies are shown as stacks of closed hexameric rings (**A**) or as pseudo-hexameric spirals (**B**).

Even at the highest concentrations tested, a clear plateau in D_{app}/D_1 is not observed, suggesting that more highly aggregated states are in equilibrium with assemblies of roughly 24 subunits. As a precise MW determination of large aggregates is not possible, we suggest that hexameric species self-associate to form 24-subunit complexes (Fig. 6). However, it is important to stress that we cannot distinguish between 18-mers and 24-mers due to the uncertainty inherent in our experiments (Fig. 5B-D, red horizontal scale bars). Additionally, we cannot rule out that larger species are formed by the continuous binding of dimers to smaller forms ($Rca_k + Rca_2 = Rca_{k+2}$), leading to the formation of helical arrangements as observed in the Rca AAA+ crystal structure [21] (Fig. 6).

DISCUSSION

Frequently, the mechanistic enzymology of complex protein assemblies is less well understood. A detailed molecular description of subunit cooperativity requires knowledge of the stoichiometries and fractional concentrations of all component species. We have developed an FCS-based method to address this question, and have tested our methodology on the highly aggregate-prone chaperone-like protein Rca. We demonstrate that the low-nanomolar to mid-micromolar range is accessible by mixing labeled and unlabeled protein, and anticipate that this method will be particularly useful for the examination of dynamic equilibria that involve fast subunit exchange. Although in higher plants, the ATPase activity of Rca regulates the level of carbon fixation, its assembly process has remained intractable for many years [47, 52]. The FCS experiments presented here provide a framework for the development of appropriate thermodynamic models for

self-association. Our data provide clear evidence of multiple oligomeric states under equilibrium conditions. Although the FCS results appear to mimic a wealth of SE-HPLC data, chromatography cannot provide equilibrium stoichiometries for a highly dynamic system, a feature that has impeded a thorough mechanistic investigation of Rca.

In the present work, we have monitored the Mg·ADP-mediated assembly process of cotton β-Rca at pH 7.6. Our PCH and FCS results demonstrate that the monomer is the dominating species at subunit concentrations below 0.5 μM, in line with EM images of amorphous particles at 0.5 - 1 μM tobacco Rca-R293V with Mg·ADP [21]. Within the 100 μM concentration range, our data are consistent with the formation of high-MW aggregates comprising roughly 24 subunits. As the assembly process appears to be continuous, with no obvious plateau at any concentration, stoichiometries of intermediate states cannot be extracted by FCS (Fig. 5B-D). Therefore, we use evidence from other sources to interpret the FCS results. Experimental evidence of Rca hexamers has been accumulating [21, 22, 59]. By EM, the ATP-dependent formation of closed 6-unit rings has been observed for 0.5 - 1 μM tobacco Rca-R293V, whereas the crystal structure provided a spiral hexameric model [21]. In general, members of the AAA+ classic clade have been crystallized in both helical and closed-ring hexameric arrays, although only toroids are considered part of the ATPase cycle. For example, the AAA+ domain of FtsH with attached protease forms a toroidal hexamer [84], whereas the absence of the protease domain leads to helical assemblies [85]. The assembly pathway of the related AAA+ protein ClpA involves dimeric and tetrameric forms [42, 86], providing support for Rca assembly according to Model 3. However, in contrast to Rca, the ClpA Mg·ATP-

dependent pathway involves tight complexes that can be isolated, with a dimer-tetramer K_d of ~ 1 nM and a tetramer-hexamer K_d below 0.17 nM [42].

Rca assembly mechanisms

We have considered three different assembly mechanisms that proceed to the ~24-subunit stage (Table 1), and find that the formation of hexamers must occur through at least one (Model 2), or more likely, two intermediate steps (Model 3). Model 3 suggests that the rise of the hexamer is initiated at 1 μM Rca (Fig. 5D), and dominates between 10 and 30 μM, whereas the tetramer contributes significantly between 5 and 20 μM. Higher-order assemblies ($k > 6$) start forming at about 15 μM, and become dominant above 40 μM. In the high range, the data do not provide sufficient detail to warrant the modeling of forms intermediate to 6- and 24-subunits, and therefore cannot be used to distinguish between models of closed hexameric rings (Fig. 6A) and pseudo-hexameric spirals (Fig. 6B). Although some AAA+ proteins contain stacks of toroidal rings [57], the helical filament model appears particularly attractive, because it allows for the continuous assembly of extended structures by the sequential addition of dimers. The K_d values extracted for the dimer-tetramer and tetramer-hexamer equilibria are estimated to be about 1 μM (Table 1), suggesting that the addition of a dimer to either a dimer or a tetramer is energetically equivalent. This feature supports a spiral assembly (Fig. 6B) and provides a rational for the broad size distribution observed by SE-HPLC [47, 52]

The present work provides evidence that under typical kinetic assay conditions (5 - 20 μM Rca), as much as four different Rca species coexist (Fig. 5B-D). Our method has

opened the door to the future investigation of Rca assembly as a function of different ADP/ATP ratios, as modulated by the energy charge of the chloroplast stroma. Long-term, we expect that a complete thermodynamic description of Rca assembly will aid in elucidating the mechanism of Rubisco reactivation.

SUPPORTING MATERIALS: PROTEIN OLIGOMERIZATION MONITORED BY FLUORESCENCE FLUCTUATION SPECTROSCOPY: SELF-ASSEMBLY OF RUBISCO ACTIVASE

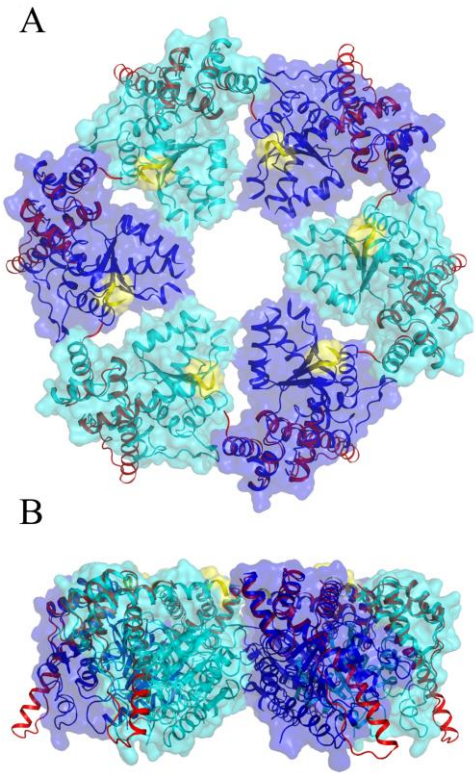


Figure 7. A model of the closed ring tobacco Rca apo-AAA+ hexamer (alternating cyan and blue protomers; PDB ID 3ZW6) from cryo electron microscopy data with the creosote recognition domain x-ray structure (red; residues 250-351; PDB ID 3THG) superimposed. Shown in yellow are the locations of the Tobacco AAA+ fragment C-termini. **A)** View down the 6-fold axis of the hexameric model. **B)** Side-on-view with the model rotated up 90°.

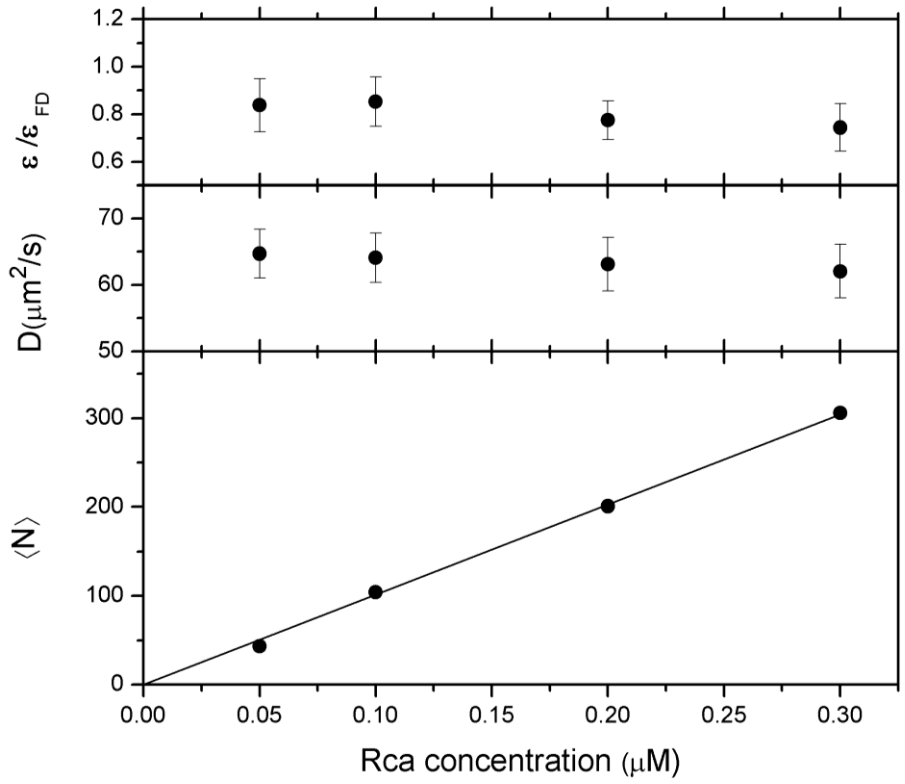


Figure 8: Rca is a monomer below 300 nM. Top: Ratio of the molecular brightness measured in experiments with labeled Rca (ε) and free Alexa 546 dye (ε_{FD}) as a function of Rca concentration. Middle: Diffusion coefficients obtained from fitting the FCS decays measured with labeled Rca in the 50-300 nM range. Bottom: mean number of diffusing particles obtained from the amplitude of the FCS decays measured with labeled Rca (Eq. 2). Put together, these results are consistent with a monomeric form of Rca.

Analytical data on Alexa-labeled Rca preparations.

To calculate the labeling stoichiometry of Alexa-derivatized Rca, protein preparations were analyzed by reverse phase HPLC (Figure 9). Eluting protein fractions were collected and absorbance scans were measured (Figure 10). The molecular masses of labeled Rca and control reactions in the presence of DTT were determined by MALDI (Figure 11). For experimental details, see Methods.

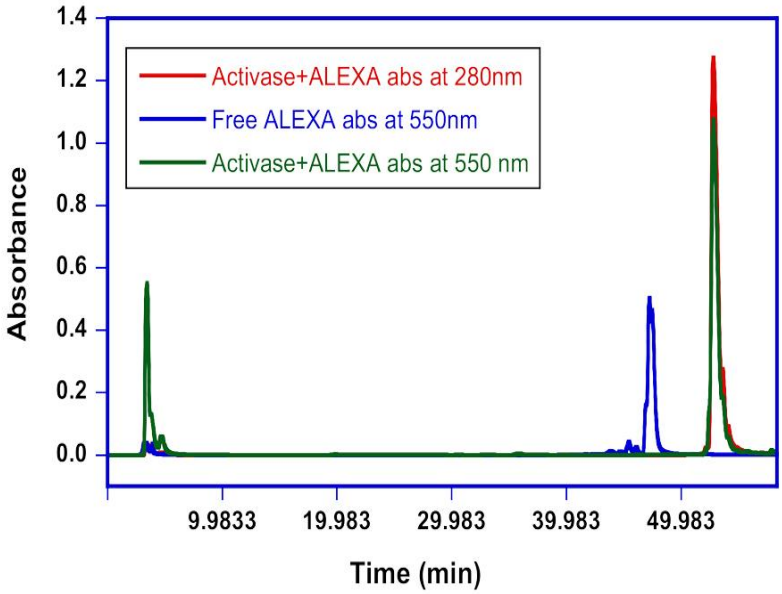


Figure 9. HPLC chromatograms. Reverse-phase HPLC chromatograms of Alexa-labeled Rca monitored at 280 nm (red) and at 550 nm (green), and free Alexa dye (blue). The chromatograms indicate that labeled Rca preparations do not contain any traces of free Alexa dye.

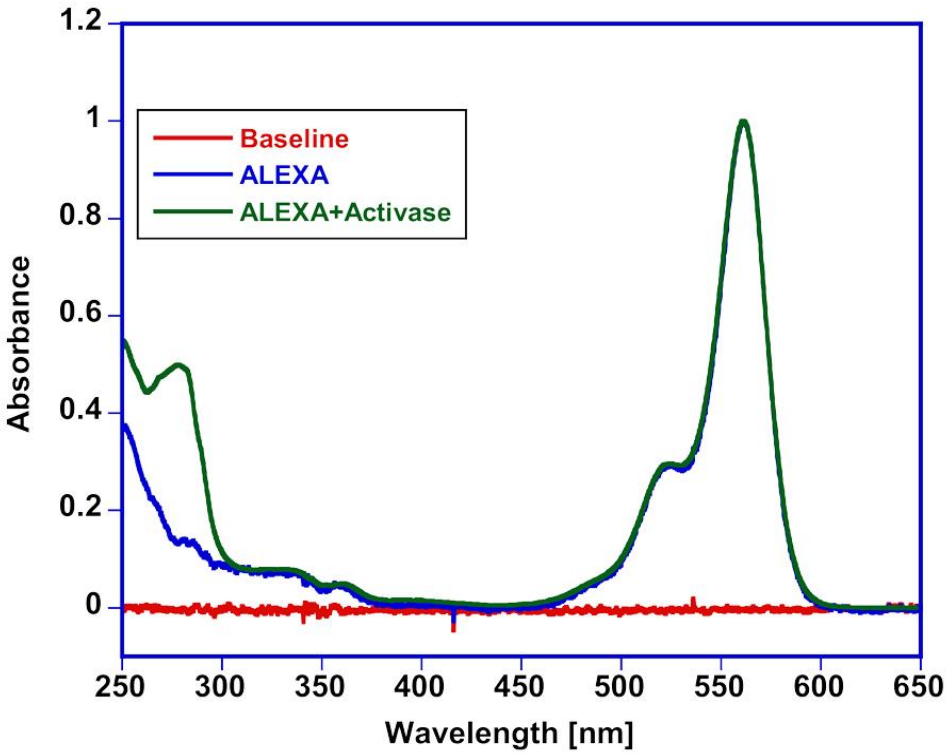


Figure 10. Absorbance spectra. The 55-min peak eluting from the HPLC column (labeled protein) was collected and analyzed by absorbance. The molar ratio of Alexa:Rca in covalently labeled preparations was determined by absorbance as described in the Methods section.

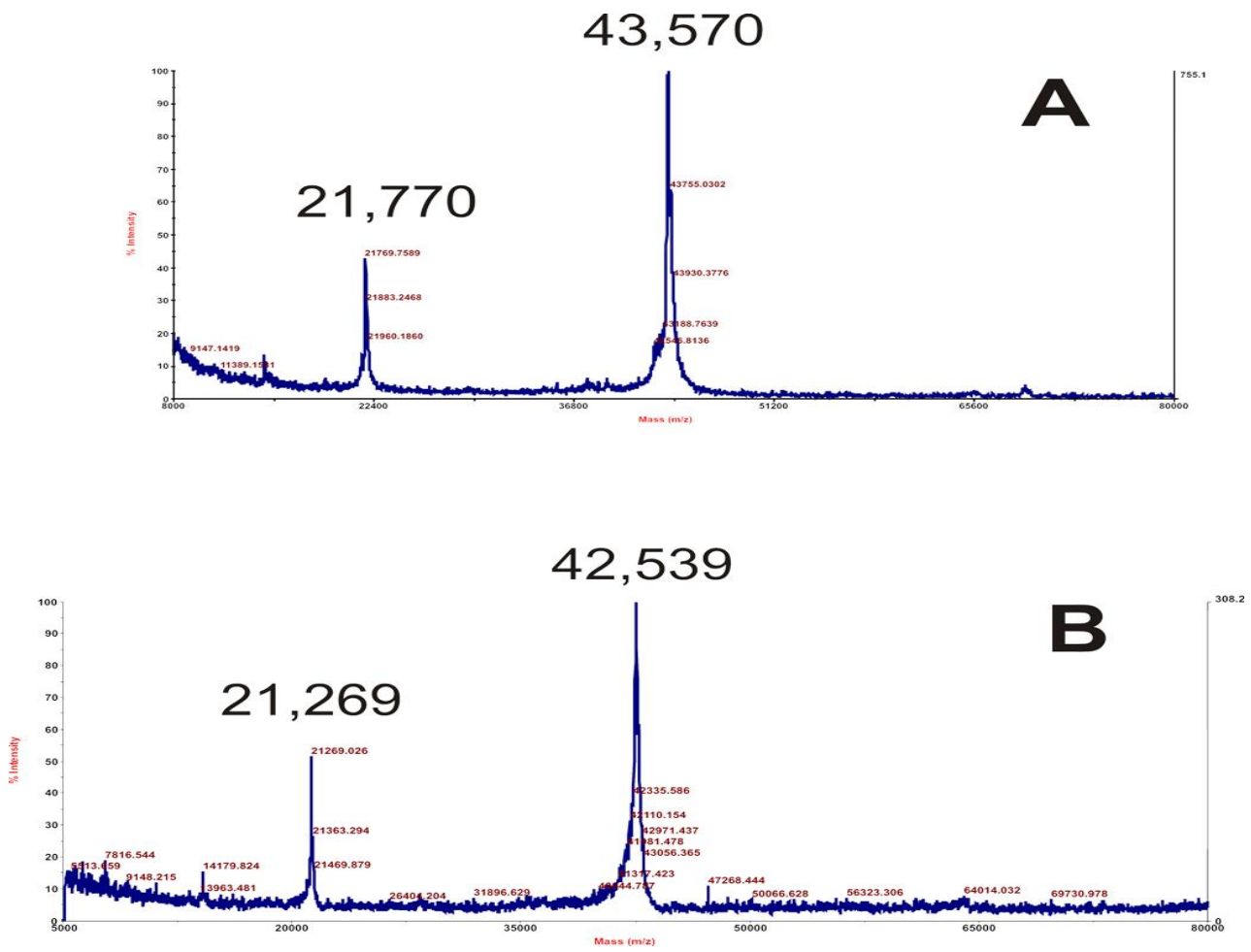


Figure 11. MALDI spectra. Spectra were collected on labeled and unlabeled Rca preparations, and compared to their theoretical values. (A) Labeled protein spectra exhibit main peaks with m/z = 43,570 (z=1), and m/z = 21,770 (z=2). The calculated mass is 43,493 g/mol. (B) Unlabeled protein spectra exhibit main peaks with m/z = 42,539 (z=1), and m/z = 21,269 (z=2). The calculated mass is 42,459 g/mol. The error of the instrument is estimated to be about 90 Da for a protein of this size.

Diffusion coefficient of the monomer and oligomers

In this section we analyze the different approximations made in the calculation of the diffusion coefficients of the monomer and oligomers, and their impact in the interpretation of the experimental results.

Effect of molecular shape.

We have assumed that the different oligomeric states of Rca are approximately spherical. In order to evaluate the possible effect of molecular shape on the values of the ratios D_k/D_1, we modeled the protein as an ellipsoid. The diffusion coefficient of a non-spherical particle is smaller than the diffusion coefficient of a sphere of the same volume, and for an ellipsoid, the diffusion coefficient can be expressed in terms of the ratio of the axial and equatorial semiaxis of the molecule (p) in terms of Perrin's friction factors [87]

$$\frac{D_{sphere}}{D_{ellipsoid}} = \frac{p^{2/3}\xi}{\operatorname{atan}\xi} \quad \text{(oblate ellipsoid)} \qquad \text{(Eq.S1)}$$

$$\xi = \frac{\sqrt{\left|p^2 - 1\right|}}{p} \qquad \text{(Eq.S2)}$$

The structure of the Rca hexamer published by Stotz *et al.* [21] indicates that the particle can be modeled as an oblate ellipsoid with an axial ratio p = 0.42 (Figure 7). The diffusion coefficient of a particle with this shape is 94% of the value one would measure for a spherical particle of the same volume ($D_{sphere}/D_{ellipsoid}$ = 1.07, Eqs. S1 and S2). The influence of molecular shape in the absolute values of the diffusion

coefficients is therefore small, and it is expected to be insignificant in the estimated values of the ratios D_k/D_1.

Effect of hydration water.

In addition, we considered the effect of hydration water, which increases the radius of the diffusing particle and therefore affects the diffusion coefficient. The volume of the hydrated protein (V_h) is related to the volume of the dry molecule (V_{dry}) and the volume of hydration water (V_{water}) as

$$V_h = V_{dry} + V_{water} = \frac{4}{3}\pi R^3 + \frac{M_w}{N_{av}\delta_{water}} h \quad \text{(Eq.S3)}$$

where R is the radius of the dry sphere, M_w is the molecular weight of the protein, N_{av} is Avogadro's number, δ_{water} is the density of water and h is the mass of hydration water expressed as grams of water per gram of protein (typically in the range 0.2-0.6, [88]).

Because the diffusion coefficient is inversely proportional to the hydrodynamic radius,

$$\frac{D(h_1)}{D(h_2)} = \left(\frac{V_h(h_2)}{V_h(h_1)} \right)^{1/3} \quad \text{(Eq. S4)}$$

where h_1 and h_2 are two values of hydration water. Using R = 6.35 10^{-7} cm for the hexamer (as described in materials and methods) and M_w = 257.3 kDa, we obtain

$$\frac{D(h=0.6)}{D(h=0.2)} = 0.96$$

This shows that different amounts of bound water in this range change the absolute values of the diffusion coefficients by about 4%. If the amount of bound water is of the same order for the different oligomers, we expect that the effect in the relative values of diffusion coefficients (D_k / D_1 ratios) will be insignificant.

Comparison between different estimates of the diffusion coefficients.

As discussed in the manuscript, we considered two alternative approaches to estimate the diffusion coefficients of the different oligomeric forms of Rca. In the first approach, we use the relationship $D_k = D_1/k^{1/3}$ with the experimentally determined value of D_1 (64.7 $\mu m^2 s^{-1}$). In the second approach, we used the published structure of the reconstructed closed-ring hexamer of tobacco Rca to obtain approximate models for the monomer, dimer, trimer and tetramer (see Eqs. 3 and 4 in the manuscript)

The ratios D_k/D_1 obtained using these two procedures are compared below:

	Method 1 ($D_k = D_1/k^{1/3}$)	Method 2 (from structural data)
D_2/D_1	0.79	0.80
D_3/D_1	0.69	0.66
D_4/D_1	0.63	0.59
D_6/D_1	0.55	0.54

Because both approaches bear significant approximations, it is critical to demonstrate that our interpretation of the experimental results does not depend on the approximations used in the analysis of the data.

To illustrate this point Fig. 12 shows a comparison of the results obtained using the D_k values estimated from the structure of the reconstructed closed-ring hexamer of tobacco Rca (green line), and using the formula $D_k = D_1/k^{1/3}$ with $D_1 = 64.7\ \mu m^2 s^{-1}$ (red line).

This comparison demonstrates that the dissociation constants reported in the manuscript do not depend significantly on the assumptions behind the estimation of the diffusion coefficients.

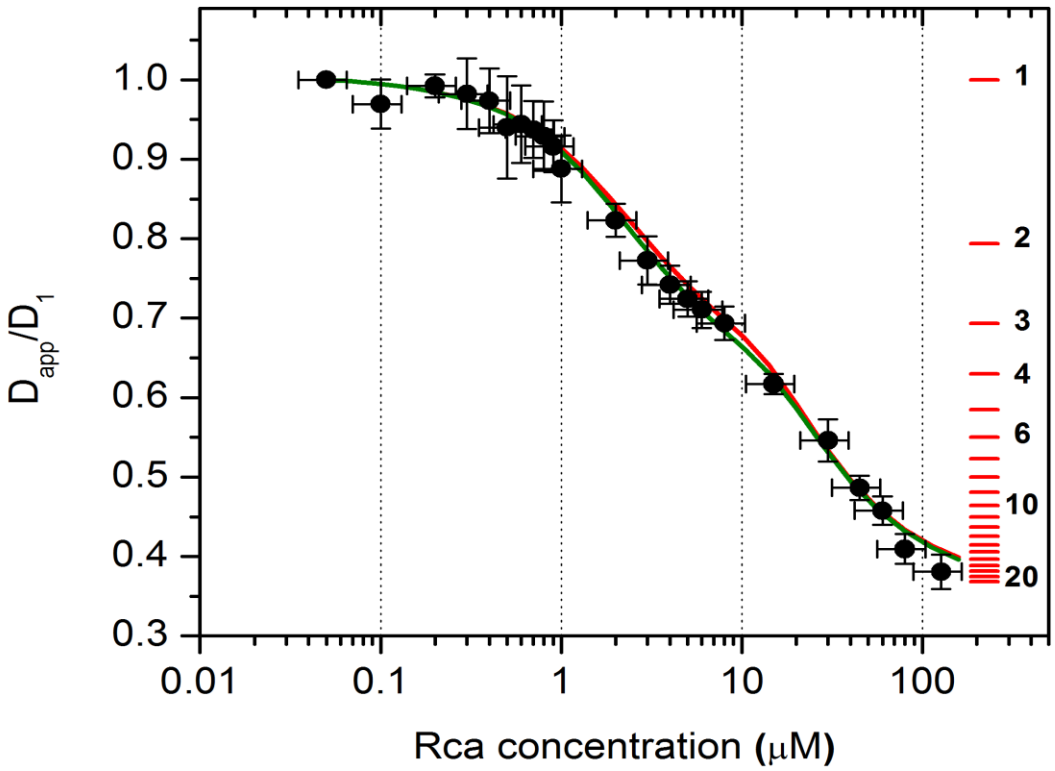


Figure 12. Results of modeling according to model 3 using K_{d1} = 3.5 μM, K_{d2} = 1 μM K_{d3} = 1 μM and K_{d4} = 25 μM[3]. The line in green is identical to the one presented in Figure 3D in the manuscript, and has been calculated using the diffusion coefficients calculated from the structure of the reconstructed closed-ring hexamer of tobacco Rca. The line in red has been calculated using $D_k = D_1/k^{1/3}$ with a value of D_1 equal to the diffusion coefficient determined at 50 nM Rca.

FCS analysis of polydisperse samples.

In the manuscript we stated that we could obtain adequate fits using a one-component FCS model even in situations where we strongly believe solutions contain two or more oligomeric species in equilibrium. Meseth et al [89] analyzed the issue of resolvability in the FCS of two-component solutions, and concluded that the diffusion times of the two species must differ by a factor of ~1.6. Although this value depends on a variety of experimental factors, it provides a reasonable "rule of thumb". When mixtures cannot be resolved, one obtains an apparent good fit using a one-particle model with an intermediate diffusion time. We performed simulations of FCS decays using the software simFCS (developed by Enrico Gratton) to show that it is reasonable to expect that the FCS decays of mixtures of monomers, dimers, etc will not be distinguishable from the FCS of monodisperse samples. To illustrate this point, Figure 12 shows the result of a simulation that would represent a condition close to what we expect for Rca at 10 μM. We set the simulation assuming that dimers, tetramers and hexamers are present at a 1/3 fractional concentration each. In addition, because the concentration of labeled Rca is much lower than the concentration of unlabeled Rca, the probability that an oligomer will contain more than one fluorescent probe is very small, so it is reasonable to assume that the brightness of dimers, tetramers and hexamers are the same. For the purpose of the simulation, we further assume that the diffusion coefficients of the species are related by $D_2/D_4 = 2^{1/3}$ and $D_2/D_6 = 3^{1/3}$, and we set D_2 at 52 $\mu m^2 s^{-1}$.

Figure 12-A shows the simulated FCS decay of the mixture of dimers, tetramers and hexamers described above (black line), and the result of fitting this decay with one component model (red line, and residuals below). The apparent diffusion coefficient

recovered from the fit is 40.7 $\mu m^2 s^{-1}$. We then performed a simulation where we assumed all particles had the same diffusion coefficient of 40.7 $\mu m^2 s^{-1}$ (Figure 12-B, black) and fitted the result with a one component model. We recover a diffusion coefficient of 40.3 $\mu m^2 s^{-1}$, which is very close to our input. Interestingly, the qualities of the fits (judged by the residuals and χ^2) are not significantly different. This exercise shows that FCS is not very sensitive to small differences in diffusion coefficients, and it is expected that all our decays can be fitted with a one-component model.

The quality of our experimental fits, as judged by the residuals, is not as good as the ones obtained in the simulations (Figure 13). This is not surprising giving the complications of experimental work, especially when working with proteins. Yet, we obtain similar fits routinely with other proteins that are known to be monomeric (and therefore solutions are in principle monodisperse). In addition, the quality of the fit is similar at Rca concentrations of 50 nM (where we have strong evidence that Rca is a monomer), and at higher concentrations, where we believe two or more oligomeric species co-exist. Using a two-model component improves the quality of the fit as expected, but this is also true in situations where we know proteins are monomeric. With this in mind, we believe that the fits obtained with the one-component model are good enough for the purpose of estimating the dissociation constants of Rca, and introducing more components in the model would not necessarily reflect the actual physical composition of the solution.

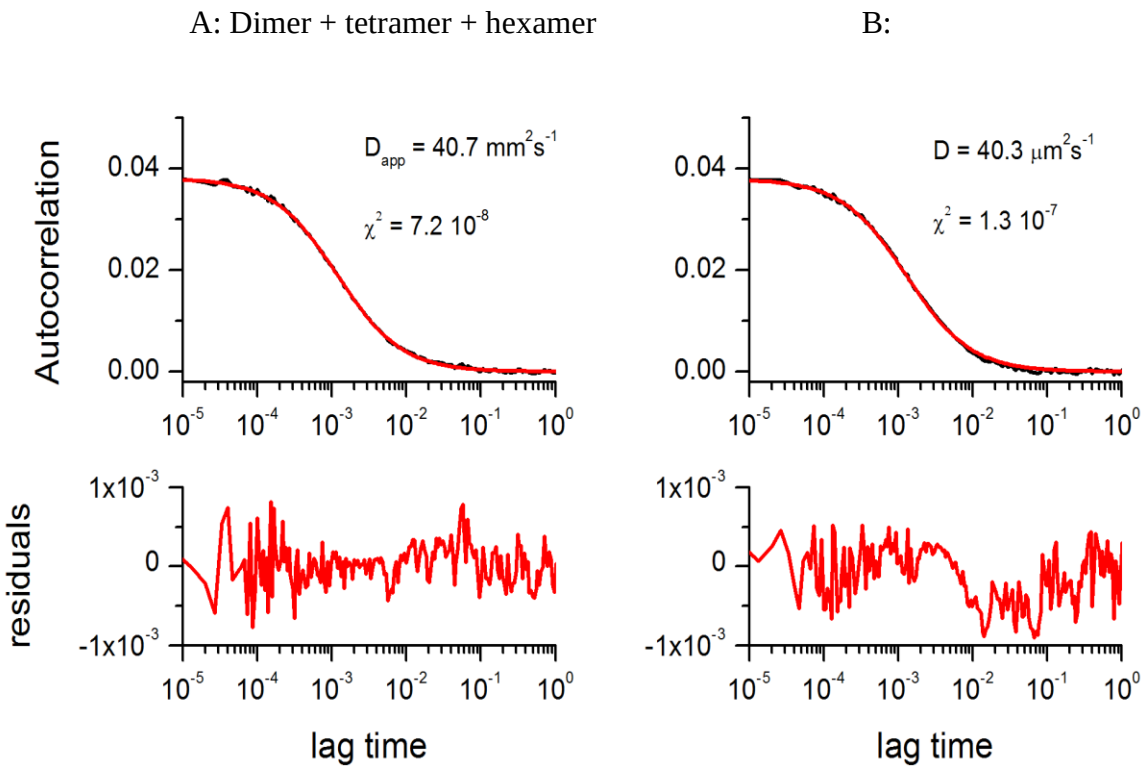


Figure 13. Stimulated FCS decays of a mixture of dimer, tetramer and hexamer (A) and a monodisperse sample (B) with the same diffusion coefficient as the apparent diffusion of the mixture. The mixture (fractional concentration of each component = 1/3) produces an FCS decay that can be fitted with a one-component model to give an apparent diffusion coefficient of 40.7 μm^2s^{-1}. The quality of the fit, judged by the residuals and χ^2 value, is excellent, and certainly not worse than the fit obtained for the monodisperse sample.

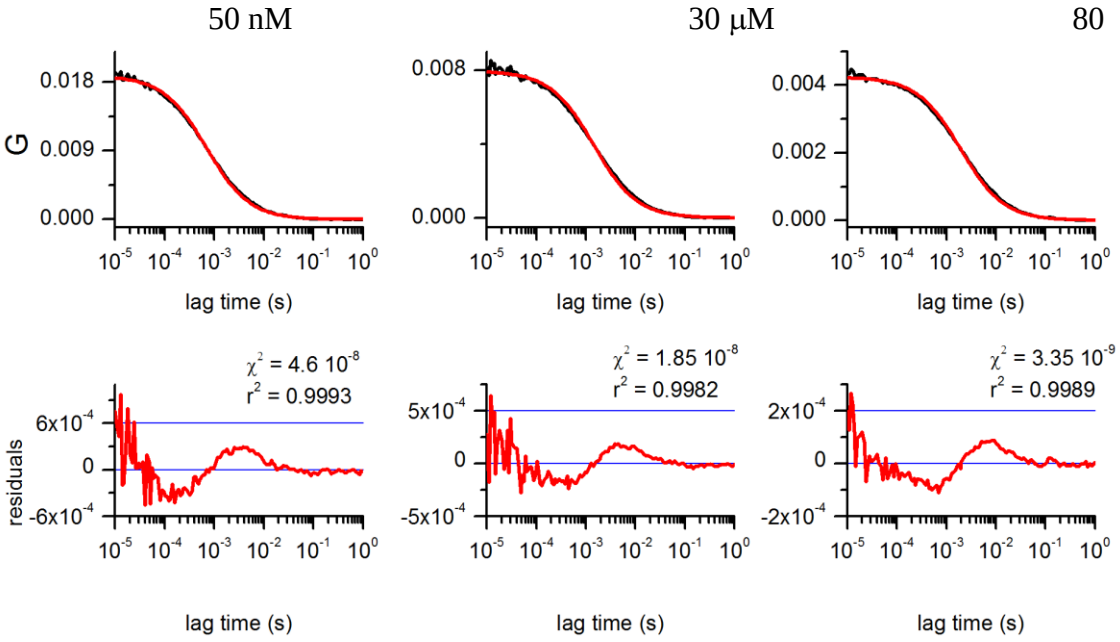


Figure 14. Representative experimental FCS decays (black lines) fitted with one component model (red lines). The apparent diffusion coefficients extracted from these fits were used to construct figure 1 in the manuscript. The value of χ^2 is calculated as the residual sum of squares divided by the degrees of freedom (N-2, where N is the size of the lag time vector). R^2 is the coefficient of determination. The qualities of the fits are similar at low and high concentrations, where we believe Rca is monomeric (50 nM) or highly polydisperse (high μM).

Uncertainties in the determination of the dissociation constants.

For models 2 and 3, we estimated the range of values of K_{d1} (first step in the assembly) that predict D_{app}/ D_1 values consistent with the experimental data. The other dissociation constants were then optimized in each case by keeping the value of K_{d1} fixed. In the case of model 2, we obtained a good fit to the experimental data using values of K_{d1} in the 3.5-10 μM^2 range (Fig. 14). This range has been determined arbitrarily by visual inspection considering the error bars in the D_{app}/D_1 and Rca concentration values, and represent only

estimates. The values of K_{d2} and K_{d3} used to generate these curves were optimized for the lowest and highest K_{d1} in this range (K_{d2} = 15 μM, K_{d3} = 2 μM^3 for K_{d1} = 3.5 μM^2, and K_{d2} = 1.3 μM, K_{d3} = 70 μM^3 for K_{d1} = 10 μM^2). The concentration profiles on the top panel were obtained with these two sets of K_d values.

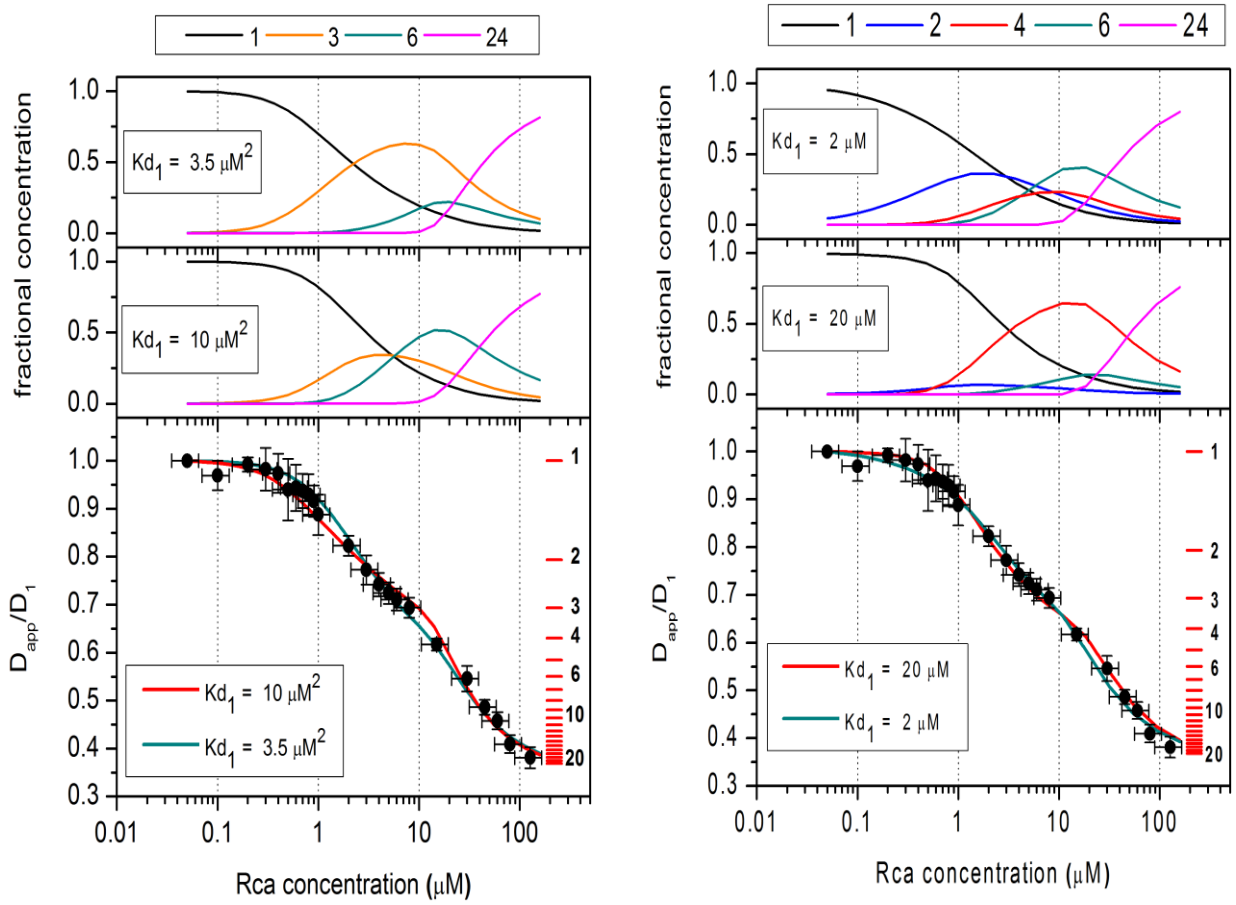


Figure 15. Results of modeling according to models 2 (panel A) and 3 (panel B). The solid circles are the results of experiments, and the red and green solid lines behind them are the results of modeling. **A:** model 2 using K_{d1} = 3.5 μM^2, K_{d2} = 15 μM and K_{d3} = 2 μM^3 (green) and K_{d1} = 10 μM^2, K_{d2} = 1.3 μM and K_{d3} = 70 μM^3 (red). **B:** model 3 using K_{d1} = 2 μM, K_{d2} = 2 μM, K_{d3} = 1 μM and K_{d4} = 20 μM^3 (green) and K_{d2} = 20 μM, K_{d3} = 0.03 μM, K_{d4} = 2 μM^3 for K_{d1} = 0.8 μM. The concentration profiles shown on top of graphs A and B show the fractional concentrations calculated using these sets of K_d values.

In the case of model 3, we can fit the experimental data using values of K_{d1} in the range of 2.0-20 μM (Fig. S9B). This range is larger than in the case of model 2 because model 3 has an additional adjustable parameter, and the number of parameters in model 2 is already enough to obtain a good fit to the data. The values of of K_{d2} - K_{d4} used to generate these curves were optimized for the values of K_{d1} on each extreme of this range (K_{d2} = 2 μM, K_{d3} = 1 μM, K_{d4} = 20 μM^3 for K_{d1} = 2 μM, and K_{d2} = 0.03 μM, K_{d3} = 2 μM, K_{d4} = 8 μM^3 for K_{d1} = 20 μM).

CHAPTER 3

ATP AND MAGNESIUM PROMOTE RUBISCO ACTIVASE HEXAMER FORMATION AT LOW MICROMOLAR CONCENTRATIONS[2]

[2] In preparation for publication as Kuriata, A. M., Chakraborty, M., Henderson J.N., Suratna H, Serban A., Levitus, M. Wachter, R. M. ATP and magnesium promote rubisco activase hexamer formation at low micromolar concentrations.

Abstract

We have characterized the dependence of variable nucleotide ratios on the complex assembly pathway of the chaperone-like AAA+ ATPase Rubisco activase (Rca) by fluorescence correlation spectroscopy (FCS) obtaining apparent diffusion coefficients from polydisperse distributions of labeled and unlabeled Rca under equilibrium solution conditions as a function of Rca concentration. FCS data are quantitatively compared to existing assembly models yielding estimates for the fractional concentrations of intermediates and corresponding dissociation constants (K_d). We collected FCS data for Rca from cotton in the presence of Mg•ATPγS and Mg•ADP and for the catalytically compromised D173N Walker B motif mutant in the presence of Mg•ATP. Our data are consistent with promotion and stabilization of hexamers by Mg•ATPγS and Mg•ATP, whereas Mg•ADP facilitates continuous assembly. Between 8 and 70 μM, 60-80% of Rca is predicted to exist as hexamers in the presence of Mg•ATP compared to just 30-40% with Mg•ADP. Dimer fractions peak between 1-4 μM and comprise 15-18% and 26-28%

of total protein for the Mg•ATP and Mg•ADP conditions, respectively. At 30 μM large aggregates begin to form comprising approximately 10% of total protein for Mg•ATP and 25% for Mg•ADP. We present an assembly model where Mg•ATP promotes and stabilizes hexamerization at low micromolar Rca concentrations relative to Mg•ADP, and suggest that this results from closed ring hexamer formation in Mg•ATP and open hexameric spiral formation in Mg•ADP.

Introduction

Rubisco [ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase] is a central enzyme in photosynthesis where it catalyzes the rate-limiting step in assimilation of inorganic CO_2 into organic building blocks responsible for all the world's biomass. Implications for agronomic practices, biofuel production and global carbon cycling ensure persistent interest in engineering or selecting for Rubisco enzymes with improved catalytic efficiency and increased crop yields. At issue are Rubisco's notoriously underwhelming characteristics as a catalyst. Low turnover numbers relative to other Calvin cycle enzymes necessitate Rubisco's high abundance and many side-reactions undermine its carboxylation activity, especially the competing oxygenation of RuBP that feeds the unproductive photorespiration pathway. When Rubisco first appeared on the early biotic earth, atmospheric CO_2 levels were many times higher and O_2 much lower than today, so O_2 competing with CO_2 for incorporation into RuBP was not problematic. As eons of oxygenic photosynthesis gradually raised O_2 levels at the expense of CO_2, evolution pushed Rubisco to bind RuBP more tightly and opened avenues for regulation

via non-productive substrate and inhibitory phosphorylated metabolite binding. Hence, a catalytic chaperone, known as Rubisco activase (Rca), was recruited to restore activity to inactive Rubisco by releasing inhibitors from Rubisco active sites through an as yet undetermined mechanism.

Rca belongs to the AAA+ (ATPase associated with diverse cellular activities) superfamily of ATPases and is part of the extended group [90]. Despite diversity in sequence and function, a common theme with AAA+ proteins is the use of energy derived from ATP hydrolysis to do chemo-mechanical work on substrates macromolecules. AAA+ proteins contain a structurally conserved ATP-binding cassette of approximately 200-250 residues consisting of an N-domain with the Walker A ($GX_4GK[T/S]$, where X is any amino acid) and Walker B (F_4DE, where F is a hydrophobic residue) motifs attached to a less conserved C-domain typically involved in substrate recognition. A cleft between the two domains provides the pocket for ATP binding, which frequently directs the formation of active arrays [15, 91]. Stoichiometry and assembly pathways in AAA+ oligomers are directly relevant to biological activity and therefore have attracted much interest. Oligomerization brings neighboring protomers together forming a bipartite ATP binding-site with a conserved arginine (arginine finger) from one protomer interacting directly with the γ-phosphate of the ATP in the adjacent protomer [15]. Presumably, this arrangement allows cycles involving ATP binding, hydrolysis and ADP release to facilitate conformational changes in the assembly that in turn mediate substrate remodeling or translocation events. Some AAA+ proteins require macromolecular binding partners to act as scaffolds facilitating proper assembly [55] and often there is a requirement for nucleotide binding [30, 42, 92, 93]. Numerous reports

suggest that assemblies with subunits in hexameric arrangements are most common [14, 30, 42, 94, 95], though there are exceptions [96-98].

The highly polydisperse nature of Rca protein preparations has long hampered studies of self-association. Size exclusion chromatography experiments describe broad, asymmetric bands consisting of concentration-dependent distributions with molecular weights ranging from monomers to extremely large assemblies of nearly 50 subunits [25, 99-101]. Such a wide range in estimated sizes reflects some species-specific differences in the Rca enzymes studied, but mainly it underscores a strong dependence on assay conditions (type of nucleotide, Mg^{2+} concentration, ionic strength, poly ethylene glycol, temperature, Rca concentration, etc). Blayney et al. [102] have used NanoESI mass spectrometry to demonstrate polydispersity in the assembly states of ADP-bound wild-type (WT) tobacco Rca (NtβRca) with sizes ranging up to hexamers and reported stable formation of hexameric species for the NtβRca R294A variant in the presence of ATPγS. A hexameric species forming a closed ring structure was also observed in negative-stain electron microscopy (EM) images by Mueller-Cajar et al. [103] in the presence of ATP and RuBP for the red-type Rca from *Rhodobacter sphaeroides* (RsRca), also known as CbbX. In the same work, the 3.0 Å x-ray crystal structure of unassembled RsRca was reported along with mutagenesis data that are consistent with a model whereby Rubisco activation involves the transient pulling of the Rubisco large subunit's carboxyl-terminal peptide into a RsRca hexameric pore, facilitating release of inhibitory sugar phosphates. Two higher plant Rca crystal structures exist both showing classic AAA+ architecture. One comprises the *Larrea tridentata* (creosote bush) AAA+ module's substrate recognition domain [20] and a second, representing the nucleotide free state of the

Nicotiana tabacum (tobacco) core AAA+ module with N- and C-terminal truncations, crystallized in a hexameric spiral arrangement [21]. From the spiral structure a closed ring hexameric model was produced, which fit well into negative-stain EM images obtained from the Nt□Rca R294V variant in the presence of ATP or ATPγS [21]. As with RsRca, mutational studies of predicted central pore residues reduced or completed abolished Rubisco reactivation activity [21], suggesting a possibly role in peptide threading; however, the details of the interaction between Rca and Rubisco await an explicit definition.

Despite progress in the form of Rca crystal structures, numerous questions remain before the mechanism of Rca-catalyzed Rubisco activation is resolved. Chief among these are the conditions governing Rca assembly states and how to directly relate subunit stoichiometry with ATPase or Rubisco activation activity. Within the chloroplast stromal energy charge regulates Rca activity via the ATP:ADP ratio and further modulates activity, in species containing the Rca α isoform, by redox potential in a light-dependent manner. It is therefore understandable that many studies have sought to define the effect of binding by adenine nucleotide phosphates on Rca self-association and activity. The increase in both ATPase and Rubisco activation with increasing Rca concentration is well documented [25, 104] and suggests that oligomerization is essential for proper biological function. Gel filtration analysis revealed that nucleotide binding had little or no effect on NtβRca oligomerization [105], which is in contrast to the observed stimulation by ATP of spinach and cotton Rca assembly [25, 100, 105]. ATP, ADP and Mg^{2+} were each observed to increase self-association of cotton Rca by dynamic light scattering (DLS) [101]. Negative-stain electron microscopy images of RsRca (CbbX) displayed the

striking dependence of structure upon nucleotide state revealing amorphous particles of ~600-900 kDa in the absence of nucleotide, long fibers of ~5-10 MDa in the presence of Mg•ATP or closed ringed hexamers with Mg•ATP and RuBP [103]. In similar experiments, the NtβRca R294V variant formed hexameric rings in the presence of Mg•ATP or Mg•ATPγS and amorphous particles with Mg•ADP [21].

Given the hexameric arrangements seen for Rca constructs in EM images and the NtβRca AAA+ module crystal structure, it is interesting that activity assays mixing WT NtβRca with unidirectional subunit interface mutants concluded that as little as 3-6 subunits are necessary for activation of inhibited Rubisco [21], suggesting that hexamerization may not be a strict requirement for proper function. Keown et al. arrived at a similar estimation of 2-4 subunits using static light scattering (SLS) and analytical ultracentrifugation (AUC) to monitor the concentration-dependent assembly of WT NtβRca in the absence of nucleotides, which they related to ATPase and Rubisco reactivation experiments [27]. This work demonstrated continuous formation of larger assemblies as protein concentration increased with no plateaus to suggest accumulation of particular intermediates, leading the authors to favor a model where Rca forms an open, helical, spiral in solution similar to that observed in the NtβRca AAA+ module crystal structure [21]. Consistent with this, we have presented a qualitatively similar result and model based on earlier fluorescence correlation spectroscopy (FCS) studies of cotton β-isoform Rca in the presence of Mg•ADP [106]. The FCS data were analyzed in terms of alternative Rca assembly mechanisms with estimated dissociation constants (K_d) obtained through iterative refinement of simulated data for agreement with experimental data. This approach allowed the exclusion of an assembly mechanism involving three

species (monomers, hexamers and 24 mers) and produced a model comprising five species (monomers, dimers, tetramers, hexamers and 24-mers) with four associated K_d values, thus providing a means of calculating the fractional concentrations of each species as a function of Rca concentration.

Here we apply the previously developed FCS method [106] to further our understanding of Rca's role in biological carbon fixation by investigating the effect of ATPγS, ATP:ADP ratio and free Mg^{2+} concentration on Rca assembly over a range of protein concentrations (0.05-75 μM). FCS relies on temporal fluctuations in fluorescence intensity that arise from a small number of molecules present in an optically restricted volume and relates these fluctuations to dynamic processes such as Brownian motion through autocorrelation analysis. As such, FCS is widely used to study dynamic processes including protein oligomerization [107, 108]. To facilitate FCS studies, the residue Asp173 within the cotton Rca β-isoform's non-canonical Walker B sequence motif (CLFIND) was mutated to Asn. In AAA+ proteins, the aspartate residue of the □$_4$DE Walker B motif coordinates Mg^{2+} and the glutamate is thought to activate a water molecule for nucleophilic attack on the β-γ phosphodiester bond of ATP [14]. Mutants in the glutamate residue often block hydrolysis while retaining nucleotide binding activity, and numerous studies have employed these so-called 'substrate trap' mutants to address biological questions [92, 109-112]. Our FCS results show a clear trend indicating advancement of hexamerization and suppression of larger oligomers through ATPγS and ATP binding which is further influenced by the presence of free Mg^{2+} and contrast with the relatively expeditious formation of larger species in the presence of ADP or absence of free Mg^{2+}.

Results

Alexa 546 derivatization and characterization of Rca variants

Previously, we described a FCS method [106] to study the ADP-dependent assembly of a *gossypium hirsutum* (cotton) β-isoform Rca variant harboring a C-terminal Ala-Cys insert [113], hereafter referred to as GhβRca*. This entailed the development of a protocol for labeling GhβRca* with a maleimide-functionalized Alexa 546 fluorophore via the introduced cysteine residue for which we demonstrated efficient incorporation of one label per GhβRca* protomer for reactions carried out in the presence of ADP [106]. In the present study, MALDI spectra from Alexa 546-labeled GhβRca* D173N samples containing either ADP or ATP indicate a molecular mass consistent with protein containing a single fluorescent label (Figure 16). The same is true for GhβRca* samples in the presence of ATPγS; however, by contrast, MALDI spectra from GhβRca* in the presence of AMP-PNP showed signs of double label incorporation, suggesting destabilization and partial unfolding. Therefore, FCS analysis of GhβRca* in the presence of AMP-PNP was not pursued. Absorbance spectra of reverse phase HPLC-purified GhβRca* D173N (Figure 18) yielded molar ratios of 1.1 ± 0.1 (n = 2) and 0.9 ± 0.1 (n = 3) labels per protein chain in the presence of ADP and ATP, respectively. For GhβRca* with ATPγS, similar analysis gave a ratio of 1.1 (n = 1) label per protein chain (Figure 17).

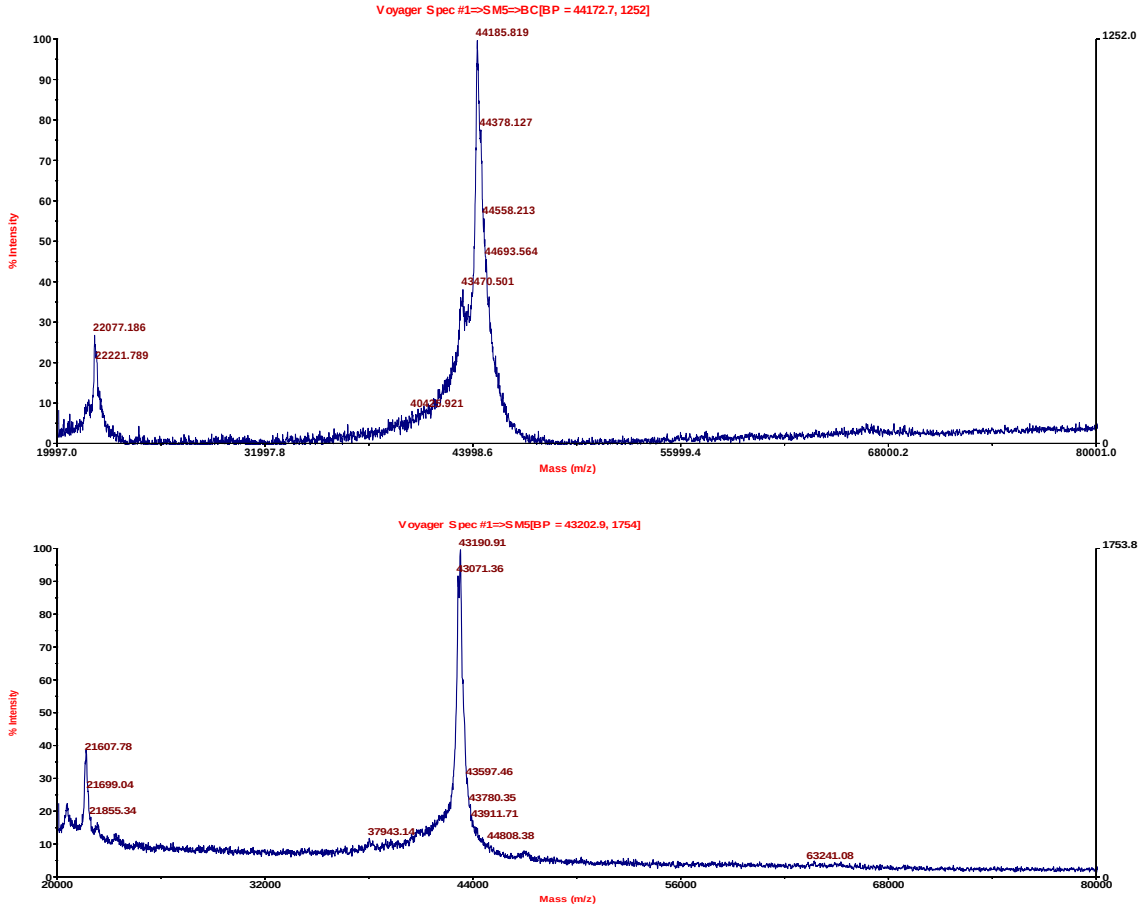


Figure 16. MALDI spectra. Spectra were collected on labeled and unlabeled Rca-D173N preparations, and compared to their theoretical values. (A) Labeled protein preparation. (B) Unlabeled protein spectra preparation. The error of the instrument is estimated to be about 100 Da for a protein of this size.

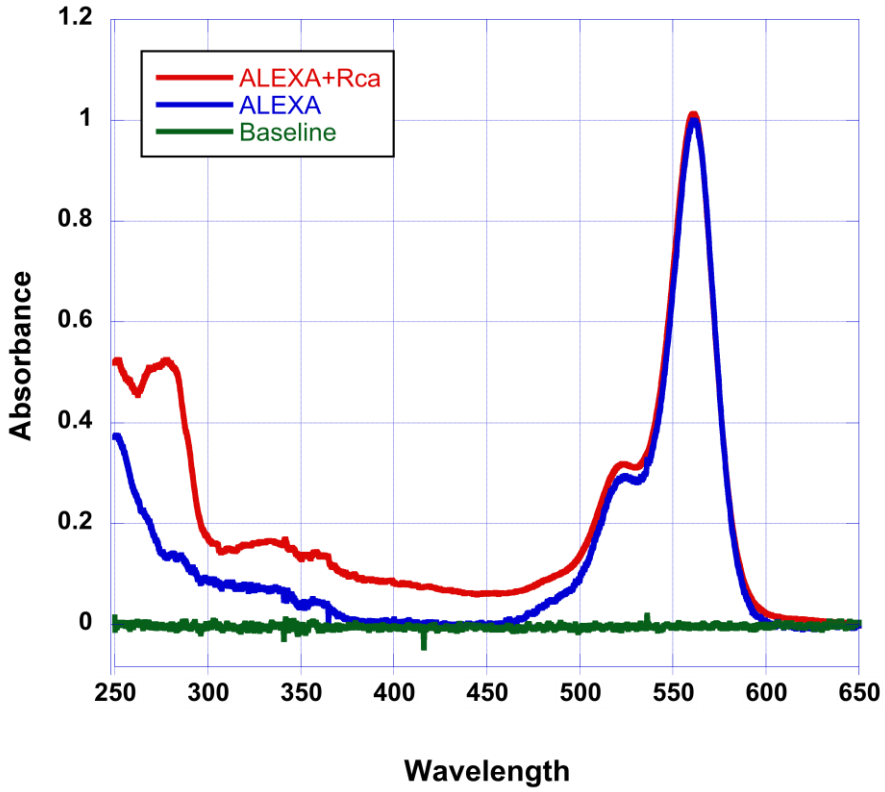


Figure 17. Absorbance spectra. The peak eluting from the HPLC column (labeled protein) was collected and analyzed by absorbance. The molar ratio of Alexa:Rca-D173N was calculated using Beer-Lambert Law.

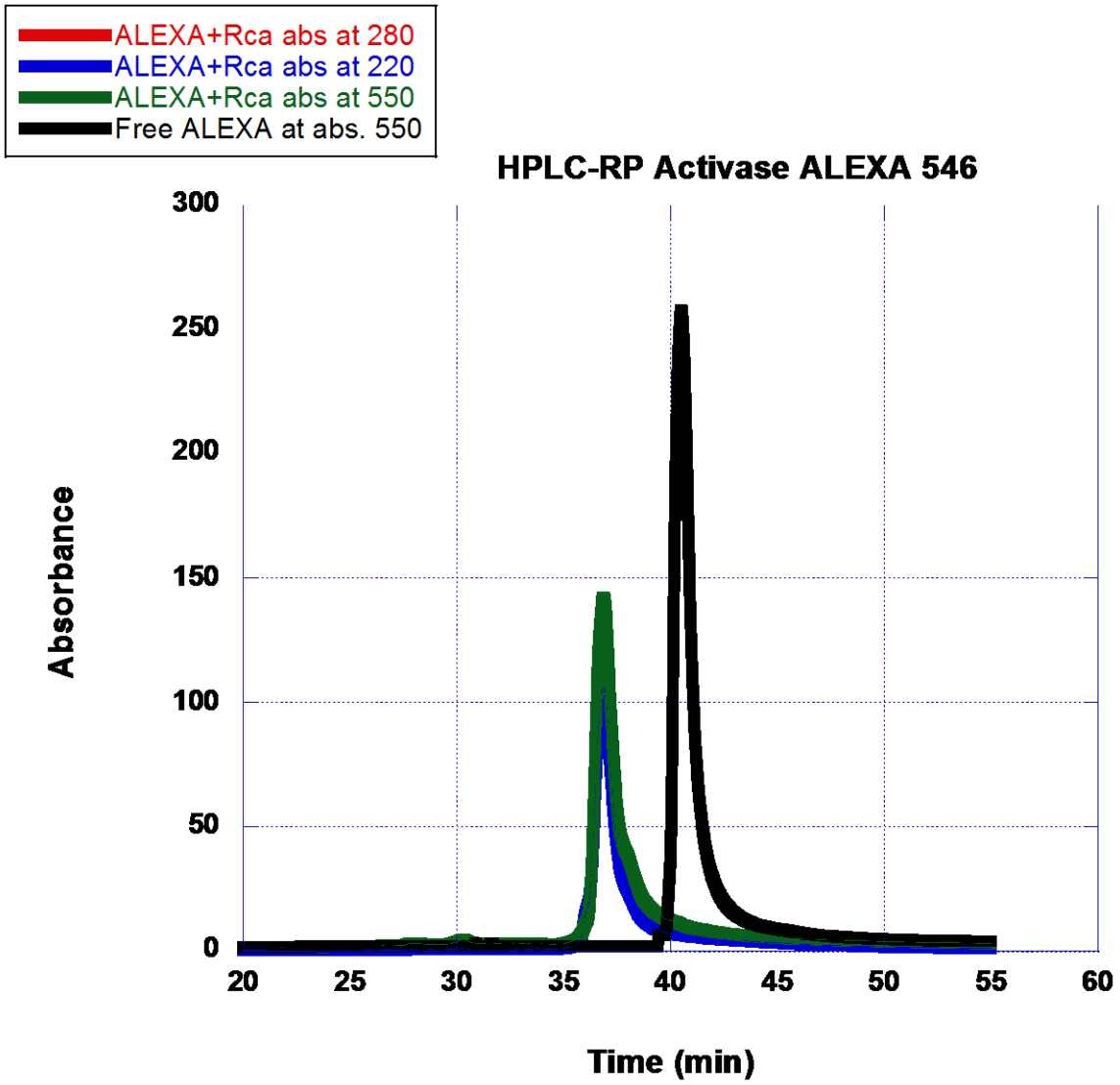


Figure 18. HPLC chromatograms. Reverse-phase HPLC chromatograms of Alexa-labeled Rca-D173N monitored at 280 nm (red), at 220 nm (blue), at 550 nm (green), and free Alexa dye (black). The chromatograms indicate that labeled Rca preparations do not contain any free Alexa dye.

ATPγS promotes Rca hexamer formation while ADP favors larger oligomers

FCS experiments with GhβRca* in the presence of Mg•ATPγS were carried out by analogy to the previously reported Mg•ADP study [106]. As before, each experimental protein concentration contained 50 nM labeled protein with higher concentrations obtained by mixing labeled and unlabeled Rca and assuming complete subunit equilibration leading to random label distribution throughout the assemblies over the course of a brief incubation on ice. Both experiments were carried out in the presence of 2 mM total nucleotide with 5 mM $MgCl_2$ and buffer at pH 7.6. Because slow hydrolysis of ATP□S to ADP was observed under conditions used for FCS sample preparation (Table 2), an average ATPγS:ADP ratio of 8 was estimated based on enzymatic assays measuring ADP production (see section below) with a range in ratios of 7-11. Between 50 nM and 300 nM, FCS decays for GhβRca* containing Mg•ATPγS do not change significantly in time scale (supplemental info), consistent with observation in the presence of Mg•ADP and indicating that the protein exists primarily in a monomeric state in this concentration regime. Indeed, a monomer in Mg•ATPγS over this protein concentration range has been confirmed via the same photon counting histogram (PCH) analysis used for the Mg•ADP bound samples [106]. As the concentration of GhβRca* increases above 300 nM, the FCS decays show a continuous shift to longer timescales (supplemental info) in the presence of Mg•ATPγS much the same way as Mg•ADP [106], indicating that under both conditions increasingly large oligomers are growing in population as a function of protein concentration with little indication of accumulation of one species. As previously described [106], the FCS decays are nonetheless adequately fit

to a single apparent diffusion coefficient (D_{app}) containing contributions from all the species present in solution.

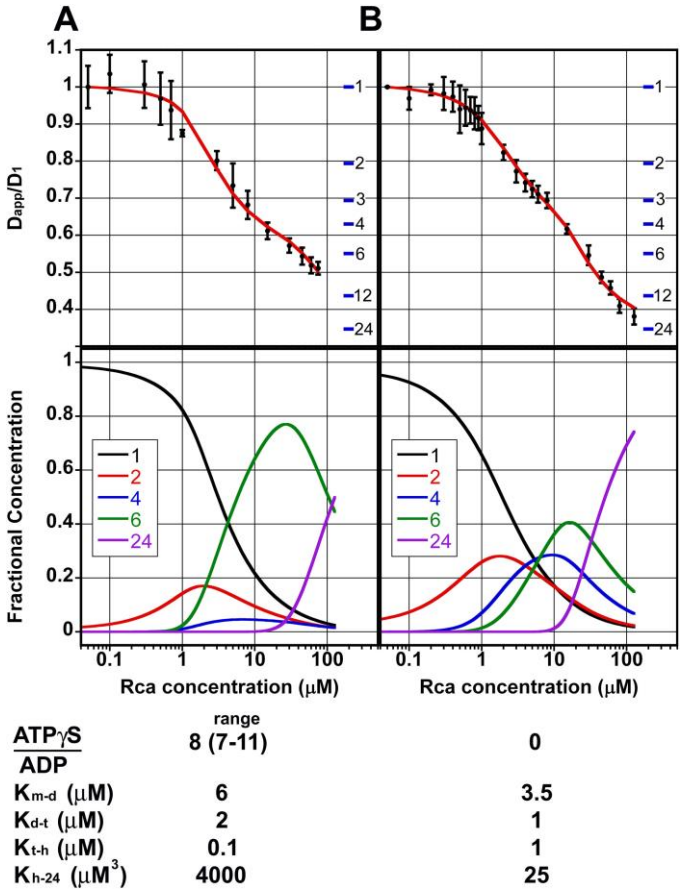


Figure 19: GhRca* assembly mechanisms under different nucleotide conditions: Each sample contained 25 mM HEPES-NaOH pH 7.6, 250 mM KCl, 5 mM $MgCl_2$, 10% glycerol and 2mM ATPS (A) or 2 mM ADP (B). Top panels: experimental (*black circles*) data for the relative apparent diffusion coefficient (D_{app}/D_1) obtained by taking the ratio of the apparent diffusion coefficient at each GhβRca* subunit concentration (D_{app}), and the diffusion coefficient obtained at 50nM (D_1). For each data point, 50nM of labeled GhβRca* was mixed with increasing amount of unlabeled GhβRca* to reach the desired protein concentration. The horizontal blue lines are placed at values of $D_{app}/D_1 = k^{-1/3}$ for k = 1-4, 6, 12, and 24; which represent the expected D_{app}/D_1 values for monomers, dimers, etc. The solid red curves simulate the experimental Rca assembly data under the different conditions and are calculated using K_d values given in the table below. Bottom panels: the corresponding total subunit concentration (*labeled* + *unlabeled*) showing the evolution of the fractional concentration of oligomeric species as they appear in each model, again, calculated from the total Rca concentration and the K_d values in the table below. Note, an estimated average ratio of 8 ATP-γS /ADP with a range of 7-11 is

reported due to slow hydrolysis during the on ice incubation of sample described in the text.

Figure 19 show that between 50 nM and 700 nM GhβRca*, samples containing Mg•ATPγS or Mg•ADP display very similar D_{app}/D_1 values, the ratio of the apparent diffusion coefficient normalized to the diffusion coefficient of a monomer species (D_1). For reference, the blue horizontal bars on the right side of the graph (Figure 19) delineate the expected ratios for D_{app}/D_1 for monomers (1), dimers (2), trimers (3), etc. At GhβRca* concentrations above 700 nM the D_{app}/D_1 data show the variable influences of Mg•ATPγS and Mg•ADP on GhβRca* assembly (Figure 19). Over the range of 0.7-15 μM, Mg•ATPγS-mediated assembly proceeds to higher-order oligomers of approximately 3-4 subunits more quickly than assembly in the presence of Mg•ADP. Above 30 μM GhβRca*, the opposite trend is observed with further oligomerization retarded in Mg•ATPγS when compared to the more continuous assembly in Mg•ADP.

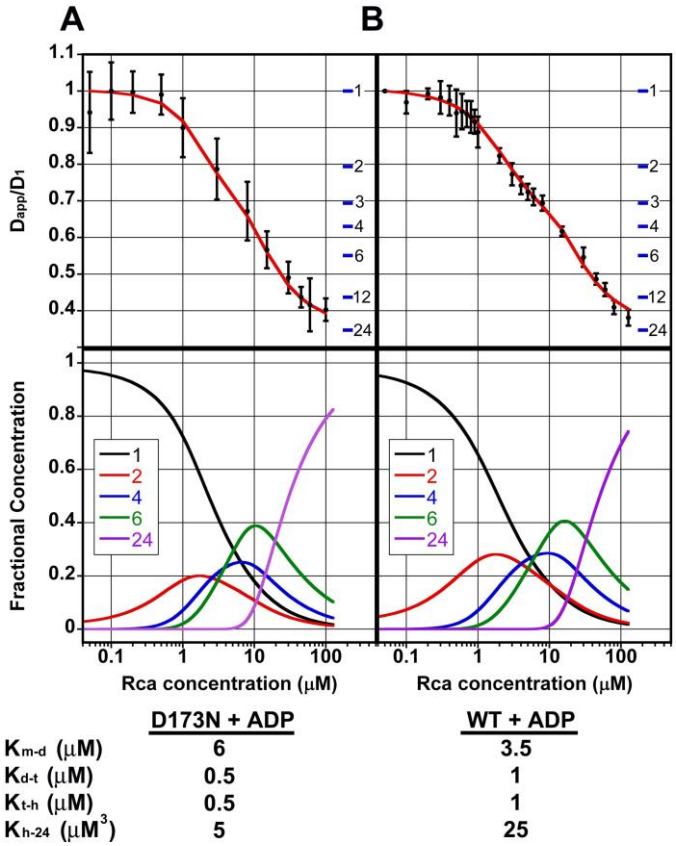


Figure 20: Assembly mechanism of GhβRca* and GhβRca* D173N in ADP: Each sample contained 25 mM HEPES-NaOH pH 7.6, 250 mM KCl, 5 mM $MgCl_2$, 10% glycerol and 2 mM ADP. Top panels: GhβRca* D173N (A) and GhβRca* (B) experimental data plotted as D_{app}/D_1 and presented as described above for Figure 19*. Bottom panels: the fractional concentration of oligomeric species for GhβRca* D173N (A) and GhβRca* (B) as they appear in each model calculated from the total Rca concentration and the K_d values in the table below.

In order to interpret the FCS data in the context of Rca assembly, we previously developed a mathematical model describing the total autocorrelation function with contributions from each oligomeric species present at a particular protein concentration [106]. The limitation in this analysis is that it requires some prior knowledge of the number of oligomeric species and their mechanism of formation. While this prevents explicit determination of the constituents of the polydisperse distribution, it allows us to

test different assembly mechanisms for agreement with experimental FCS data. For GhβRca* in the presence of Mg•ATPγS we have tested two of the pathways modeled for Mg•ADP-mediated assembly [106] and carried out analogous simulations of FCS data with iterative rounds of K_d optimization to yield values that are consistent within error bars of the experimental D_{app}/D_1 values. The first model involves concerted hexamerization of six monomers to a hexamer and subsequent association of four hexamers into a 24-mer. As was the case with the Mg•ADP assembly [106], data were difficult to simulate using this two-K_d model. The second model involves monomer-dimer, dimer-tetramer, tetramer-hexamer and hexamer-24-mer equilibria and, with the choice of appropriate K_d values, adequately describes the data. The Mg•ATPγS and Mg•ADP experimental and simulated data are shown in the top panel of Figure 19A and 19B, respectively. In the bottom panel fractional concentrations of assembly intermediates, derived from optimized K_d values found at the bottom of Figure 19, are plotted against GhβRca* concentration. This plots show that our data are consistent with models of assembly whereby Mg•ATPγS promotes and stabilizes the formation of hexameric species relative to Mg•ADP, which enhances the formation of larger oligomeric forms.

Nucleotide and Mg^{2+} binding monitored by Thermofluor assay

Potential solution conditions for FCS analysis of Rca assembly were examined with Thermofluor stability assays that compared the effect of increasing concentrations of various nucleotides in the presence or absence of excess free Mg^{2+} to determine how their binding influences the thermal stability of GhβRca* and its D173N variant. The Themofluor assay is a fluorescent dye-based measurement used to report on protein

unfolding. As temperature increases proteins begin to unfold baring their hydrophobic core. This increasingly exposes a solvatochromic dye to a medium of relatively lower dielectric constant giving rise to a temperature-dependent fluorescence increase. Although the apparent melting temperature (T_m) extracted is not a thermodynamic quantity, the technique does allow for rapid comparison between buffer conditions for the relative effect on the apparent T_m.

Apo GhβRca* and GhβRca* D173N protein preparations exhibit a strong increase in fluorescence in response to heating yielding apparent T_m values of 37.0 ± 0.6°C and 34.4 ± 0.4°C (Table 2), respectively. This indicates that the mutation is slightly destabilizing. As shown previously for GhβRca* [101], titration of the D173N variant with either ADP or ATP leads to incremental increases in apparent T_m values (data not shown) confirming that disruption of the catalytic Walker B aspartate residue does not eliminate nucleotide binding. Addition of ADP to apo protein raises the apparent T_m to 46.0 ± 0.2°C for GhβRca* and 45.4 ± 0.1°C for GhβRca* D173N (Table 2) suggesting a similar binding affinity for both variants. Furthermore, consistent with earlier data for GhβRca* [101], Table 2 shows that when excess free Mg^{2+} is added to either variant prepared with or without ADP, the effect is a drop in apparent T_m that is more pronounced with ADP (~ 3.1-3.3°C) than for apo (~ 0.9-1.3°C) preparations. For both variants ADP is more stabilizing than ATP, though the difference is much more pronounced for GhβRca* (~ 5.7 °C) than for GhβRca* D173N (~ 0.6°C). In fact, at 2 mM ATP, the D173N variant has an apparent T_m of 44.0 ± 0.2°C compared to 40.3 ± 0.5°C for the parent protein (Table 2). This indicates stronger ATP binding in the D173N variant and is likely attributable in part to a diminishment of unfavorable electrostatic

interactions between the negatively charged triphosphate tail and the Walker B motif in the mutant. Including an excess of Mg^{2+} (3 mM) in the presence of ATP affects the T_m for the two variants differently. Mg^{2+} in the presence of ATP stabilizes GhβRca* by ~ 5.3°C relative to ATP alone (Table 2) and contrasts with the mild destabilization ~ 0.2°C observed for the D173N variant under that same conditions.

To characterize the binding competence of the ATP analogues ATPγS and AMP-PNP, we collected Thermofluor data on GhβRca* with increasing concentrations of these nucleotides in both the absence and presence of 5 mM Mg^{2+}. Increasing the concentration of ATPγS and AMP-PNP both afforded increasing stabilization to GhβRca* (Table 2) indicating that both nucleotides bind, but the stabilizing effect of ATPγS was more pronounced - 44.4 ± 0.2°C at 2 mM compared to 39.5 ± 0.3°C for AMP-PNP at the same concentration. The apparent T_m of GhβRca* in ATP□S was slightly lower than for ADP (46.0 ± 0.2°C), but appreciably larger than for ATP (40.3 ± 0.2°C) suggesting that ATPγS binds more tightly than ATP. In contrast to previous results with ADP and ATP [101], Mg^{2+} had relatively minor effects on the thermal stability of GhβRca* in the presence of ATPγS. Differences in apparent T_m with and without Mg^{2+} for the ATPγS samples are essentially negligible, while Mg^{2+} in the presence of AMP-PNP destabilizes GhβRca* by ~ 1.4°C relative to AMP-PNP alone (Table 2).

Table 2. Thermofluor data for GhβRca* with various nucleotides

ATPγS [mM]	AMP-PNP [mM]	Mg^{2+} [mM]	ATP [mM]	ADP [mM]	Apparent T_m (°C)	
					GhβRca*	D173N
					37.0±0.6	34.5±0.4
		5			35.7±0.4	33.6±0.8
				2	46.0±0.2	45.4±0.1
		3		2	42.7±0.2	42.3±0.0

			2		40.3±0.5	44.0±0.2
		3	2		45.6±0.2	43.8±0.3
0.01					37.8±0.3	n/a
0.1					39.3±0.3	n/a
1					43.1±0.2	n/a
2					44.4±0.2	n/a
4					45.4±0.1	n/a
0.01		5			36.5±0.0	n/a
0.1		5			38.9±0.2	n/a
1		5			43.4±0.2	n/a
2		5			44.5±0.0	n/a
4		5			45.7±0.2	n/a
	0.01				36.7±0.3	n/a
	0.1				36.9±0.2	n/a
	1				38.5±0.4	n/a
	2				39.5±0.2	n/a
	4				40.6±0.3	n/a
	0.01	5			35.5±0.0	n/a
	0.1	5			35.6±0.1	n/a
	1	5			37.1±0.2	n/a
	2	5			38.1±0.2	n/a
	4	5			39.1±0.3	n/a

Analysis of ATPase activity in GhβRca* and its D173N variant

The effectiveness of the D173N mutation of the Walker B catalytic aspartic acid in eliminating ATPase activity from GhβRca* was tested by collecting room temperature turnover data at two different ATP concentrations for both GhβRca* and the D173N variant. Under all conditions tested, the D173N variant exhibit diminished ATPase activity relative to GhβRca* (Table 3), ranging from about 4-fold to 13-fold lower. As a negative control, BSA from Sigma was test as shown to have about 15-fold lower ATPase activity than GhβRca* assayed under the same conditions.

ATP and ATP□S hydrolysis activities of GhβRca* D173N and GhβRca* on ice

To test the effect on ice incubation for periods up to several hours has on the resulting ratios of nucleotide triphosphate to nucleotide diphosphate, we used end point analyses to determine the final phosphate content in GhβRca* D173N samples with Mg•ATP or ADP content in GhβRca* samples with Mg•ATPγS as a function of incubation time. Tables S2 and S3 show that the ratios of ATP to ADP and ATPγS to ADP, as respectively estimated from the Enzcheck phosphate release assay and NADH enzyme-linked ADP assay, change significantly over the course of five hours of incubation on ice. Notably, despite the fact that measures were taken to minimize the amount of ADP initially present in each sample, upon thawing and short (5-7 minute) incubation on ice, the calculated ratios are 6.8 for ATP:ADP in GhβRca* D173N and 11.2 for ATPγS:ADP in GhβRca*. These values are unexpectedly low for samples that nominally should contain entirely nucleotide triphosphates based on preparation method. These lower than expected ratios are likely attributable in part to contamination inherent to the commercially available nucleotide samples and in part to incomplete interchange between nucleotides during buffer exchange steps. Furthermore as incubation times are increased, these ratios decreased. The rate of change in calculated ATP:ADP or ATPγS:ADP ratios is more pronounced early in the time course and becomes less significant with increased incubation times (Tables 3 and 4). The results are consistent with the presence of ATP and ATPγS hydrolyzing activity in the GhβRca* D173N and GhβRca* samples, respectively. In both cases, the activity is relatively fast initial before

leveling off, likely reflecting lower initial concentrations of the competitive inhibitory ADP, which increases in concentration as hydrolysis proceeds.

Table 3. ATPase activity comparison of GhβRca* and GhβRca* D173N

	GhβRca* ATPase Turnover (min^{-1})	**GhβRca* D173N ATPase Turnover (min^{-1})**	**BSA ATPase Turnover (min^{-1})**
4 mM ATP*	3.8 ± 0.1	0.52 ± 0.28	n/a
0.5 mM ATP	2.8 ± 0.6	0.21 ± 0.15	n/a
4 mM ATP#	3.7 ± 1.9	0.98 ± 0.33	0.25 ± 0.18

*Samples included 3 mM "free" Mg^{2+} and 50 μM ADP. #Samples included 5 mM "free" Mg^{2+} and 65 μM ADP.

Table 4. ATPase activity of GhβRca* D173N samples on ice

Time (min)	**ATP:ADP**	**ATP (mM)**	**ADP (mM)**
7	6.76	1.74	0.26
20	5.24 ± 1.52	1.68	0.32
175	2.41	1.41	0.59
300	2.16 ± 0.57	1.37	0.63

Table 5. ATPγS hydrolysis activity of GhβRca* samples on ice

Time (min)	**ATPγS:ADP**	**ATPγS (mM)**	**ADP (mM)**
5	11.15	1.74	0.26
20	9.37 ± 5.55	1.72	0.28
160	7.11 ± 1.62	1.67	0.33
300	5.51 ± 1.02	1.61	0.39

Comparison of GhβRca* and GhβRca* D173N assembly with ADP

The GhβRca* D173N variant was constructed to facilitate FCS studies of Rca in the presence of ATP. Mutation of the catalytic Asp173 in the Walker B motif is expected to diminish ATPase activity while having less effect of the nucleotide binding properties. ATPase activity assay and Thermofluor stability assays confirm both reduced ATPase activity and nucleotide binding activity for this variant. In general, the mutant protein is more difficult to work with, providing lower yields and precipitating more easily. This

necessitated some adjustments to the previously published protocol for Alexa dye conjugation in order to obtain efficiently labeled GhβRca* D173N preparations. To test whether the GhβRca* D173N was suitable for use in FCS experiments, assembly data were collected in the presence of ADP and compared with the previously published data from GhβRca* under identical conditions. A side-by-side comparison (Figure 20, top panels) shows that the D173N variant has a larger uncertainty than the GhβRca* parent protein as judged by standard deviation. Although the apparent diffusion times determined with the D173N variant have a larger experimental standard deviation than the values determined with GhβRca*, we could determine that the D_{app}/D_1 versus concentration plots of both variants show the same trend. Likewise, refined K_d values for the GhβRca* D173N in ADP experimental data (Figure 20A, top) gave numbers very similar to GhβRca* in ADP (bottom of Figure 20). This result suggests that the assembly mechanism previously published for GhβRca* with ADP [106] is not altered significantly in the D173N variant and that this variant can be used for investigating the ATP-bound state under conditions where ATP hydrolysis is diminished.

FCS studies of GhβRca* D173N assembly at variable ATP:ADP ratios

We collected FCS data on GhβRca* D173N samples prepared with different ATP:ADP ratios as a function of protein concentration to assess the effect of nucleotide ratio on assembly. All samples contained 2 mM total nucleotide and 5 mM $MgCl_2$ with pure ATP, 3:1 ATP:ADP or pure ADP. Unexpectedly, the GhβRca* D173N variant shows significant residual ATPase activity (Table 4), which is ongoing under the conditions used for FCS sample preparation. Therefore, an average ATP:ADP ratio and a

range of ATP:ADP ratios (Figure 21, bottom) in GhβRca* D173N FCS experiments collected from 50 nM to 8 μM were estimated based on experimentally measured ATPase activity (Table 4). Samples initially containing all ATP have an estimated ATP:ADP ratio of 3.6 that ranges from 3-6, those prepared at 3:1 contain an approximate ratio of 2.3 with a range of 2-3 and those prepared with ADP retain a ratio of 0. Despite the ongoing ATPase activity, a clear trend in the data is discernable based the ATP:ADP ratio at which the GhβRca* D173N samples were prepared. These differences are quantifiable with simulations of D_{app}/D_1 versus Rca concentration using refined K_d values for monomer to dimer ($K_{m\text{-}d}$), dimer to tetramer ($K_{d\text{-}t}$) and tetramer to hexamer ($K_{d\text{-}h}$) equilibria. Doing so, yields identical $K_{m\text{-}d}$ and $K_{d\text{-}t}$ values of 10 μM and 2 μM, respectively, for the GhβRca* D173N samples with the highest (3.6) and intermediate (2.3) ATP:ADP ratios (Figures 21A and 21B, top panel). For the pure ADP sample, optimized $K_{m\text{-}d}$ and $K_{d\text{-}t}$ values of 6 μM and 0.5 μM point to an increased propensity towards dimerization and subsequent tetramerization than when ATP is present. The most readily apparent difference when comparing samples prepared at the highest ATP:ADP ratio to those containing lower ratios is the pronounced drop in the measured D_{app}/D_1 values at Rca concentrations above 1 μM (Figure 21, top panels). Adjustment of $K_{t\text{-}h}$ facilitates modeling of this drop in D_{app}/D_1 and reveals the trend of increasingly tight hexamerization with increasing ATP:ADP. Modeled $K_{t\text{-}h}$ values of 0.05 μM, 0.2 μM and 0.5 μM for the highest, intermediate and lowest ATP:ADP ratios, respectively, were deemed consistent with experimental data. Based on these optimized K_d values, calculation of the fractional concentration of oligomeric species as a function of GhβRca* D173N reveals a trend where increased ATP:ADP ratio promotes hexamer formation,

which is consistent with our findings for the GhβRca* construct in the presence of ATPγS or ADP.

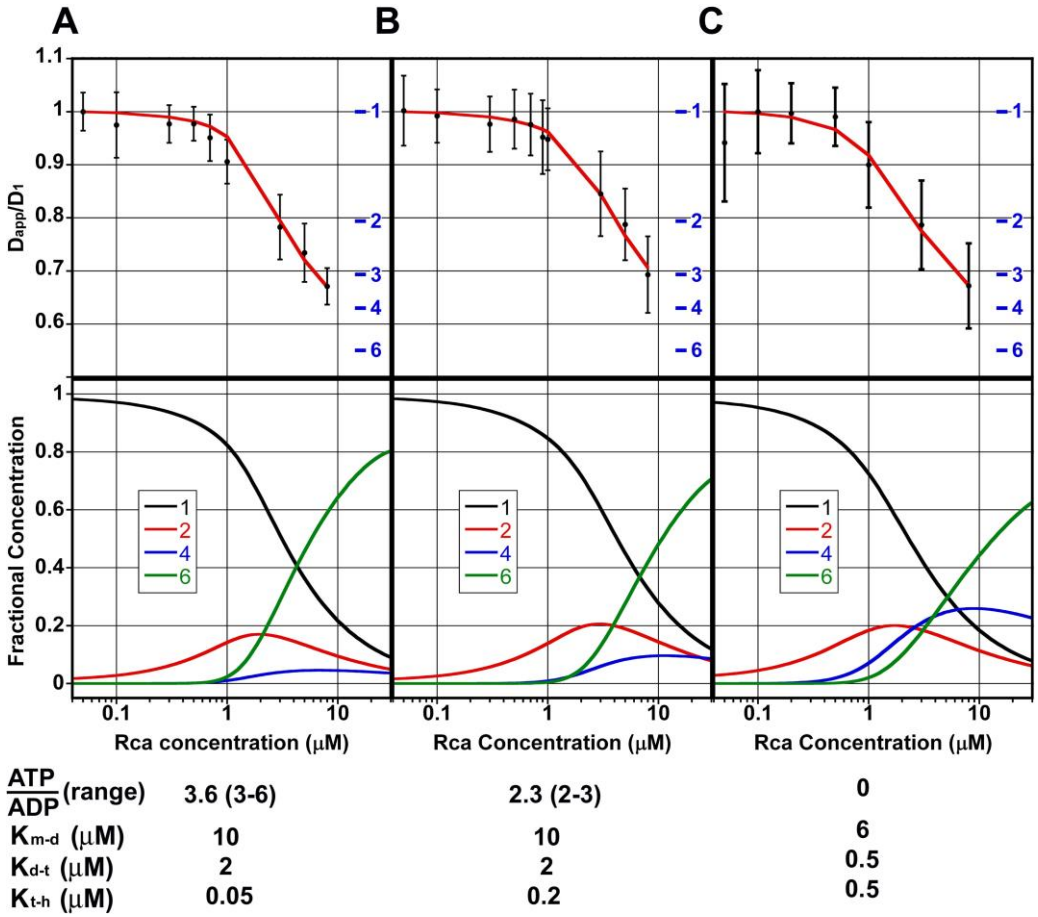


Figure 21: GhβRca* D173N assembly mechanisms at different ATP:ADP ratios: Each sample contained 25 mM HEPES-NaOH pH 7.6, 250 mM KCl, 5 mM $MgCl_2$, 10% glycerol and 2 mM total nucleotide prepared at ratios of all ATP (A), 3:1 ATP:ADP (B) and all ADP (C). Top panels: experimental data plotted as D_{app}/D_1 and presented as described above for Figure 19. Bottom panels: the fractional concentrations of oligomeric species as they appear in each model calculated from the total Rca concentration and the K_d values in the table below. Note, estimated average ATP:ADP ratios of 3.6 and 2.3 with ranges of 3-6 and 2-3 are reported due to slow hydrolysis during the on ice incubation of sample described in the text.

Nucleotide titration of GhβRca* D173N

To further examine the effect of nucleotide ratio on assembly, we used FCS to measure the apparent diffusion coefficients (D_{app}) of GhβRca* D173N samples at 60 μM

Rca in the presence of five different ATP:ADP ratios: all ATP, 3:1, 1:1, 1:3 and pure ADP. Each sample contained 2 mM total nucleotide and 5 mM $MgCl_2$. Figure 22 shows the experimentally measured D_{app} as well as the normalized D_{app}/D_1 values plotted against the ATP:ADP ratio for GhβRca* D173N. A gradual decrease in D_{app} is observed with decreasing ATP:ADP ratio (Figure 22), indicating that GhβRca* D173N at 60 μM forms smaller assemblies, with approximate sizes between 6-12 subunits, when the ratio is high. This finding agrees with the concentration-dependent studies of GhβRca* and GhβRca* D173N where a propensity towards formation of smaller oligomers (possibly hexamers) was found at high ratios of adenine nucleotide triphosphate to diphosphate (Figures 19 and 21).

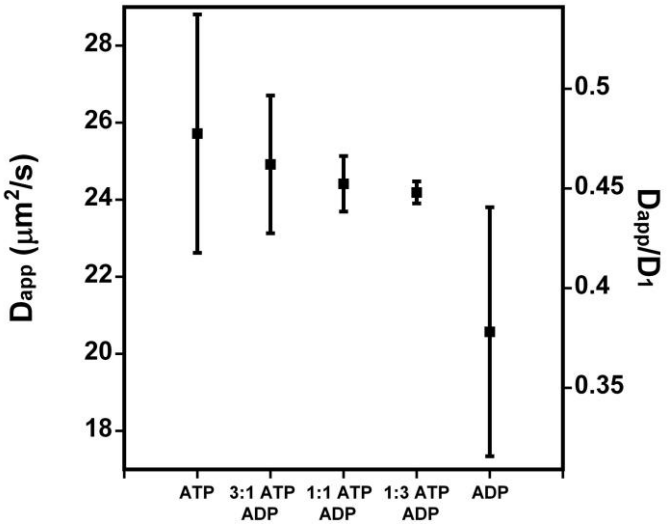


Figure 22: Nucleotide titration results: The apparent diffusion coefficients (D_{app}, *black squares*) measured for 60μM mutant (D173N) Rca under different nucleotide conditions as outlined in the 'x' axis of the graph above. The error bars represent the experimental standard deviations of at least three identical determinations. The relative apparent diffusion coefficients (D_{app}/D_1) scale is provided in the right hand Y-axis.

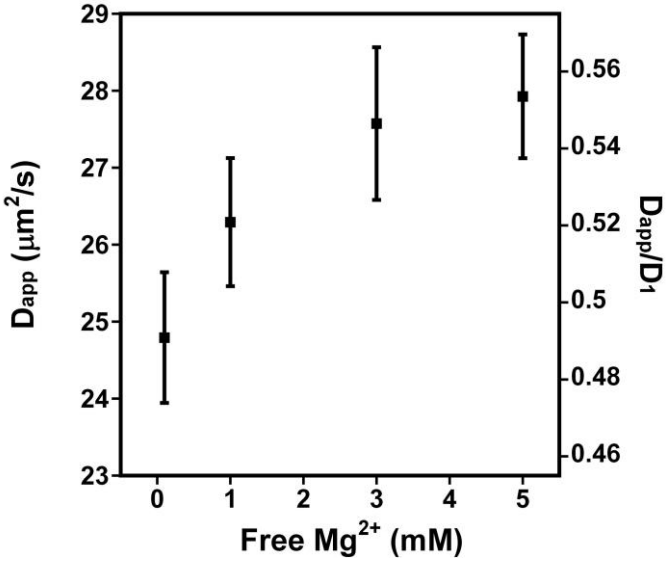


Figure 23: Mg^{+2} titration results: The measured apparent diffusion coefficient (D_{app}) of 45μM Rca (*44.95μM unlabeled +0.05μM labeled protein*) in presence of 1.5mM ATP, 0.5mM ADP, and increasing concentration of Mg^{+2} ion. The error bars represent standard deviation of experimental determinations.

Effect of free Mg^{2+} on GhβRca* D173N assembly

In order to assess how free Mg^{2+} affects the assembly process, we used FCS to measure the D_{app} values of 45 μM GhβRca* D173N samples in the presence of a nominal 3:1 ratio of ATP:ADP (2 mM total nucleotide) while varying the concentration of free Mg^{2+} from 100 μM to 5 mM. Figure 23 shows D_{app} versus the free Mg^{2+} concentration and reveals the strong influence that Mg^{2+} ions have on the size of Rca assemblies. As free Mg^{2+} increases from 100 μM to 5 mM, D_{app} increases from 24.8 $\mu m^2/s$ to 27.9 $\mu Mm^2/s$. When normalized to the diffusion coefficient of the monomeric species (50.4 μ^2/s), D_{app}/D_1 is around 0.49 for the low free Mg^{2+} and 0.55 for the high free Mg^{2+} condition. By comparison the predicted D_{app}/D_1 for the hexamer is about 0.55. Therefore, the free Mg^{2+} titration data indicate that, under the given conditions, low free Mg^{2+} favors the formation of larger, possibly aggregated species, and that increasing the free Mg^{2+} to

5 mM results in assembly to smaller oligomeric species with D_{app}/D_1 values consistent with that of a hexamer. This difference in assembly suggests a role played by Mg^{2+} in regulation of Rca activity that goes beyond serving as a cofactor to facilitate ATP binding and hydrolysis.

Dynamic Light Scattering

To test the effects of incubation time and temperature on the GhβRca* assembly state, dynamic light scattering (DLS) data were collected for samples incubated either at room temperature or on ice for various lengths of time with data collection at either 4°C or 25°C. In general, the Dynamics 7.0.3.12 software fits autocorrelation functions derived from fluctuations in scattered light intensity of GhβRca* samples best using multiple exponential time constants. These regularization fits [114] give multimodal distributions containing discrete polydisperse peaks. However, given the lack of reproducibility in these size estimates, coupled with contrary evidence from size exclusion chromatography and FCS suggesting continuous assembly, the intermediates of which are not resolvable by regularization analysis, we report the single exponential cumulant fits [115] for DLS data from GhβRca* and variants. The DLS data in Table 6 show little variability in the mean hydrodynamic radii (R_h) under the majority of conditions, ranging from about 13.2 to 14.1 nm with a relatively large (~25%) standard deviation reflecting the polydispersity of the distributions. An exception is the sample incubated on ice for 30 minutes before data collection at 25°C, which exhibited a significantly larger R_h around 17.1 nm with similar polydispersity.

Table 6. DLS data for GhβRca*

Sample	Radius (nm)	%Pd	MW-R (kDa)
Gh□Rca* 4°C, after 10 min on ice	13.150	26.6	1403
Gh□Rca* 4°C, after 10 min on ice	14.124	25.6	1656
Gh□Rca* 25°C, after 30 min on ice	17.013	24.6	2687
Gh□Rca* 25°C, after 5 min in RT	13.739	23.4	1550
Gh□Rca* 25°C, after 10 min at RT	13.816	23.9	1569

Discussion

By nature, multisubunit macromolecular assemblies are more complex than their monomeric counterparts. It is therefore not surprising that complex biochemical tasks required for cellular homeostasis are often carried out by macromolecular assemblies, which evolution favors for reasons including error control and regulation [116, 117]. In self-associating proteins, protomer-protomer interactions and protomer-substrate interactions can be symmetric or asymmetric. As such, understanding subunit cooperativity and biological function often requires careful experimental design and data interpretation When a protein forms a continuum of self-associated states with subunits that undergo rapid exchange, such as Rca, it becomes increasingly difficult to relate functional data to subunit stoichiometry and the fractional concentration of active species, which may be difficult to quantitate and/or unknown. This work builds upon the previously reported FCS-based method used to study GhβRca* assembly in ADP [106], here characterizing the effect of ATP, ATPγS, different ATP:ADP ratios and free Mg^{2+}. Unique to this method is the framework developed for modeling thermodynamics of self-association. With this method, we are able to interpret FCS data in terms of assembly

models with specific intermediates for which there are independent experimental evidence, particular dimeric and hexameric species.

In order to properly investigate the ATP-bound state of an enzyme with ATPase activity, it is often necessary to ensure that hydrolysis is minimized. Common approaches involve the use of nonhydrolyzable ATP analogue and mutagenesis of the conserved Walker B motif acid residue that is required for catalysis, but not for nucleotide binding. We have used both strategies to characterize Rca samples for use in FCS experiments. Thermofluor assays suggest that nucleotides bind with decreasing affinity to GhβRca* as follows: ADP > ATPγS > ATP > AMP-PNP (Table 2). These results are consistent with previous reports demonstrating that Rca binds tighter to ADP than to ATP in the absence of Mg^{2+} and that Mg^{2+} increases the affinity for ATP [101, 118, 119] while decreasing the affinity for ADP [101]. The observation that Mg^{2+} had essentially no effect on stability in the presence of ATPγS and slightly destabilized the AMP-PNP bound state gives some reason for pause when considering these analogues as potential mimics of the Mg•ATP-bound state. Apparent T_m values from Thermofluor assays indicate that 2 mM AMP-PNP provides a modest stabilization of ~ 2.5°C (Table 2) and therefore is likely weakly bound. In the absence of structural information, it is difficult to ascribe a reason for the low affinity GhβRca* displays for AMP-PNP; however, the presence of a bridging nitrogen has been shown to promote side chain rearrangements in the nucleotide binding pocket of myosin [120]. Our attempts to prepare Alexa-labeled GhβRca* in the presence of AMP-PNP resulted in double labeling (data not shown) indicating partial unfolding to expose one or more buried cysteine residues, which made us conclude that this nonhydrolyzable ATP analogue is not appropriate for studying Gh□Rca* assembly by FCS. Unlike AMP-

PNP, labeled ATP□S samples incorporated one molecule of Alexa dye per GhβRca* subunit, which is consistent with the notion that ATP□S supports a folded state that is significantly more thermally stable (~ 4.9°C at 2 mM nucleotide) than AMP-PNP. Despite this increased stability, our activity studies show that, like other ATPases [111, 120], GhβRca* is able to catalyze slow hydrolysis of ATPγS to ADP (Table 5). The hydrolysis is slow enough that ATPγS is a practical analogue for studying the effect of ATP on GhβRca* assembly, however, on-going hydrolysis means that the sample is not completely ADP-free. Therefore, we have enzymatically assayed ADP production from GhβRca* samples prepared in ATPγS and estimated the average and range of ATPγS:ADP ratios (Figure 19).

To study the effect of ATP binding on Rca assembly by FCS, an isosteric mutation of GhβRca* Asp173 to Asn was designed that binds ATP and assembles like wild-type, but lacks ATPase activity. This GhβRca* D173N variant, while slightly destabilized relative to the parent protein, nonetheless retains the ability to bind ATP and ADP (Table 2). In a previous study, van de Loo and Salvucci [119] made mutants of the homologous Asp174 in NtβRca including D174Q, D174A and D174E, but not D174N and measured the binding affinity by ANS fluorescence. Similar to their findings for D174Q and D174A, we observe that the affinity for ADP is similar between GhβRca* and GhβRca* D173N. However, GhβRca* D173N appears to bind ATP substantially tighter than GhβRca* as judged by the 3.7°C increase in apparent T_m (Table 2) for the mutant in the presence of 2 mM ATP, which also agrees with the NtβRca D174Q and D174A results [119]. The apparent T_m for GhβRca* D173N in 2 mM ATP (44.0 ± 0.2°C) is very close to the T_m for GhβRca* in 2 mM ATPγS (44.4 ± 0.2°C), indicating similar nucleotide

affinities and potentially related conformational states. In addition, unlike GhβRca* or NtβRca [119], free Mg^{2+} has no stabilizing effect on GhβRca* D173N or the NtβRca D174Q and D174A variants in the presence of ATP (Table 2), which is the same result seen for GhβRca* in the presence of ATPγS. Interestingly, while no ATPase activity was reported for the NtβRca D174Q and D174A variants [119], we have observed residual ATPase activity in the GhβRca* D173N variant (Tables 3 and 4), which is actually faster than the rate of ATPγS hydrolysis catalyzed by GhβRca* (Tables 4 and 5). Taken as a whole, these results suggest that the GhβRca* in ATPγS and Gh□Rca* D173N in ATP represent related assemblies of nucleotide-bound Rca that are amenable to study via FCS.

We have monitored the Mg•ATPγS-mediated assembly of GhβRca* at pH 7.6 using the slowly hydrolyzed ATP analogue ATPγS. Consistent with the Mg•ADP study [106], below 500 nM our data indicate that the monomer is the primary species in solution. By contrast, Stotz et al. [21] observed hexameric ring structures in negative-stain EM images of NtβRca R294V at approximately 500 nM in the presence of Mg•ATP or Mg•ATPγS. No EM images were reported for the parent protein (NtβRca) under these conditions and presumably hexameric rings are not formed, suggesting that the R294V substitution is required to form such structures at 500 nM Rca. We introduced the corresponding R293V mutation into the GhβRca* construct with the aim of using this variant as a background for FCS studies of Rca assembly. Interestingly, we found that this mutation substantially destabilized GhβRca* with Thermofluor assays indicating partially denatured enzyme that was incapable of nucleotide binding and therefore did not pursue characterization of this variant.

Our data show that at concentrations of GhβRca* above 700 nM and below 30 μM samples with high ratios of Mg•ATPγS:Mg•ADP (7-11) promote quicker assembly to higher order oligomers than samples containing only Mg•ADP (Figure 19*). Furthermore, above 30 μM GhβRca*, high Mg•ATPγS appears to partially arrest further oligomerization in comparison to Mg•ADP alone, where continued assembly appears facilitated. As described previously [106], here we have tested our data for agreement with two models of GhβRca* assembly in the presence of high Mg•ATPγS, one involving monomeric, hexameric and large (24 subunit) components and the second containing monomeric, dimeric, tetrameric, hexameric and large (24 subunit) components. In line with the earlier study of Mg•ADP-mediated GhβRca* assembly [106], only the model involving intermediates between monomers and hexamers was adequate to describe the data. In contrast, small angle x-ray scattering (SAXS) data collected for apo NtβRca were used to construct *in initio* models describing an equilibrium between monomer (52%) and hexamer (48%) at a protein concentration of 2.4 μM [27], yet models involving dimeric and tetrameric intermediates fit the data poorly.

Between 1-4 μM data are consistent with a model where roughly 15-18% and 26-28% of GhβRca* exists as dimers in the presence of Mg•ATPγS and Mg•ADP, respectively. At higher GhβRca* concentrations from 8 μM and 70 μM, the data are consistent with roughly 60-80% GhβRca* hexamers in high Mg•ATPγS, compared to about 30-40% hexamer supported with Mg•ADP alone (Figure 19). The similar $K_{m\text{-}d}$ (6 μM versus 3.5 μM) and $K_{d\text{-}t}$ (2 μM versus 1 μM) in the presence of high Mg•ATPγS versus Mg•ADP alone reflects the relatively small differences in the D_{app}/D_1 ratio for both samples at protein concentrations between 50 nM and 30 μM compared to the larger

D_{app}/D_1 differences above 30 μM. Indeed, iterative cycles of data simulation with the $K_{m\text{-}d}$ and $K_{d\text{-}t}$ values suggest that these are highly correlated. By contrast, large differences in the D_{app}/D_1 ratios at protein concentrations are observed above 30 μM and are manifested in the $K_{t\text{-}h}$ and $K_{h\text{-}24}$ values, $K_{t\text{-}h}$ = 0.1 μM and 1 μM while $K_{h\text{-}24}$ = 4000 μM^3 and 25 μM^3 for high Mg•ATPγS and Mg•ADP alone, respectively. It is important to note that the $K_{t\text{-}h}$ and $K_{h\text{-}24}$ values for this model are not highly correlated as judged by the difficultly in compensating for poor agreement between experimental and simulated data using a given $K_{t\text{-}h}$ thorough the systematic adjustment of the $K_{h\text{-}24}$ value and vice versa. Therefore, our experimental data are well modeled by an assembly pathway whose discrepancies in optimized $K_{t\text{-}h}$ and $K_{h\text{-}24}$ values suggest promotion of a hexameric species (green curves in the bottom panel of Figure 19) in the presence of high Mg•ATPγS relative to Mg•ADP alone.

We also employed the Walker B D173N mutant of GhβRca* study ATP-mediated assembly under conditions of limited ATP hydrolysis. Interestingly, the GhβRca* D173N retains significant residual ATPase activity (Tables 4 and 5), a somewhat surprising result given the wealth of literature describing analogous mutations in homologous AAA+ proteins that block hydrolysis while retaining binding to ATP and the association with macromolecular partners [92, 109-112, 121, 122]. While the structural details of ATP binding to Rca remain to be elucidated, the apo structure of the NtβRca AAA+ module (PDB ID 3T15) and a hexameric ring model created using this structure in combination with EM data (PDB ID 3ZW6) [21] reveal the presence of a number of highly conserved acidic residues (Glu121, Asp176, Asp204, Asp232 and Asp242) which, upon nucleotide binding, might be within close enough proximity to electrostatically compensate for the

loss of negative charge in the GhβRca* D173N mutant and thus maintain a low level of ATPase activity. Unfortunately, the residual ATPase activity in the GhβRca* D173N variant was significant enough that concentration-dependent FCS assembly data were measured only up to 8 μM Rca to ensure that samples were not incubated too long on ice and that a reasonably high ATP:ADP ratio was present (Figure 21 and Table 4). Still the FCS data to 8 μM for GhβRca* D173N in the presence of high ratios of ATP:ADP are in agreement with observations from GhβRca* with high ATPγS:ADP ratios and are consistent with a model where Mg•ATP promotes a more rapid assembly to hexameric species (Figure 21). Optimized K_d values from experiments with GhβRca* D173N support models where at ATP:ADP ratios of 3.6, 2.3 and 0 (2 mM total nucleotide with 5 mM total Mg^{2+}) Rca hexamers comprise > 60%, ~ 50% and ~ 40% of the assembly states, respectively (Figure 21). In separate experiments at higher Gh□Rca* D173N concentrations (60 μM) samples prepared at five different ATP:ADP ratios were assayed by FCS (Figure 19). These experiments show that decreasing the ratio of ATP:ADP is accompanied by an incremental decrease in the D_{app}, which is consistent with the observation of facilitated higher-order oligomerization (eg. formation of 24-mers) seen at similar GhβRca* concentrations in the presence of increasing Mg•ADP (Figure 19).

As is the case for several AAA+ proteins, the importance of Mg^{2+} in the activity of Rca is believed to center on its acting as a co-substrate to facilitate ATP binding and hydrolysis. Indeed, published assays for Rca-catalyzed Rubisco reactivation and ATPase activity contain excess free Mg^{2+} for exactly this reason. However, the effect of free Mg^{2+} in modulation of enzymatic activity through promoting changes in assembly states is not well understood. Changes in the free Mg^{2+} concentration of the chloroplast stroma, where

Rubisco and Rca reside, are believed to occur in response to different light/dark illumination conditions [100, 123]. This raises the possibility that Rca activity is modulated in response to changes in the concentration of free Mg^{2+}, which is known to be essential for activating Rubisco [124]. We used FCS to examine the effect of increasing concentrations of free Mg^{2+} in the presence of a 3:1 ratio of ATP:ADP had on the assembly of GhβRca* D173N at 45 μM. The incremental increase in D_{app} observed with increasing free Mg^{2+} is consistent with excess free Mg^{2+} favoring specific conformations of Rca subunits that promote interactions leading to hexamer formation (Figure 23). Conversely, when the free Mg^{2+} concentration is low at the same GhβRca* D173N concentration (45 μM), assembly of Rca to species significantly larger than hexamers is facilitated (Figure 23), again underscoring a likely difference in protein conformation that is associated with the availability of free Mg^{2+}.

Based on FCS data from GhβRca* and GhβRca* D173N collected under variable nucleotide conditions, we propose a revised model for concentration-dependent assembly of Rca in vitro that is consistent with all our observations (Figure 21). The model consists of six distinct assembly states and four dissociation constants. Starting from predominantly monomers at Rca concentrations below 300 nM, ADP and ATP promote dimerization (Figure 21, top) equally well with an average dissociation constant $K_{m\text{-}d}$ = 7 μM across all simulations. Self-association of Rca dimers into tetramers likewise is independent of the nucleotide-bound state with an average dissociation constant ($K_{d\text{-}t}$) for this step in around 1 μM. $K_{t\text{-}h}$ describes the equilibrium association of Rca tetramers with dimers to yield hexamers, and unlike the first two steps, this step is strongly dependent on the nucleotide bound state. The rapid drop in D_{app}/D_1 over the intermediate range of Rca

concentration in the presence of high ratios of ATPγS:ADP or ATP:ADP (Figures 16 and 18) is reproduced only in simulations where the $K_{t\text{-}h}$ value is set significantly lower than for the corresponding ADP data [106], amounting to a 10-fold difference in $K_{t\text{-}h}$ for ATP-bound (0.1 μM) and ADP-bound states (1 μM). This result suggests that there are nucleotide-dependent differences in the way dimers are added to tetramers as Rca forms hexamers. Our model interprets this difference as a reflection of ATP-facilitate formation of a closed ring hexamer versus the ADP-facilitated formation of an open hexameric spiral (Figure 21, center). Indeed, the tighter $K_{t\text{-}h}$ value in the presence of ATP is consistent with the notion that formation of a closed ring should bury twice the exposed surface area and therefore is likely energetically more favorable than ADP-mediated formation of opened hexameric spirals, which is supported by the identical $K_{d\text{-}t}$ and $K_{t\text{-}h}$ values indicating energetically equivalent addition of dimers to dimer and dimer to tetramers. As the Rca concentration increases further, larger assemblies here modeled as 24-mers, form from self-association of hexameric complexes (Figure 21, bottom). Similar to the hexamerization step, finding good agreement between experiment and simulations required the use of vastly different $K_{h\text{-}24}$ values of 4000 μM^3 and 25 μM^3 (Figure 19) for data collected in ATPγS and ADP, respectively. This result suggests that large aggregate formation is blocked in the presence of ATPγS and facilitated in the presence of ADP, or looking from the opposite point of view, ATPγS stabilizes the hexameric structure while ADP does not. A reasonable interpretation of this effect in the context of our model is that, at certain Rca concentrations, closed ring hexamers and open hexameric spiral structures are in a dynamic equilibrium (Figure 24, left center) that is strongly influenced by the ratio ATP:ADP with high ATP favoring rings and high ADP favoring spirals.

Assembly to larger aggregates proceeds through self-association of hexameric spirals necessitating high Rca concentration in the presence of ATP to drive the accumulation of aggregates through mass action. In addition to our FCS experiments, this model incorporates and is consistent with knowledge of dimeric species directly detected by SE-HPLC [101], hexameric rings observed in negative stain EM images [21] and the six protomer per turn spiral structures seen by X-ray crystallography and EM [21]. It should be noted however that this model is by no means the only one capable of providing good agreement with our experimental data. In particular the data are likely well fit by models involving alternative or addition intermediates in the assembly from monomer to hexamer (i.e. trimeric and/or pentameric species). While we favor the assembly of hexamers to larger aggregates through a spiraling pathway, our data do not rule out the possibility of large assemblies formed from the stacking of hexameric ring on top of one another. Similarly, our data are insufficient to distinguish whether the mechanism of large aggregate (24-mers) formation involves self-association of pre-formed hexamers or incremental addition of smaller oligomers (e.g. dimers) to build larger complexes.

This methodology is thus far the only one that has allowed for a thermodynamic description of protein self-association in the highly polydisperse enzyme Rca in the form of modeled dissociation constants. From these values fractional concentrations of relevant oligomeric species can be estimated under variable solution conditions and that information can potentially be used to aid in the interpretation of functional data. Here we used the FCS method to demonstrate that high, physiologically relevant ratios of ATP:ADP facilitate the formation and stabilization of a likely hexameric species. Furthermore, we show that free Mg^{2+} plays a role in attenuating hexamer formation under

similar conditions. In the future we expect that continued application of this technique will allow a comprehensive analysis of the factors governing Rca assembly including modulation via interaction with Rubisco and that this knowledge will contribute to a complete understanding of the mechanistic enzymology behind Rca-catalyzed Rubisco reactivation.

Materials and Methods

Site-directed Mutagenesis

A modified version of the QuickChange method (Stratagene, Lo Jolla, CA) was used to make a D173N substitution into a *gossypium hirsutum* (cotton) □-isoform Rca variant harboring a C-terminal Ala-Cys insert, hereafter referred to as Gh□Rca*, previously cloned into a pET151-dTOPO expression plasmid [106]. The following forward and reverse oligo nucleotide primers were ordered from Integrated DNA Technologies (Coralville, IA): 5'-CTCTTCATCAACAATCTCGACGCTGGAG-3' and 5'-CTCCAGCGTCGAGATTGTTGATGAAGAG-3'. PfuTurbo DNA polymerase and 10X Cloned Pfu DNA polymerase reaction buffer were purchased from Agilent Technologies (Santa Clara, CA), dNTPs from Life Technologies (Grand Island, NY) and DpnI restriction enzyme from New England Biolabs (Ipswich, MA). Thermal cycling, DpnI digestion and heat shock transformation into chemically competent Top10 *E.coli* was carried out according to the QuickChange Site-directed Mutagenesis guidelines. Single colony transformants from overnight LB (Luria-Bertani) agar plates containing 100 □g/mL carbenicillin were cultured in 25 mL of LB with 100 □g/mL carbenicillin

and grown overnight at 37°C and 250 rpm. After centrifugation (5000 rpm, 4°C, 10 minutes), plasmid was purified from overnight cultures using the centrifugal method from the Qiagen Plasmid Mini Prep Kit (Germantown, MD). The correct mutation was confirmed by DNA sequencing.

Expression, purification and dye conjugation of Rca variants

GhβRca* variants were recombinantly expressed from pET151-dTOPO or pET23 plasmids, with and without N-terminal 6XHis affinity tags, respectively, as previously described [106]. GhβRca* expressed without an affinity tag was purified according to methods published earlier [101]. Nucleotides and other reagents used were of the highest purity reasonably obtainable. Frozen cell pellets (-80°C) from 3 L of culture expressing GhβRca* or GhβRca* D173N were thawed on ice and suspended in 50 mL of 25 mM Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 10% glycerol, ~30 µM hen egg white lysozyme (Sigma-Aldrich), 1 mM phenylmethanesulfonylfluoride, 1 mM DTT, 0.1 mM ADP and 0.1 mM EDTA; then stirred at 4°C for 30 minutes and disrupted by sonication. Cell lysate was pelleted by centrifugation, and the supernatant passed through a 0.8 µm syringe filter before being loaded onto a Ni^{2+}-nitrilotriacetic acid (Ni-NTA) column (Qiagen, Valencia, CA). Protein was purified using an imidazole buffer step gradient (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT and 0.1 mM ADP, plus variable amounts of imidazole). After washing the columns with 10 mM and 80 mM imidazole buffer, His-tagged GhβRca* and GhβRca* D173N were eluted from the column with 35 mL of 200 mM imidazole buffer. The pooled elution fractions were digested with tobacco etch virus protease to remove the 6XHis tag, dialyzed overnight, pass over a second Ni-NTA column and concentrated to 2.5 mL in Centriprep concentrators (Millipore, Milford,

MA) as previously described [106]. Samples were then buffer exchanged into 25 mM HEPES-NaOH pH 7.5, 250 mM KCl, 2 mM nucleotide (ADP, ATP, ATPγS or AMP-PNP), 5 mM $MgCl_2$ (or no $MgCl_2$) and 10% glycerol using PD-10 columns (GE Healthcare). Protein concentrations were determined using the Bradford method, with typical yields of 3-4 mg/L cell culture and 1.6 mg/L cell culture for GhβRca* and GhβRca* D173N, respectively. GhβRca* and GhβRca* D173N conjugation with an ALEXA 546 C_5-maleimide fluorophore and analysis for efficiency of dye labeling followed the published method [106], except that ALEXA dye stock was prepared at 350 µM in 50 mM HEPES-NaOH pH 7.2 rather than at 1.93 mM in 50 mM Na_2HPO_4/NaH_2PO_4 pH 7.2.

HPLC, spectrophotometric and mass spectrometric analysis of labeled protein

Alexa-labeled GhβRca* and GhβRca* D173N samples were analyzed by reverse-phase HPLC (Agilent Technologies 1260 Infinity Quaternary LC system, Agilent Technologies 1100 Series Diode-Array detector) on a C18 analytical column (Phenomenex Prodigy, 5 µm ODS-3, 100 Å pore size, 250 x 4.6 mm) using a linear water/acetonitrile gradient with 0.1% trifluoroacetic acid (TFA). Protein was monitored by optical density (O.D.) 220 nm and O.D. 280 nm, and Alexa by O.D. 550 nm. Whereas free Alexa eluted at 40.5 min, all protein eluted at 36.5 min. Protein fractions were collected and the absorbance spectra collected using 50% acetonitrile/TFA as a blank. From these spectra, molar ratios for Alexa/protein were calculated using the previously described method [106]. To verify correct molecular mass of labeled samples, MALDI

spectra were collected on a Voyager DE STR mass spectrometer as previously described [106].

FCS sample preparation

In typical FCS experiments, 500 µL of sample buffer (25 mM HEPES pH 7.6, 250 mM KCl, 2 mM ATP, 5 mM $MgCl_2$, and 10% glycerol) was placed on ice and used for subsequent dilutions. GhβRca* or GhβRca* D173N samples for FCS were prepared in sample buffer at concentrations typically 85-90 µM for labeled and 110-135 µM for unlabeled protein prior to flash freezing with liquid nitrogen and storage at -80°C. Frozen stocks of labeled and unlabeled protein were thawed on ice. Once thawed, 1 µM and 10 µM of unlabeled protein were prepared on ice for use in subsequent dilutions. For each experiment, an appropriate amount of labeled and unlabeled GhβRca* or GhβRca* D173N were mixed in siliconized low-retention microcentrifuge tubes containing 1X buffer to give the desired final protein concentration. Regardless of total protein concentration, all samples were prepared so as to contain exactly 50 nM labeled protein. To allow subunit equilibration, samples were incubated for approximately 15 minutes on ice before performing experiments. ADP-containing samples were prepared in a similar manner.

For the nucleotide ratio experiments, thawed labeled and unlabeled GhβRca* D173N solutions containing ATP or ADP were first mixed to give the desired ratio then diluted to the appropriate concentration as mentioned above to prepare samples for FCS measurements. Buffer solutions used for nucleotide ratio experiments (Figures 1B) contained 25 mM HEPES pH 7.6, 250 mM KCl, 1.5 mM ATP, 0.5 mM ADP, 5 mM $MgCl_2$ and 10% glycerol; for low $MgCl_2$ experiment (Figures 1C) buffer contained 25

mM HEPES pH 7.6, 250 mM KCl, 1.5 mM ATP, 0.5 mM ADP, 2.02 mM $MgCl_2$ and 10% glycerol. All incubations were done on ice.

FCS sample preparation for Mg^{2+} titration experiments

A set of four 1X ATP-containing sample buffer solutions (see above) with variable total $MgCl_2$ concentrations (1.7, 3, 5 or 7 mM) and Gh□Rca* D173N at 45 μM were prepared by mixing labeled protein with unlabeled protein following a dilution protocol analogous to the one mentioned above. The concentration of free Mg^{2+} in solution was calculated assuming multiple equilibria and using the reported dissociation constants for ADP and ATP (ref) of 0.241 mM and 0.028 mM, respectively. By fixing the ADP and ATP concentrations, this allowed the concentration of free Mg^{2+} in solution to be adjusted simply by dilution the appropriate amounts of one of the four variable $MgCl_2$ buffers.

FCS sample preparation for Nucleotide titration experiments

Labeled and unlabeled stock solutions for this experiment were prepared by mixing appropriate amount of thawed cotton GhβRca* D173N in ATP or in ADP and following dilution protocols mentioned above. Each experiment contained a total of 60 μM GhβRca* D173N, 50 nM of which was labeled. Different dilution buffers were used to vary the nucleotide ratios in each sample. The buffers had all the ingredients in common except their nucleotide composition, which were as follows: 2 mM ATP, 1.5 mM ATP + 0.5 mM ADP, 1 mM ATP + 1 mM ADP, 0.5 mM ATP + 1.5 mM ADP, and 2 mM ADP. All samples were incubated on ice.

Fluorescence Fluctuation Spectroscopy data collection

FCS measurements were carried out using a home-built confocal optical setup. Excitation was achieved with a Compass 215M-10 532 nm CW laser (Coherent GmbH, Germany) attenuated to 50 μW to minimize triplet dynamics. The output of the laser was expanded, collimated, and directed, via a dichroic lens, into an Olympus PlanApo 100X/1.4NA Oil objective. Samples were placed into silicone perfusion chambers (Grace Biolabs, Bend, OR) pre-treated with BSA to minimize Rca adsorption onto the cover glass. Fluorescence from samples was collected via the objective, separated from excitation light through the dichroic, and reflected into a 70 μm pinhole. The emission was then focus into an avalanche photodiode detector (Perkin-Elmer Optoelectronics, SPCM-AQR14). A bandpass filter was used before the detector to minimize stray light (Omega 3RD560-620). The recorded fluorescence signal was autocorrelated in real time using an ALV7002 USB-25 correlator (ALV, Germany).

FCS data analysis

The autocorrelation function of a single species that diffuses freely in solution is given by:

$$G(\tau) = G_0\left(1+\frac{4D\tau}{r_0^2}\right)^{-1}\left(1+\frac{4D\tau}{z_0^2}\right)^{-1/2}$$

where τ is the correlation lag-time, G_0 is the amplitude of the decay, D is the diffusion coefficient of the diffusing particle, and r_0 and z_0 are the radial and axial semi-axes of the Gaussian confocal volume.

The instrument was calibrated daily by measuring the FCS decays of free TAMRA dye (D = 414 $\mu m^2 s^{-1}$) in buffer, and fitting the experimental decays using r_0 and

z_0 as fitting parameters. The diffusion coefficient of Rca was then determined by fitting the experimental FCS decays using the values of r_0 and z_0 measured on the same day. As shown in our previous work, solutions of Rca are in general polydisperse and therefore the diffusion coefficients obtained from the fits represent apparent values for a given mixture of oligomeric species. Data in Figures 1-3 are presented as normalized ratios D_{app}/D_1, where D_1 is the diffusion coefficient of the monomer measured with a 50 nM Rca solution.

GhβRca* in Adenosine 5'-(gamma-thiophosphate) (ATPγS)

Experiments were conducted exactly as with the D173N variant except that the 1X buffer contained 2 mM ATPγS instead of 2 mM ATP. For the time series experiment, 45μM GhβRca* was prepared by mixing labeled and unlabeled protein as described above. All incubations were done on ice.

Enzcheck Phosphate Assay to follow Rca ATPase activity

Inorganic phosphate (Pi) production was monitored with a UV 2401 PC Shimadzu spectrophotometer using the Enzcheck phosphate assay kit (Life Technologies) where the substrate 2-amino-6-marcapto-7-methyl purine riboside (MESG, I_{max} = 330 nm), in the presence of P_i, is converted by Purine nucleoside phosphorylase (PNP) to Ribose-1-phosphate and 2-amino-6-mercapto-7-methyl purine (I_{max} = 360 nm). ATPase activity is reflected by temporal changes in sample absorbance at 360 nm.

Standard curve

Reference and sample cells were auto-zeroed with 1X Enzcheck reaction buffer (50 mM Tris-HCl pH 7.5, 1 mM $MgCl_2$ and 0.1 mM NaN_3) supplemented with 55 mM HEPES pH 8, 1.6 mM DTT and 4 mM $MgCl_2$. MESG then PNP were added to 200 μM

and 3 U/mL, respectively and the mixture equilibrated for 30 seconds before starting the time scan. Once the signal was stable, known concentrations of P_i standard (KH_2PO_4) were added, mixed and time scans continued until saturation was reached. Each reaction volume was 1 mL and is reflected in concentrations given above. For each P_i concentration, initial and final 360 nm absorbance values were recorded and the change plotted as a function of [P_i] to give a standard curve, which was fit to a linear equation and used to calculate P_i release upon Rca-catalyzed ATP hydrolysis.

Comparison of ATPase activity in GhβRca* and its D173N variant

ATPase activity was assay with 5 μM Rca, 4 mM ATP (saturating) and 5 mM $MgCl_2$. Each condition was measured in triplicate. Substrate ATP for this assay was prepared in 100 mM HEPES pH 8 to avoid acidic hydrolysis. Reference and sample cell were auto-zeroed as described above for the P_i standard curve. Next 200 nmoles MESG, 5 nmoles Rca and 3 Units PNP were added sequentially and the mixture equilibrated for 1 minute to consume contaminating phosphate and achieved steady signal. The time scan was started and, after 1 minute, substrate ATP (4 μmoles) was added, the sample mixed and the time scan continued for ~4 minutes. P_i contamination from ATP gives a significant signal even in absence of Rca. Therefore, a blank experiment was performed where the Rca buffer (25mM HEPES, 300mM KCl, 1 mM DTT, 2mM ADP and 10% glycerol) was added instead of Rca. A control experiment with 5 μM bovine serum albumin (BSA, Sigma Aldrich) suspended in 25mM HEPES, 300mM KCl, 1mM DTT, 2mM ADP and 10% glycerol) was also performed. Buffer blank data were subtracted from Rca turnover data in the calculation of initial rates, using the initial steepest linear portion of the activity plots. Generally the time range for these calculations was 20-30

seconds. The absorbance value of the initial slopes were converted to the amount of P_i using the standard curve and normalized by dividing the Pi concentration by the concentration of Rca.

Analysis of ADP in GhβRca* D173N ATP samples through measurement of Pi

A phosphate measurement assay was used to quantify the P_i content of GhβRca* D173N samples under various conditions with a UV 2401 PC Shimadzu spectrophotometer. Frozen (-80°C) samples containing 113 μM GhβRca* D173N in 25 mM HEPES-NaOH pH 7.5, 250 mM KCl, 5 mM MgCl2, 10% glycerol and 2 mM ATP were thawed on ice alongside corresponding buffer samples. Once thawed, 22.1μL of GhβRca* D173N was added to a cuvette containing a final volume of 1 mL with 0.2 mM MESG, 3 Units of PNP, 50 mM Tris pH 8.0, 55 mM HEPES pH 8.0, 1 mM $MgCl_2$, 0.1 mM NaN_3. Blank samples were prepared by adding 22.1 μL of protein buffer instead. Initially the signal at 360 nm was observed for 60 seconds (scan speed = 1 sec^{-1}) to ensure the absorbance was increasing as expected. Samples were then maintained at room temperature for 240 seconds additional seconds, at which point the 60 sec time scans were recollected. These experiments were repeated, for GhβRca* D173N and buffer blank samples, following thawing and incubation for 10 minutes, 20 minutes, 40 minutes, 80 minutes, 120 minutes, 160 minutes, 200 minutes and 240 minutes on ice. For each time point following the 5 minute room temperature incubation, the average A_{360} over the 60 second scan for the GhβRca* D173N sample was subtracted from the blank sample and this absorbance difference converted to an amount of P_i using a standard curve. Since the

amount of P_i contamination is negligible for this assay the amount of total phosphate is equivalent to the amount of ADP.

Measuring the ADP content of ATP-γS incubated GhβRca* samples

A reduced nicotinamide adenine dinucleotide (NADH) enzyme-linked assay was used to quantify the ADP content of Gh□Rca* samples under various conditions with a UV 2401 PC Shimadzu spectrophotometer. Frozen (-80°C) samples containing 5 mg/mL Gh□Rca* in 25 mM HEPES-NaOH pH 7.5, 250 mM KCl, 5 mM $MgCl_2$, 10% glycerol and 2 mM ATPγS were thawed on ice alongside corresponding buffer samples. Once thawed, 25.5 µL of GhβRca* sample was added to a cuvette containing a final volume of 1 mL with 0.3 mM NADH, 2 mM phoshoenolpyruvate, 100 mM Tricine pH 8.0, 5 mM $MgCl_2$, 20 mM KCl, 4.8-8 U pyruvate kinase and 7.2-11.2 U lactate dehydrogenase. Blank samples were prepared by 25.5 µL of protein buffer instead. Initially the signal at 340 nm was observed for 60 seconds (scan speed = 1 sec^{-1}) to ensure the absorbance was decreasing as expected. Samples were then maintained at room temperature for 240 seconds additional seconds, at which point the 60 sec time scans were recollected. These experiments were repeated, for GhβRca* and buffer blank samples, following thawing and incubation for 20 minutes, 175 minutes and 300 minutes on ice. For each time point following the 5 minute room temperature incubation, the average Abs_{340} over the 60 second scan for the GhβRca* samples was subtracted from the blank samples and this absorbance difference converted to an amount of ADP using an extinction coefficient (e_{340}) for NADH of 6220 $M^{-1}cm^{-1}$ [125].

Dynamic Light Scattering (DLS)

Stocks of purified GhβRca* stored in 25mM HEPES pH 7.5, 250 mM KCl, 10% glycerol, 5mM $MgCl_2$, and 2mM ATP□S or 2 mM ATP (D173N variant) were thawed on ice then filtered through centrifugal spin columns with a 0.1 □m pore size (Milipore) by centrifugation at 4°C prior to DLS measurements. The filtrates were diluted to the desired concentration using filtered, ice-cooled protein buffer then transferred to disposable UVette cuvettes (Eppendorf, Hauppauge, NY) for data collection. DLS was collected with a DynaPro NanoStar (Wyatt Technologies, Santa Barbara, CA) and data analyzed using the software Dynamics 7.0.3.12. Scattered light intensity fluctuations were averaged from forty acquisitions (5 seconds each) and fit to both monomodal and multimodal distributions.

Thermofluor Stability Assays

Data were collected with ABI Prism 7900HT Sequence Detection System (Applied Biosystems). Each condition contained 20 µL of sample at 0.25 mg/mL Rca in 25 mM HEPES-NaOH pH 7.5, 150 mM KCl and 16X Sypro Orange (Invitrogen). GhβRca* and GhβRca* D173N premixes at 1.05X were made as described previously (BBA paper) and 19 µL pipetted into the wells of a 384-well polypropylene TempPlate PCR plate, which was kept on ice. Nucleotides (ADP, ATP, ATPγS or AMP-PNP) at 2 mM, $MgCl_2$ at 5 mM or both were introduced by addition of 1 µL stocks containing either 40 mM nucleotide, 100 mM $MgCl_2$ or both. Each condition was setup in triplicate. The plate was then sealed with optically clear tape and data collected and processed as previously described [101].

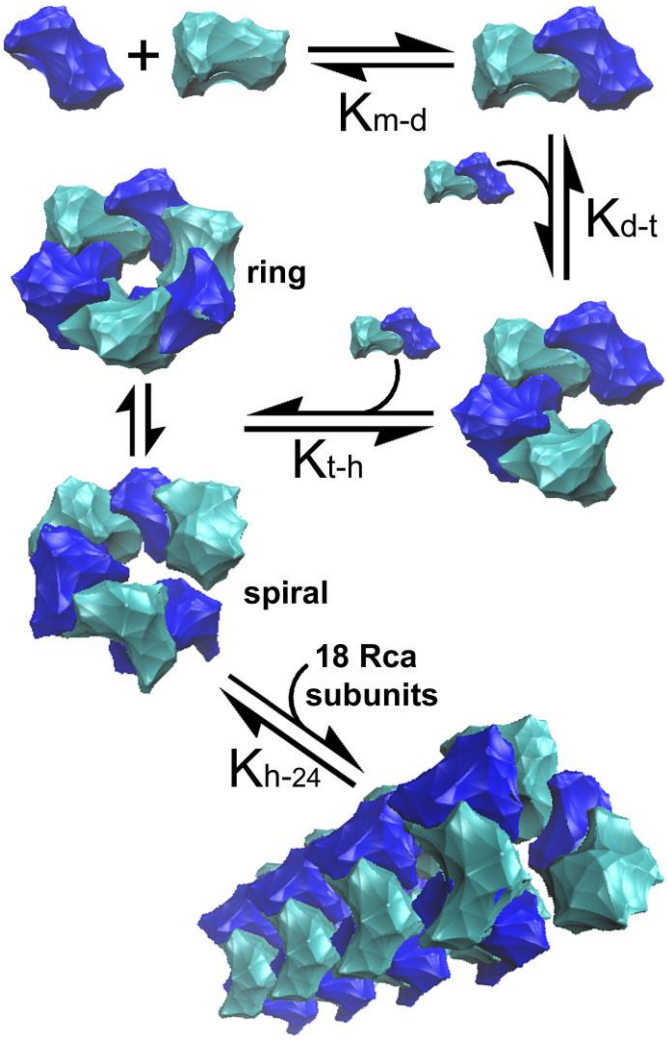


Figure 24: Schematic representation of the Rca assembly model consistent with experimental results: The model consists of a dimer and a tetramer as intermediates before the formation of a hexamer. Hexameric species are shown as an equilibrium distribution of closed ring and open spiral forms, which are presumably populated differentially depending on the particulars of nucleotide triphosphate or diphosphate binding.

CHAPTER 4

ADDITIONAL TOOLS FOR ASSESSING OLIGOMERIZATION OF RUBISCO ACTIVASE

Determination Of Self-Assembly Of Rca By Analytical Ultracentrifugation

Theory

Analytical ultracentrifugation (AUC) is one of the oldest and most powerful techniques for the quantitative characterization of macromolecules in solution. AUC has been extensively applied to the study of biomacromolecules in different solvents and concentrations. The observation of sedimentation in real time is available by three optical systems: absorbance, interference, and fluorescence. In AUC, the sedimentation occurs in free solution, without the need of surfaces, matrices, or attached labels. The molecular weights detected by AUC range from a few hundred Daltons to several hundred-million Daltons [126].

Sedimentation velocity and sedimentation equilibrium

There are two approaches to detecting or characterizing macromolecular interactions: sedimentation velocity (SV) and sedimentation equilibrium (SE).

Sedimentation velocity is a hydrodynamic method in which a high gravitational field is applied to a solution of the protein sample with real-time imaging. The separation of proteins is observed due to their different rates of migration in the centrifugal field.

Mathematical analysis is used to study the concentration gradients formed from the centrifugal force. As a result, the sedimentation boundaries determine the separation of protein species based on size dependence, providing the user with a sedimentation

coefficient which is related to the molar mass of each sedimented species [126]. In SV experiments, a rapidly interconverting dynamic equilibrium of different assembly states yields a broad distribution of s-values. Whether static or dynamic, the integration of c(s) provides a weight-averaged sedimentation coefficient (s_w) for each experiment (SEDFIT). The composition- and concentration- dependence of s_w may be evaluated in terms of protein self-association models [127] .

Sedimentation equilibrium is a thermodynamic method where lower centrifugal speeds are used on solutions to identify molecular mass, stoichiometry, and association constants. In short, at small centrifugal forces the equilibrium concentration distribution of molecules is reached when the sedimentation is exactly balanced against the diffusion [128].

The availability of interference optics is essential for this work when UV- absorbing nucleotides are present in the sample buffer. Additionally, interference optics allow for the detection of a broader range of protein concentrations.

To estimate the thermodynamic binding constants for assembly equilibria, sedimentation equilibrium (SE) experiments are carried out at different protein loading concentrations and different velocities. Depending on the level of complexity, this approach may allow for the determination of accurate molar masses for each species. Under favorable circumstances, equilibrium experiments can be used to extract association constants by global non-linear regression fitting of multiple data sets [129]. The fit of quality of different models can be evaluated based on the distribution of residuals (the difference

between the experimental data and the theoretical curve) and the physical reasonableness of extracted values. A useful program for this approach is called SEDPHAT [129].

Methods

The experiments were performed at the University of Arizona in Tucson using analytical ultracentrifuge equipment, a Beckman Coulter XL-I with monochrometer and interference scanning optics. Rca-WT was expressed in E. Coli without an affinity tag, and purified by traditional procedures as described previously. For sedimentation velocity experiments the protein was prepared in 25mM HEPES pH 7.5, 250mM KCl, 10% glycerol, and 0.05mM ADP.

The sample protein and buffer solution (400 µl each) were loaded into the appropriate sides of a two-sector shaped cell. Sedimentation velocity experiments were carried out with one sector loaded with sample, the second being the reference sector that contains the solvent. After cooling the instrument to 4 °C for >1 h, the sample was spun at 40 000 rpm. Absorbance scans were taken continuously for 17 h. The data were analyzed with SEDFIT, allowing for a continuous $c(s)$ distribution [130]. The Table XXX presents the partial specific volume of the protein. The buffer density and viscosity were calculated from SEDNTERP or measured by the instrument (Table 7). Initial Rca concentrations were 1.85, 4.63, 10.88, 15.28, and 17.36 µM.

Table 7. Density and viscosity of buffer for sedimentation velocity experiments: 25mM HEPES pH 7.5, 250mM KCl, 10% glycerol, and 0.05mM ADP.

	Measured by a high-precision density meter	**SEDNTERP**
Density (g/mL)	1.03887	1.03946
Viscosity (P)	0.012233	0.012348

A partial specific volume (V-bar) at 25°C was calculated by SEDNTERP from amino acid sequence:

AKEIDEDTQTDQDRWKGLAYDISDDQQDITRGKGMVDSLFQAPMNDGTHYAV
MSSYEYISQGLRTYDLDNNMDGFYIAPAFMDKLVVHITKNYMTLPNIKVPLILGI
WGGKGQGKSFQCELVFAKMGINPIMMSAGELESGNAGEPAKLIRQRYREAADII
KKGKMCCLFINDLDAGAGRMGGTTQYTVNNQMVNATLMNIADNPTNVQLPGM
YNKEENPRVPIIVTGNDFSTLYAPLIRDGRMEKFYWAPTREDRIGVCTGIFRTDNV
PVDDIVKLVDTFPGQSIDFFGALRARVYDDEVRKWIGEVGVNSVGKKLVNSREG
PPSFEQPTMTIEKLLEYGNMLVAEQENVKRVQLADKYLSEAALGNANDDAIKRG

AF is **0.733793.**

The sedimentation equilibrium experiments were performed as above for sedimentation velocity with the following changes. Samples (110 µl) containing varied concentrations of apoprotein Rca-WT (0.7-1.4 mg/mL) dialyzed overnight in 25mM HEPES pH 7.5 and 250mM KCl were loaded into three chambers of a six chamber centerpiece, with buffer occupying the remaining three. The system was allowed to equilibrate for at least 1 hour after reaching the target temperature of 4°C. Samples were then subjected to centrifugal

speeds of 7000, 17000, and 28000 rpm for a period of almost three weeks. High-resolution absorbance and interference scans at 295, and 300 nm were taken every 4 hours, with a step size of 0.001 cm. The establishment of equilibrium was determined when a reasonable (<0.07) and constant RMSD was found between sequential scans. The data were analyzed using global and single fitting in SEDPHAT with two or three species. Density and partial specific volumes were determined as above.

Results

Sedimentation velocity

A sedimentation velocity experiment on the Rca-WT protein sample was run at 40 000 rpm for 17 hours at 4°C. Five different concentrations, 1.85, 4.63, 10.88, 15.28, and 17.36 μM, were chosen in order to study the protein size dependence on the concentration. Although the first experiment at 0.08 mg/mL had barely enough signal to be 'detectable', the higher concentrations gave good signal to determine Svedberg values. Figure 24 shows the sedimentation velocity results for the Rca-WT as a function of the sedimentation coefficient distribution ($c(s)$, y-axis) versus the Svedberg constant ([S], x-axis). Table 9 shows fitting the data that yields an apparent MW of 44, 84.1, 95.1, 110, 153 kDa at 1.85, 4.63, 10.88, 15.28, and 17.36 uM respectively. The molecular weight of a monomer is 42.5 kDa, therefore the species in the solution may primarily be monomers, dimers, or trimers / tetramers. Surprisingly, these MW values are inconsistent with published AUC SV data on Rca protein; measurements at similar Rca concentrations provided different molecular weights. The published values are almost double the size of MWs presented here. For example, at a concentration of 1.2 μM, the MW was 121 kDa,

4.8 µM corresponds to 225 kDa, 9.5 µM to 300 kDa, and 19 µM gave 310 kDa [27]. The peaks in figure 24 are not sharp, but broad and elongated, which indicates that there must be more than one species in solution. The determination of accurate diffusion coefficients and molar masses of the sedimenting species is only possible for highly homogeneous samples. Moreover, for heterogeneous solutions, the determination of D and s may not be possible due to the rapid association-dissociation kinetics of this protein. The possible interaction over time of different oligomeric species of Rca provides only apparent sedimentation and diffusion coefficients of coexisting species, which makes it difficult to provide a good description of measured values for each oligomeric component [131].

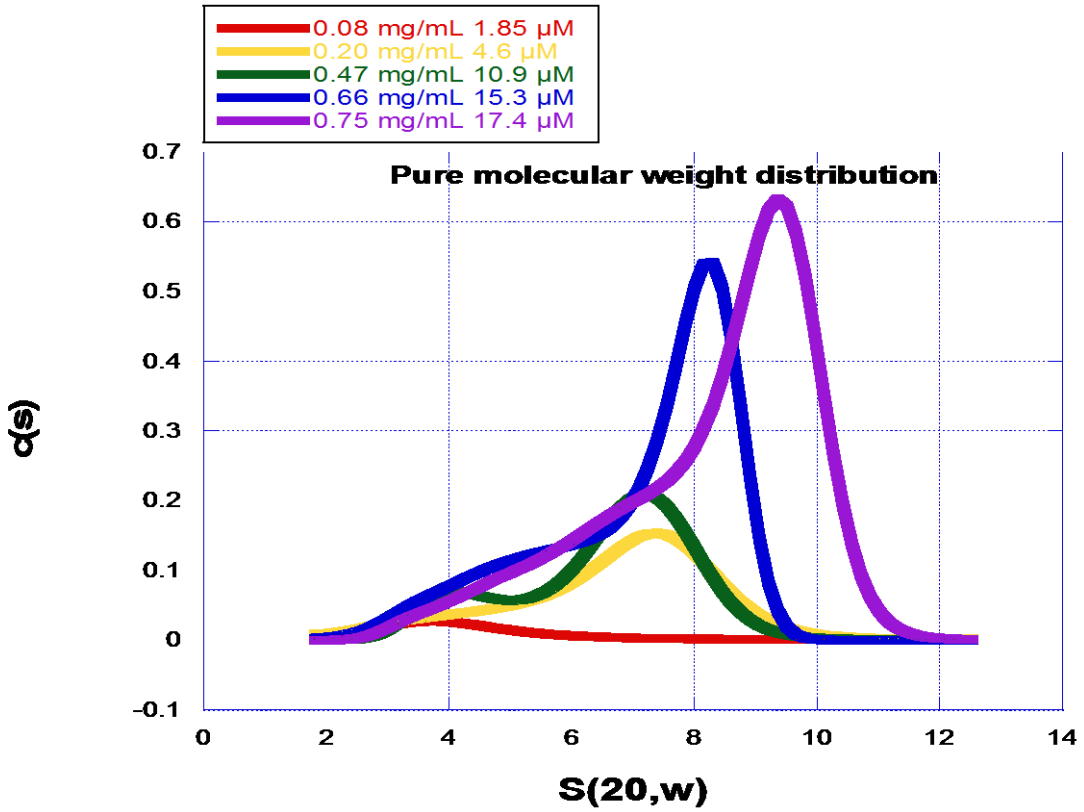


Figure 25. Sedimentation velocity on Rca-WT as a function of the sedimentation coefficient distribution ($c(s)$, y-axis) versus the Svedberg constant ([S], x-axis).

Table 8. The calculated molecular masses from sedimentation velocity experiments.

Concentration (mg/mL) µM	f/f0	MW (kDa)	$s_w(20,w)$	integration c(s) peak	Filename
0.08 1.85	1.0315	44.00	4.3130	0.07870	RAW-V008A-s20w.dist
0.20 4.63	1.0267	84.10	6.6730	0.52130	RAW-V009A-s20w.dist
0.47 10.88	1.0480	95.10	7.0950	0.61540	RAW-V010A-s20w.dist
0.66 15.27	1.1440	110.00	7.1500	1.2587	RAW-V011A-s20w.dist
0.75 17.36	1.2350	153.00	8.2810	1.8144	RAW-V012A-s20w.dist

Sedimentation equilibrium

Sedimentation equilibrium experiments were performed on apoprotein Rca-WT (0.7-1.4mg/mL) at three speeds: 7 000, 17 000, and 28 000 rpm. The runs were performed at 4°C and it took three weeks until the system reached its equilibrium. Data collected at 28 000 rpm were incomplete and noisy. To analyze the data collected at 7,000 and 17,000 rpm, we decided to try both a single fit and a global fit. Since we had knowledge about coexisting multispecies in solution, we agreed to fit the data to double or triple species and skip the single species fit. The absorbance wavelengths that gave the cleanest raw data were 295 and 300nm, along with interference optics data.

The global fit and single fit to triple species (absorbance 295 nm) using data from two speeds (7000 and 17ooo rpm) at concentrations of 1.4 mg/mL, 1mg/mL and 0.7mg/mL resulted in stoichiometries of around monomer/dimer and hexamer (Table 9). The fits to

triple species may be over-fitted because two species are dominant in the solutions, with a very minimal fraction of aggregates that give gigantic molecular weights. Nevertheless, by comparing AUC and Fluorescence Correlation Spectroscopy (FCS, described in chapter 2) data at these three concentrations, there is some consistency in the stoichiometries.

Table 9. Single and global fits for three species of Rca samples measured at absorbance 295nm.

A B S O R B A N C E 2 9 5 n m				
Single fit for 2 species			Global fit for 2 species and 2 speeds (7k+17k)	
Concentration (mg/mL) (µM)	MW (kDa)	Stoichiometry	MW (kDa)	Stoichiometry
0.7 16.5	234.0 53.9	5.5 1.2	550.0 98.7	12.9 2.3
1 23.5	245.3 53.3	5.8 1.2	441.0 76.7	10.4 1.8
1.4 33.0	240.2 101.3	5.6 2.4	513.3 86.1	12.0 2
Single fit for 3 species (7k+17k)			Global fit for 3 species and 2 speeds (7k+17k)	
Concentration (mg/mL) µM	MW (kDa)	Stoichiometry	MW (kDa)	Stoichiometry
0.7 16.5	239.6 47.5	5.6 1.1	197.4 49.1	4.6 1.1
1 23.5	243.0 53.3	5.7 1.2	303.6 67.7	7.1 1.6
1.4 33.0	239.5 125.9	5.6 2.9	293.2 83.4	6.9 2.0

Table 10. Summary of SE data on absorbance and interference optics for Rca.

	Absorbance 295 nm		Interference optics
Concentration (mg/mL) µM	Stoichiometry (single fit for 2 or 3 species)	Stoichiometry (global fit for 2 or 3 species)	Stoichiometry (single fit for 3 species)
0.7 16.5	5.55 and 1.15	12.9 and 2.3 4.6 and 1.1	18.3, 7.2, and 1.4
1 23.5	5.75 and 1.2	10.4 and 1.8 7.1 and 1.6	25.9, 7.7, and 1.5
1.4 33.0	5.6 and 2.6	12.0 and 2.0 6.9 and 2.0	

A single fit to three species on interference optics at protein concentrations of 0.7 and 1.0 mg/mL provided stoichiometries of monomer/dimer, heptamers, and some species in a range of low 20mers (Table 11). Although these results are very preliminary, they agree to some extent with the FCS Rca-WT•ADP "model 3" in which we can see species between monomer/dimer, tetramer/hexamer and 24mer.

Table 11. Interference optics data utilized to calculate species in a solution containing Rca samples

INTERFERENCE OPTICS		
Single fit for 3 species		
Concentration (mg/mL) µM	MW (kDa)	Stoichiometry
0.7 16.5	776.2 308.4 58.0	18.3 7.2 1.4
1 23.5	1,102.8 326.8 62.1	25.9 7.7 1.5

.

SUMMARY

In general, due to high polydispersity, long experimental runs, and very unstable Rca in the absence of any nucleotide; it is very difficult to analyze this protein.

The sedimentation velocities data, although providing Svedberg values for different Rca concentrations, could not be used to calculate reliable molecular weights. The appearance of peaks provides information about species in solution; however, a heterogenous lightly dynamic mixture is likely to provide sedimentation coefficients that are controlled by protein interaction. During the experiment, Rca subunits probably interact with each other, exchanging rapidly.

In sedimentation equilibrium experiments, the data can be fitted to more than one model with equally acceptable results. Therefore, different assembly models cannot be distinguished.

Another possibility is that the protein was not pure enough. The quantity of sample required for AUC is typically a few hundred micrograms of greater than 95% purity. At that time, we had purified the Rca without affinity tag, and all the AUC runs were performed from these purification pools. Now, we know that Rca preparations are cleaner when the affinity tag is present during the purification. If samples are not pure enough, the analysis of the data may yield inconsistent results that are not reliable.

Size Exclusion High-Pressure Liquid Chromatography (SE-HPLC)

Theory

Size exclusion chromatography, also known as gel filtration chromatography or gel permeation chromatography, is a method that separates molecules in solution based on their size. The column is packed with small, porous material and filled with the mobile phase. Injected samples diffuse into pores, where smaller particles permeate into the pores. Since larger particles are unable to enter the smaller pores, they elute from the column first.

Methods and generation of standard curve for size determination

Rca-WT-like preparations were analyzed by SE-HPLC using a Biosep S-4000 (Phenomenex) column (600 x 7.8 mm, resolving range 15 000 to 1 500 000 Daltons). The mobile phase contained 25 mM HEPES pH 7.5, 250 mM KCl, 10 % glycerol and 0.1 mM ADP or 50mM phosphate (K^+) pH 7.2, 1mM MgCl2, 0.1mM ADP. A chosen flow rate of 0.8 mL/min and absorbance of 280 nm were used to monitor all injections. In practice, to calculate particle size, a number of well-characterized protein standards were used to form a standard curve for size determination. The standard curve was generated for the second mobile phase (Fig.25), using thyroglobulin (670 kDa), γ-globulin (158kDa), ovalbumin (44 kDa), myoblobin (17kDa), Arabidopsis thaliana Rubisco (585 kDa), and GFP (28.3 kDa) as protein standards (Table 13). Blue dextran (~ 2 000 kDa) was used as a high-MW specie to fully define the excluded (V_o) volume and Vitamin B_{12} (1.35 kDa) for a low-MW species to get the fully included (V_t) volume of the column. Normalized

sample elution volume (V_e) was determined from values of volumes mentioned above and the resolving range was determined according to the partition coefficient equation $K_{av}=(V_e-V_o)/(V_t-V_o)$. Log (MW) vs. K_{av} were used to generate a standard curve and to determine the MW of the Rca samples. A comparison of mass utilized by SE-HPLC to the known mass of protein standards gave 25% relative standard error.

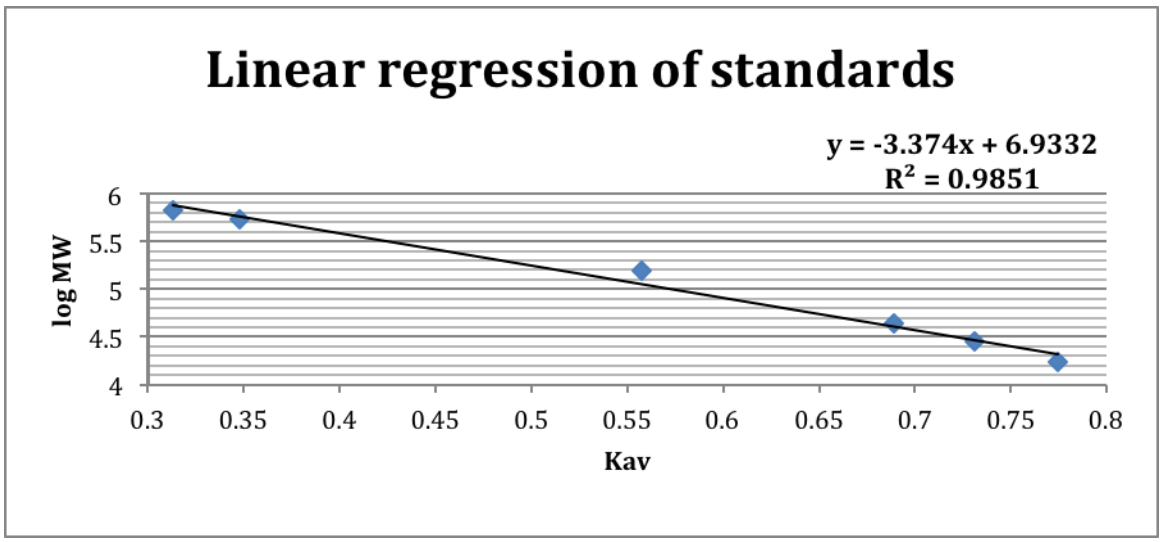


Figure 26. The representation of a linear regression of different standards for a size determination of samples.

Samples of cotton *Gossypium*–Rca (6.3 mg/mL) were prepared in buffer containing 25 mM HEPES pH 7.5, 250 mM KCl, 10 % glycerol and 0.1 mM ADP. The protein was kept on ice before injection into the HPLC system and followed by injection of four different volumes (50, 80, 230, 480 µL) (Fig.26). Molecular mass estimates (Table 12) were carried out for the peak positions of each elution band using the standard curve (Fig. 26).

	Elution time A [min]	Elution time B [min]	Elution time C [min]	Average [min]	Elution volume average [ml]	Kav	Molecular weight Manufacturer [Da]	log molecular weight
Thyroglobulin	19.05	19.04	19.16	19.08	15.27	0.31	670×10^3	5.82
Rubisco arabidopsis (8L8S)	V_e			19.55	15.64	0.35	542×10^3	5.73
γ-globulin	22.37	22.34	22.48	22.40	17.92	0.55	158×10^3	5.19
Ovalbumin	24.16	24.12	24.27	24.18	19.35	0.68	44×10^3	4.64
GFP trix	24.74	24.78		24.76	19.81	0.73	28×10^3	4.45
Myoblobin	25.31	25.29	25.43	25.34	20.27	0.77	17×10^3	4.23
Vitamin B12 V_t	28.39	28.36	28.48	28.41	22.73	1	1.35×10^3	3.13
Blue Dextran V_o	14.64	14.98	14.86	14.83	11.86		$2{,}000 \times 10^3$	

Table 12 Values of elution times for standard proteins, manufacturer MW, K_{av} and log MW values.

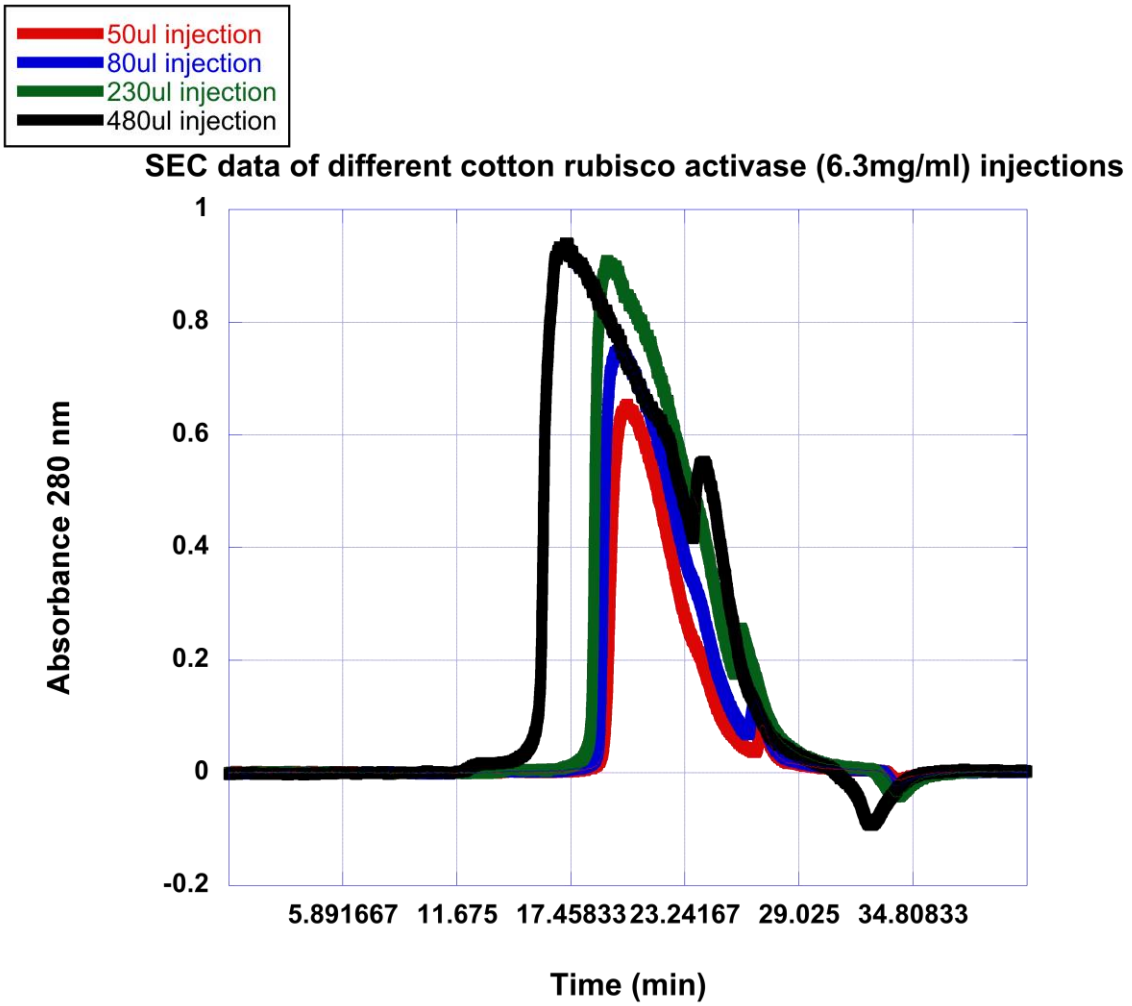


Figure 27. Size exclusion chromatograms collected on different Rca injection volumes. The approximate subunit stoichiometries at the peak positions are indicated for Rca volumes as injected: 480, 230, 80, 50 µL.

Table 13. Representation of elution times from SE-HPLC and subunit stoichiometries at different injection volumes.

Injection volumes (µL)	Time (min)	Stoichiometry
50	20.25	2
80	19.55	10
230	19.30	12
480	17.00	35

Results

Rca self-assembly monitored by size-exclusion chromatography

Size –exclusion chromatography provided us with the size distribution profiles of cotton Rca-WT as a function of protein concentration in the presence of ADP.

The peak positions of protein bands were used to calculate the molecular mass by a standard curve. Next, the subunit stoichiometries were defined and the relative standard error was determined to be 25%. All SE-HPLC elutions contained non-Gaussian bands with a long, trailing down-slope, meaning that a broad spread of different protein populations were observed.

THERMOFLUOR ASSAY

Theory

The Thermofluor assay is a quick, temperature-based assay to assess the stability of proteins. It can be used as an alternative to Circular Dichroism melts to estimate melting temperatures. The method takes a small amount of protein (0.25mg/mL) and adds SYPRO Orange dye. The dye binds to hydrophobic patches/denatured protein and fluoresces. As the temperature increases, the protein unfolds and the increase in fluorescence can be monitored and approximate melting temperatures are determined.

Methods

Applied Biosystems ABI Prism 7900 HT Sequence Detection System was used for all data collection. Samples were pipetted into the wells of 384-well polypropylene TempPlate PCR plates and covered with optically clear seals prior to data acquisition. Each well contained 20 µL total volume with concentrations of 0.25 mg/mL cotton β-Rca WT or D173N, 25mM HEPES pH 7.5, 150 mM KCl, and 16x SYPRO Orange (Invitrogen, Molecular Probes, Eugene, OR).

1.05x premixes were made from, in order of addition, Millipore deionized water, 1M HEPES pH 7.5, 3 M KCl, 1M $MgCl_2$, 0.5M ATP, 0.5M ADP (where necessary), 3.11 mg/mL cotton β-Rca WT or 2.6 mg/mL D173N, and 300x SYPRO Orange (made by diluting the 5000x stock in DMSO with Millipore deionized water). 20mM K_2HPO_4/KH_2PO_4 at pH 7.5 (P_i) was included in some of the samples as part of the premix before additive addition. In theses cases, the KCl content of the premixture was lowered to account for the effect of increased ionic strength.

To each well, 19 µL of the appropriate 1.05x premix was added followed by 1 µL of various concentrated stock solutions (e.g. ADP, ATP, $MgCl_2$). All conditions were reproduced in triplicate. Thermal denaturation runs consisted of three steps: 1) a fast ramp (~0.5 ºC/sec) from 25 ºC to 4 ºC, where the temperature was held for two minutes; 2) a slow ram (~0.03ºC/sec) from 4ºC to 80ºC with two minutes hold; 3) a fast ramp (~0.5ºC/sec) down to 4ºC. Fluorescence data were collected at ~8.5second intervals throughout each run. The fluorescence intensity of each sample at the protein "melting" transition was roughly approximated using the following equation:

I_{Tm} =B-1/2 (B-A)

Where, A is the minimum values prior to the fluorescence increase and B is the maximum in intensity following the increase. From these fluorescence values, corresponding melting temperatures were obtained.

Table 14. Stability studies of Rca-WT and D173N using Thermofluor Assay.

ADP [mM]	**ATP [mM]**	**Mg^{2+} [mM]**	**Pi [mM]**	**GhAβ_CA**	**D173N**
				Apparent T_m (°C ±STDEV)	
8				48.2±0.0	47.47±0.40
4				47.15±1.06	46.3±0.17
2				45.95±0.21	45.43±0.12
1				45.2±0.14	44.9±0.35
0.1				41.05±0.21	39.97±0.4
0.01				38.45±0.21	36±0.0
0	0			37±0.30	34.5±0.30
	8			40.6±0.0	42.70±0.20
	4			41.3±0.14	44.17±0.35
	2			40.25±0.49	44.00±0.20
	1			40.25±0.21	43.60±0.35
	0.1			38.45±0.21	40.90±0.00
	0.01			37.1±0.57	36.90±0.35
0	0	0		36.67±0.12	34.47±0.40
		9		35.43±0.12	33.30±0.36
		5		35.43±0.51	33.57±0.75
		3		35.67±0.58	33.63±0.40
		2		35.70±1.08	34.47±0.68
		1		36.47±0.29	34.83±0.29
8		9		45.47±0.40	46.27±0.12
4		5		46.23±0.29	46.60±0.66
2		3		42.70±0.17	42.30±0.00
1		2		44.47±0.12	44.53±0.12
	8	9		44.70±0.17	42.83±0.58
	4	5		44.97±0.51	43.43±0.12
	2	3		45.60±0.17	43.80±0.30
	1	2		45.77±0.12	44.20±0.17

8 4 2 1 0.5 0		3 3 3 3 3 3	20 20 20 20 20 20		49.40±0.35 47.57±0.12 45.50±0.17 44.30±0.17 42.30±0.17 35.70±0.66
8 4 2 1 0.5 0		3 3 3 3 3 3			47.37±1.06 45.87±0.40 40.10±0.17 43.70±0.17 41.60±0.17 35.87±0.12
	8 4 2 1 0.5 0	3 3 3 3 3 3	20 20 20 20 20 20		41.07±0.12 41.23±0.46 41.20±0.00 39.77±0.40 39.00±0.17 35.17±0.42
	8 4 2 1 0.5 0	3 3 3 3 3 3			39.73±0.29 40.17±0.93 40.10±0.17 39.43±0.25 38.47±0.12 34.50±0.50

Nucleotide, magnesium, and phosphate binding by the Thermofluor Assay

To illustrate the effects of nucleotides, magnesium, and phosphate, the thermal stability of cotton Rca was determined by the Thermofluor assay as a function of additives [132]. In this assay, hydrophobic groups of the protein are exposed due to high temperature that causes a fluorescent increase of the dye SYPRO Orange, up to the time of decrease of fluorescence at even higher temperatures. This assay provided information of apparent T_m values under different buffer conditions.

ADP-binding increases the thermal stability of Rca over ATP-binding.

Apoprotein GhbRca* samples depicted a strong increase in fluorescence upon heating with a midpoint (apparent T_m) value of 37.0 ± 0.30 °C. Increasing amounts of ADP caused a rise of the apparent T_m to 48.2 ± 0.00 °C at 8mM ADP. By contrast, ATP binding at 8mM impacted the stability of Rca less with a T_m of only 40.6 ± 0.00 °C (Table 14).

The difference of $T_{m\text{-}app}$ values for ADP and ATP are consistent with published K_d values for nucleotide, showing tighter association of ADP than ATP [119, 133].

Magnesium ions increase the thermal stability of ATP-bound Rca, slightly destabilizes ADP-bound Rca, and destabilizes Rca in the absence of any nucleotide.

ATP, and to a smaller degree ADP, are known to associate with magnesium ions within the active sites of many nucleotide- binding proteins. Therefore, a series of ATP and ADP titrations, with a 1mM excess of $MgCl_2$ over nucleotide, were performed with Rca and are listed in Table 15. In the absence of nucleotides, the addition of Mg^{2+} caused a destabilization of the protein by about 10 °C. In the presence of ATP, Mg^{2+} stabilized Rca by about 5°C, whereas in the presence of ADP, Mg^{2+} slightly destabilized the protein.

Stability studies on cotton Rca-ß-D173N

Exactly the same additive concentrations were used for D173N samples as for Rca-WT-like. D173N seems to be equally stable, with T_m values of only about 1°C lower than Rca-WT-like, when titrated with ADP. Mg^{2+} alone destabilizes the mutant by about 11°C, and in the presence of nucleotides, the T_m values are similar to those obtained for Rca with nucleotides alone (by around 1°C lower) (Table 14).

Phosphate increases the thermal stability of ADP-bound and ATP-bound Rca.

The effect of phosphate was tested by adding 20 mM KH_2PO_4/K_2HPO_4 pH 7.5 to protein titrated with ADP or ATP alone, and with addition of constant 3mM $MgCl_2$.

In the presence of ADP, Mg^{2+} and phosphate, the T_m values appeared to be higher by about 2°C than those without Mg^{2+}. The T_m values are slightly lower in the presence of ATP by only 0.5-1.0 °C.

Additionally, Mg^{2+} was replaced with Mn^{2+} and Rca-WT-like's stability was measured with different nucleotides. This replacement increased apparent T_m value by 3˚C. The geometries of both metals are the same - octahedral coordination is preferred, but the electronegativity differs, which might change the affinity to bound Rca (Table 15).

From the literature, a single-site mutant of tobacco Rca R294A with reduced activity [58] was shown to form hexameric structures in the presence of ATP-γS, as determined by high-resolution mass spectrometry [59]. The closely related but fully active R294V mutant in the presence of ATP and ATP-γS showed formation of hexameric assemblies in EM reconstructions, whereas ADP provided amorphous particles [21]. However, the thermofluor results on Rca-ß-R294V provided information that this variant is denatured.

This conclusion was made based on high starting fluorescence that was decreasing with increasing temperatures and nucleotides concentrations. The apparent T_m values could not be determined.

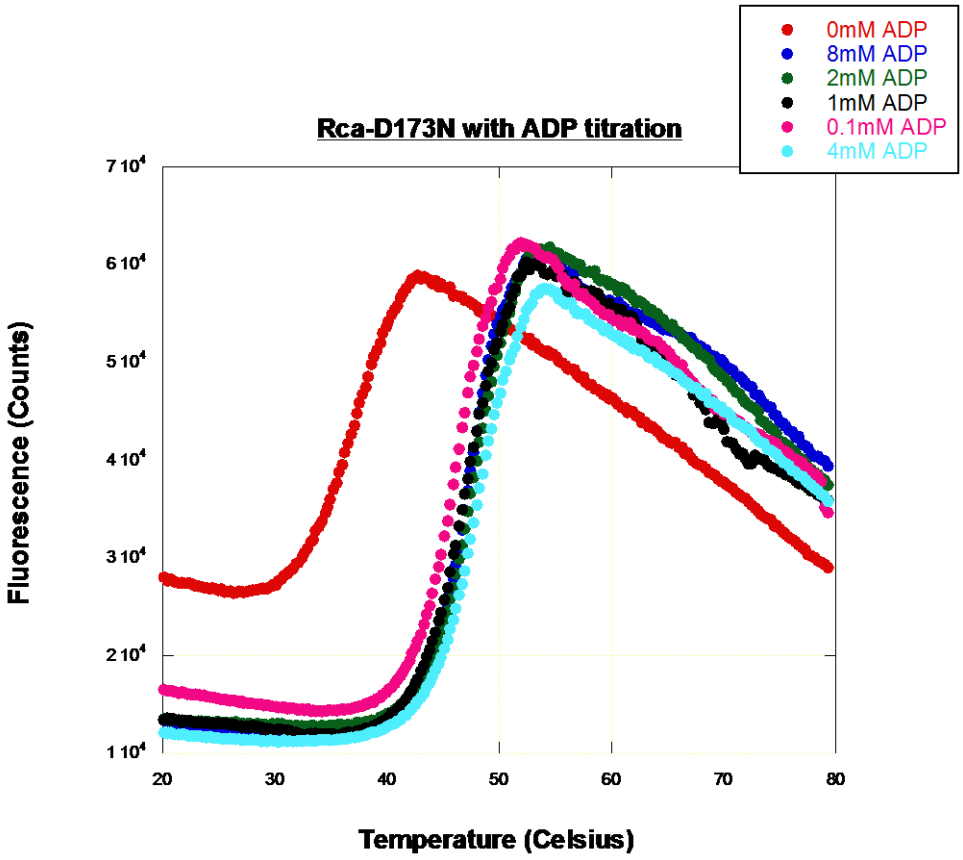
Summary

Based on the thermofluor assay, the Rca-WT like protein in the absence of nucleotides is destabilized. ADP at 2mM clearly helps to form more stable Rca by around 8 °C, wherea ATP only stabilizes by 3 °C. In the presence of Mg^{2+}•2mM ADP, the protein becomes less stable by 3°C. In contrast Mg^{2+}•2mM ATP formed a more stable complex by 5°C. This indicates that without magnesium ions, the Rca-WT-like protein may have a higher affinity for ADP than ATP. The γ-phosphate of ATP may also weaken the binding of the adenine nucleotide. The T_m values are higher when Mg^{2+} and ATP are present in solution, and this may suggest that the Mg^{2+} cofactor coordinates the γ-phosphate of ATP, possibly through an interaction with an aspartate residue [119].

The apparent T_m values for the mutant D173N in the presence of 2mM ADP are 11°C and in the presence of 2mM ATP 10°C higher than in an apo-protein solution. Rca is not stabilized by Mg^{2+} alone, but the addition of P_i, and 2mM ADP raises the T_m value by 5°C.

The aspartic acid residue located within the Walker B motif in cotton Rca (conserved glutamate for many other AAA^+ proteins) may be responsible for ATP hydrolysis. The aspartate most likely primes a water molecule for nucleophilic attack on the γ-phosphate group of ATP [134]. Based on published results, the mutant D173A [119], hinders ATP hydrolysis and causes the substrate to be trapped [135]. The mutation of aspartic acid into

asparagine blocks the hydrolysis of ATP. Binding of negatively charged ATP to the mutant enzyme appears stronger than with the wild type Rca, presumably due to the removal of electrostatic clashes accompanying the mutation of the negatively charge aspartic acid side chain to a neutral asparagine.



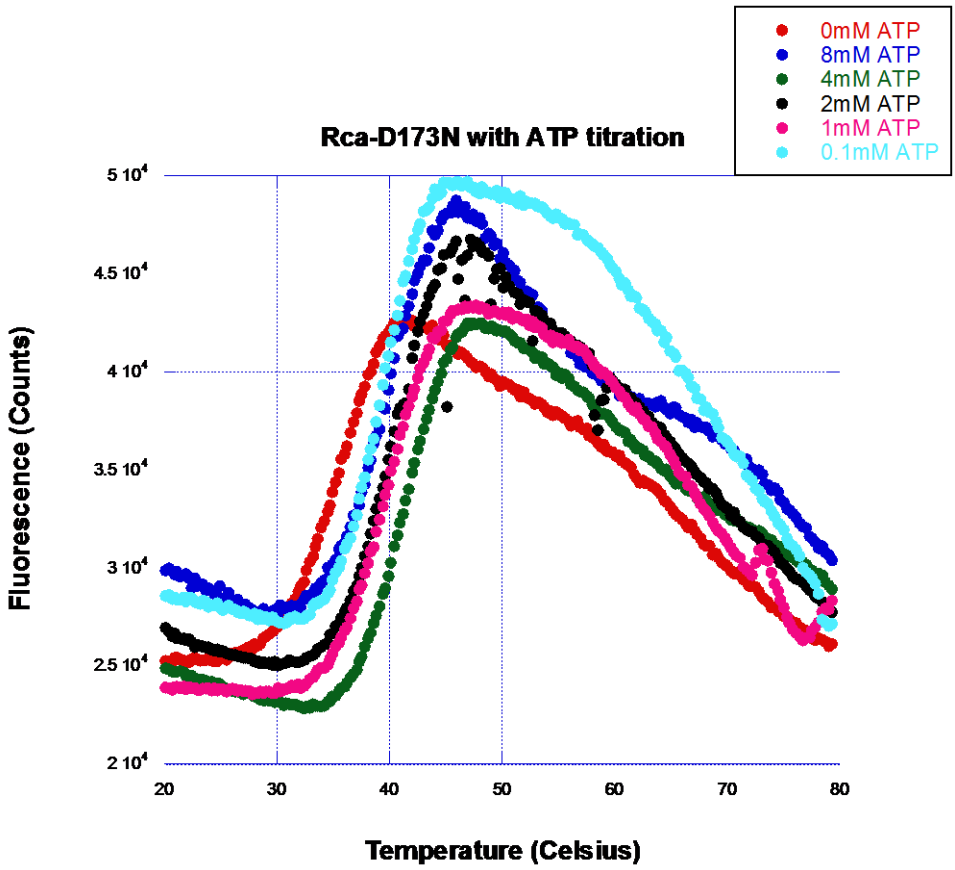


Figure 28. Thermofluor assays on Rca-ß-D173N. A) ADP titration. B) ATP titration

Discussion

Understanding the self-assembly of Rca is an important part of studying the enzymology of this protein. Currently, there are plenty of techniques available that can be used to study the stoichiometries of protein assemblies. Unfortunately, many of these techniques did not provide interpretable data. The biggest drawbacks of the Rubisco activase are high polydispersity and thermolability. These features made it hard to extract useful data from some of the techniques employed. For example, AUC data happened to be difficult to analyze and to interpret. One reason was that the protein, in order to stay active, must

contain nucleotides in solution. In the absence of nucleotide, Rca aggregates. The presence of nucleotides limits obtaining high-quality data by monitoring absorbance 280 nm, since nucleotides absorb at this wavelength. Although interference optics were also used, the data could be fit equally well to several models. Unfortunately, AUC data were not reproducible.

Another issue related to that method is very long run time of the experiment. The samples, although at 4˚C, were spinning in the rotor for three weeks, such a long time may cause loss of activity and perhaps denaturation of the protein forming very large aggregates. Indeed, fitted AUC results contained some fraction of a very large protein size, indicating formation of aggregates over time.

SE-HPLC, another method used to study the oligomerization of Rca, appeared to be uninformative mainly because of the polydisperse character that could be observed by formation of non-Gaussian bands with a long, trailing down-slope. It indicates that many different species elute in a broad long slope and therefore not resolved. SE-HPLC is also not the most precise technique, with an error of above 25%.

The Thermofluor Assay was a valuable method to study the stability of Rca protein. The apparent T_m values informed about the stability under different buffer conditions. The absence of nucleotides destabilizes Rca, whereas 2mM ADP forms most stable complexes with Rca. The Mg•ATP bound Rca has higher T_m values than these for Rca with Mg•ADP by around 3º. The T_m measurements of the mutant D173N also showed that the protein is destabilized in the absence of nucleotides, in contrast addition of 2mM ADP or ATP raised the apparent T_m by ~ 10º. This assay provided a comparison of Rca-WT and D173N, showing that both constructs are equally stable at certain conditions,

which was crucial information in order to continue to work with Rca-WT and the mutant D173N.

CHAPTER 5

DISCUSSION

In order to understand the mechanistic enzymology of complex protein assemblies and the cooperativity between subunits, it is necessary to know the stoichiometries and fractional concentrations of all component species. Therefore, a lot of effort was expended to find a useful method to study the oligomerization of Rubisco activase. This highly aggregation-prone protein remains stable only under very specific conditions, which resulted in a significant amount of time spent on sample preparation and the selection of an appropriate method. An FCS-based method was developed and Rca protein samples were tested. It was possible to access the low-nanomolar to mid-micromolar range by mixing labeled and unlabeled protein.

The FCS experiments allowed for developing thermodynamic models for self-association, where multiple oligomeric species co-exist. This method proved to be a better approach to examine stoichiometries for a highly dynamic fast-exchanging system than SE-HPLC, AUC, or DLS.

The FCS and PCH methods were used to monitor the Mg•ADP-mediated assembly process of cotton β-Rca at pH 7.6 and depicted that, at concentrations below 0.5 μM, monomers are the dominating species. These results are in agreement with EM images of molecules at 0.5 - 1 μM mutant tobacco Rca-R293V with Mg•ADP. At higher concentrations (100μM), the data are consistent with higher molecular weight aggregates (roughly 24 subunits). The data lack an obvious plateau at any concentration. Therefore, the intermediate states could not be determined by FCS (Fig. 1B-1D) and other

information had to be used to interpret the FCS results. There are plenty of studies of Rca or related proteins (ATPases) that have shown formation of hexamers. For instance, EM images of ATP-dependent formation of a closed ring hexamer were observed for 0.5 - 1 μM tobacco Rca-R293V, whereas the crystal structure of tobacco was solved in a spiral hexameric model [21]. The AAA^+ protein FtsH•protease formed a toroidal hexamer, whereas without protease, the protein formed helical assemblies. ClpA, a related protein that involves dimers and tetramers as intermediates, was found to be in agreement with Model 3 (Fig.15B) [42, 86]. The differences are that tighter complexes are formed by ClpA, with K_d values of ~1 nM (dimer-tetramer) and below 0.17 nM (tetramer-hexamer) [42].

Rca assembly mechanisms with ADP•Mg

The FCS method presented three possible assembly mechanisms that proceed to the ~24 subunit stage. In Model 3, the hexamer starts to rise at 1 μM Rca, dominates between 10 and 30 μM (this is consistent with NanoESI data of 8 μM Rca, that forms hexamers [59]), and the tetramer contributes between 5 and 20 μM. Assemblies larger than 6 subunits start to dominate at higher concentration. However, high range concentration data cannot distinguish between models of closed hexameric rings and pseudo-hexameric spirals. Based on K_d values calculated from dimer-tetramer and tetramer-hexamer equilibria (about 1 μM), it is appropriate to assume that addition of a dimer to a dimer or a tetramer is energetically equivalent. This fact suggests that a spiral assembly is formed and also explains the broad size distribution seen by SE-HPLC.

Rca assembly mechanisms with ATP•Mg

The FCS method encouraged the investigation of Rca assembly as a function of different ADP/ATP ratios and magnesium concentrations as modulated by the energy charge of the chloroplast stroma. In order to study the self-assembly of Rca with ATP, an appropriate mutant was designed. The aspartic acid residue found in the Walker B motif of Rca, residue 173, is thought to play an important role in ATP hydrolysis; it primes a water molecule for a nucleophilic attack of the γ-phosphate group of ATP. The mutation of aspartic acid into asparagine hinders the hydrolysis of ATP. Evidently, replacement of the negatively charged aspartate with an uncharged amino acid helps the negatively charged nucleotide to bind to the active-site pocket.

The FCS data of the Rca-D173N mutant were interpreted using the same association model (Model 3), and again suggest formation of monomers at low concentrations up to 1 μM, whereas rise of hexamers lies in a range of 8-70 μM (Fig.21). This result is in agreement with published electron micrographs of a hexameric ring of tobacco R294V Rca in the presence of ATP [21]. In the case where Rca mutant D173N is bound to ADP, the existing species in solution correspond to FCS results of Rca-WT-like bound with ADP, where larger oligomers (greater than hexamers) are formed. In the presence of ATP in solution, higher assemblies are not favorable and it can be concluded from the determined value of K_{d4}, which is around 120 fold larger than the K_d in the presence of ADP alone. The change in populations between hexamers and larger oligomers as a functioning nucleotide might have a physiological role during the light and dark phases

of photosynthesis. During the day, the stroma's ATP level is high and the activity of Rca is up-regulated, whereas during the night, the ADP level rises and Rca activity is down-regulated [18, 136].

The FCS results suggest that changes in nucleotide regulates Rca activity. In the experiment where five different ATP/ADP ratios were measured at 60 µM of Rca, hexamerization is observed in the presence of ATP, but drops when the ATP is replaced with ADP.

The Mg^{2+} titration data show that the hexameric state diminishes at low Mg^{2+} concentrations. The optimum concentration lies in a range of total free 3 – 5 mM Mg^{2+}, which is also the closest to mimicking the physiological conditions in plants. These results are in good agreement with the ATPase protein kinesin that is only functional in the presence of ATP•Mg^{2+} complex and the bound magnesium is crucial to the hydrolysis of ATP. At lower Mg^{2+} concentration and in the presence of ATP, the association of Rca subunits seems to be identical to that in the presence of ADP and 5 mM Mg^{2+}. The coordination of ATP to the Mg^{2+} cation plays an important role in ATP affinity for the binding pocket. The affinity without Mg^{2+} is much lower than for the complex formed with ATP, as also seen in other ATPases [137, 138]. Formation of hexamers when free Mg^{2+} are present, indicates that there is a second magnesium binding site, in addition to coordination of ATP.

The FCS assembly studies provide valuable information about the formation of active states of Rca that are formed in the presence of either ATP or ADP.

The research presented here opens an understanding of the molecular mechanisms by which Rca activity is up- and down-regulated under different physiological conditions.

The overarching research goal is to learn about the Rca structural features and oligomeric states that respond to changes in stromal energy charge, and how this response is utilized to regulate CO_2 fixation by Rubisco.

REFERENCES

1. Stocker, T.F., D.Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.Boschung, A.Nauels, Y.Xia, V.Bex and P.M. Midgley (eds), *Climate Change 2013: The Physical Science Basis. Contributon of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. .* IPCC,2013, 2013: p. 1535

2. Battisti, D.S. and R.L. Naylor, *Historical warnings of future food insecurity with unprecedented seasonal heat.* Science, 2009. **323**(5911): p. 240-4.

3. Sage, R.F., D.A. Way, and D.S. Kubien, *Rubisco, Rubisco activase, and global climate change.* J Exp Bot, 2008. **59**(7): p. 1581-95.

4. Berry, J. and O. Bjorkman, *Photosynthetic Response and Adaptation to Temperature in Higher-Plants.* Annual Review of Plant Physiology and Plant Molecular Biology, 1980. **31**: p. 491-543.

5. Spreitzer, R.J. and M.E. Salvucci, *Rubisco: structure, regulatory interactions, and possibilities for a better enzyme.* Annu Rev Plant Biol, 2002. **53**: p. 449-75.

6. Lorimer, G.H., M.R. Badger, and H.W. Heldt, *The activation of ribulose 1,5-bisphosphate carboxylase/oxygenase.* Basic Life Sci, 1978. **11**: p. 283-306.

7. Jordan, D.B. and R. Chollet, *Inhibition of Ribulose Bisphosphate Carboxylase by Substrate Ribulose 1,5-Bisphosphate.* Journal of Biological Chemistry, 1983. **258**(22): p. 3752-3758.

8. Andrews, T.J., *The bait in the Rubisco mousetrap.* Nature Structural Biology, 1996. **3**(1): p. 3-7.

9. Somerville, C.R., A.R. Portis, and W.L. Ogren, *A Mutant of Arabidopsis thaliana Which Lacks Activation of RuBP Carboxylase In Vivo.* Plant Physiol, 1982. **70**(2): p. 381-7.

10. Diemand, A.V. and A.N. Lupas, *Modeling AAA+ ring complexes from monomeric structures.* Journal of Structural Biology, 2006. **156**(1): p. 230-243.

11. Yokota, A. and N. Tsujimoto, *Characterization of Ribulose-1,5-Bisphosphate Carboxylase Oxygenase Carrying Ribulose 1,5-Bisphosphate at Its Regulatory Sites and the Mechanism of Interaction of This Form of the Enzyme with Ribulose-1,5-Bisphosphate-Carboxylase Oxygenase Activase.* European Journal of Biochemistry, 1992. **204**(2): p. 901-909.

12. Ott, C.M., et al., *Activase region on chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase - Nonconservative substitution in the large subunit*

alters species specificity of protein interaction. Journal of Biological Chemistry, 2000. **275**(34): p. 26241-26244.

13. Neuwald, A.F., et al., *AAA(+): A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes.* Genome Research, 1999. **9**(1): p. 27-43.

14. Iyer, L.M., et al., *Evolutionary history and higher order classification of AAA plus ATPases.* Journal of Structural Biology, 2004. **146**(1-2): p. 11-31.

15. Ogura, T., S.W. Whiteheart, and A.J. Wilkinson, *Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA(+) ATPases.* Journal of Structural Biology, 2004. **146**(1-2): p. 106-112.

16. Smith, C.K., T.A. Baker, and R.T. Sauer, *Lon and Clp family proteases and chaperones share homologous substrate-recognition domains.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(12): p. 6678-6682.

17. Portis, A.R., et al., *Regulation of Rubisco activase and its interaction with Rubisco.* Journal of Experimental Botany, 2008. **59**(7): p. 1597-1604.

18. Zhang, N., et al., *Light modulation of Rubisco in Arabidopsis requires a capacity for redox regulation of the larger Rubisco activase isoform.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(5): p. 3330-3334.

19. Fiser, A. and A. Sali, *Modeller: generation and refinement of homology-based protein structure models.* Methods Enzymol, 2003. **374**: p. 461-91.

20. Henderson, J.N., et al., *Atomic Resolution X-ray Structure of the Substrate Recognition Domain of Higher Plant Ribulose-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase.* Journal of Biological Chemistry, 2011. **286**(41): p. 35683-35688.

21. Stotz, M., et al., *Structure of green-type Rubisco activase from tobacco.* Nature Structural & Molecular Biology, 2011. **18**(12): p. 1366-U78.

22. Mueller-Cajar, O., et al., *Structure and function of the AAA+ protein CbbX, a red-type Rubisco activase.* Nature, 2011. **479**(7372): p. 194-9.

23. Lan, Y. and K.A. Mott, *Determination of Apparent Km Values for Ribulose 1,5-Bisphosphate Carboxylase Oxygenase (Rubisco) Activase Using the*

Spectrophotometric Assay of Rubisco Activity. Plant Physiology, 1991. **95**(2): p. 604-609.

24. Robinson, S.P. and A.R. Portis, *Release of the Nocturnal Inhibitor, Carboxyarabinitol-1-Phosphate, from Ribulose Bisphosphate Carboxylase Oxygenase by Rubisco Activase.* Febs Letters, 1988. **233**(2): p. 413-416.

25. Wang, Z.Y., R.T. Ramage, and A.R. Portis, *Mg(2+) 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.* Biochimica Et Biophysica Acta, 1993. **1202**(1): p. 47-55.

26. Podzimek, S., *Light Scattering, Size Exclusion Chromatography and Asymmetric Flow Field Flow Fractionation: Powerful Tools for the Characterization of Polymers, Proteins and Nanoparticles.* 2011.

27. Keown, J.R., et al., *Small Oligomers of Ribulose-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase Are Required for Biological Activity.* Journal of Biological Chemistry, 2013. **288**(28): p. 20607-20615.

28. Andersson, I. and A. Backlund, *Structure and function of Rubisco.* Plant Physiology and Biochemistry, 2008. **46**(3): p. 275-291.

29. Baker, T.S., et al., *The structure of form I crystals of D-ribulose-1,5-diphosphate carboxylase.* J Mol Biol, 1975. **91**(4): p. 391-9.

30. Wang, J., et al., *Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU.* Structure, 2001. **9**(11): p. 1107-16.

31. Goodsell, D.S., *Inside a living cell.* Trends Biochem Sci, 1991. **16**(6): p. 203-6.

32. Goodsell, D.S. and A.J. Olson, *Structural symmetry and protein function.* Annu Rev Biophys Biomol Struct, 2000. **29**: p. 105-53.

33. Nooren, I.M.A. and J.M. Thornton, *Structural characterisation and functional significance of transient protein-protein interactions.* Journal of Molecular Biology, 2003. **325**(5): p. 991-1018.

34. Goodsell, D.S. and A.J. Olson, *Soluble-Proteins - Size, Shape and Function.* Trends in Biochemical Sciences, 1993. **18**(3): p. 65-68.

35. Nooren, I.M.A. and J.M. Thornton, *Diversity of protein-protein interactions.* Embo Journal, 2003. **22**(14): p. 3486-3492.

36. Walden, H., et al., *Tiny TIM: A small, tetrameric, hyperthermostable triosephosphate isomerase.* Journal of Molecular Biology, 2001. **306**(4): p. 745-757.

37. Grimaud, R., et al., *Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP.* Journal of Biological Chemistry, 1998. **273**(20): p. 12476-12481.

38. Bochtler, M., et al., *The structures of HsIU and ATP-dependent protease HsIU-HsIV.* Nature, 2000. **403**(6771): p. 800-805.

39. Parsell, D.A., A.S. Kowal, and S. Lindquist, *Saccharomyces-Cerevisiae Hsp104 Protein - Purification and Characterization of Atp-Induced Structural-Changes.* Journal of Biological Chemistry, 1994. **269**(6): p. 4480-4487.

40. Kessel, M., et al., *Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome.* J Mol Biol, 1995. **250**(5): p. 587-94.

41. Maurizi, M.R., et al., *Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and ClpP.* Biochemistry, 1998. **37**(21): p. 7778-7786.

42. Kress, W., H. Mutschler, and E. Weber-Ban, *Assembly pathway of an AAA+ protein: tracking ClpA and ClpAP complex formation in real time.* Biochemistry, 2007. **46**(21): p. 6183-93.

43. Zolkiewski, M., et al., *Nucleotide-dependent oligomerization of ClpB from Escherichia coli.* Protein Science, 1999. **8**(9): p. 1899-1903.

44. Hashimoto, K., et al., *Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization.* Phys Biol, 2011. **8**(3): p. 035007.

45. Ali, M.H., and B. Imperiali, *Protein oligomerization: How and why.* Bioorg. Med. Chem, 2005. **13**: p. 5013-5020.

46. Crafts-Brandner, S.J. and M.E. Salvucci, *Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO2.* Proc Natl Acad Sci U S A, 2000. **97**(24): p. 13430-5.

47. Portis, A.R., Jr., *Rubisco activase - Rubisco's catalytic chaperone.* Photosynth Res, 2003. **75**(1): p. 11-27.

48. Wang, Z.Y. and A.R. Portis, *Dissociation of ribulose-1,5-bisphosphate bound to ribulose-1,5-bisphosphate carboxylase/oxygenase and its enhancement by*

ribulose-1,5-bisphosphate carboxylase/oxygenase activase-mediated hydrolysis of ATP. Plant Physiol, 1992. **99**(4): p. 1348-53.

49. Andersson, I., *Catalysis and regulation in Rubisco.* J Exp Bot, 2008. **59**(7): p. 1555-68.

50. Pearce, F.G. and T.J. Andrews, *The relationship between side reactions and slow inhibition of ribulose-bisphosphate carboxylase revealed by a loop 6 mutant of the tobacco enzyme.* J Biol Chem, 2003. **278**(35): p. 32526-36.

51. Salvucci, M.E., F.J. van de Loo, and D. Stecher, *Two isoforms of Rubisco activase in cotton, the products of separate genes not alternative splicing.* Planta, 2003. **216**(5): p. 736-44.

52. Barta, C., et al., *Structural changes associated with the acute thermal instability of Rubisco activase.* Arch Biochem Biophys, 2010. **499**(1-2): p. 17-25.

53. Carmo-Silva, A.E. and M.E. Salvucci, *The activity of Rubisco's molecular chaperone, Rubisco activase, in leaf extracts.* Photosynth Res, 2011. **108**(2-3): p. 143-55.

54. Snider, J. and W.A. Houry, *AAA+ proteins: diversity in function, similarity in structure.* Biochem Soc Trans, 2008. **36**(Pt 1): p. 72-7.

55. Hartman, J.J. and R.D. Vale, *Microtubule disassembly by ATP-dependent oligomerization of the AAA enzyme katanin.* Science, 1999. **286**(5440): p. 782-5.

56. Smith, D.M., et al., *ATP binds to proteasomal ATPases in pairs with distinct functional effects, implying an ordered reaction cycle.* Cell, 2011. **144**(4): p. 526-38.

57. Hattendorf, D.A. and S.L. Lindquist, *Cooperative kinetics of both Hsp104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants.* EMBO J, 2002. **21**(1-2): p. 12-21.

58. Li, C., D. Wang, and A.R. Portis, Jr., *Identification of critical arginine residues in the functioning of Rubisco activase.* Arch Biochem Biophys, 2006. **450**(2): p. 176-82.

59. Blayney, M.J., S.M. Whitney, and J.L. Beck, *NanoESI mass spectrometry of Rubisco and Rubisco activase structures and their interactions with nucleotides and sugar phosphates.* J Am Soc Mass Spectrom, 2011. **22**(9): p. 1588-601.

60. Jameson, D.M., J.A. Ross, and J.P. Albanesi, *Fluorescence fluctuation spectroscopy: ushering in a new age of enlightenment for cellular dynamics.* Biophys Rev, 2009. **1**(3): p. 105-118.

61. Muller, J.D., Y. Chen, and E. Gratton, *Fluorescence correlation spectroscopy.* Methods Enzymol, 2003. **361**: p. 69-92.

62. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence correlation spectroscopy. II. An experimental realization.* Biopolymers, 1974. **13**(1): p. 29-61.

63. Thompson, N.L., *FLuorescence Correlation Spectroscopy. In Topics in Fluorescence Spectroscopy.* Plenum Press, 1991: p. 337-355.

64. Timbs, M.M., et al., *Binding and mobility of anti-dinitrophenyl monoclonal antibodies on fluid-like, Langmuir-Blodgett phospholipid monolayers containing dinitrophenyl-conjugated phospholipids.* Biochim Biophys Acta, 1991. **1064**(2): p. 219-28.

65. Kahya, N. and P. Schwille, *Fluorescence correlation studies of lipid domains in model membranes.* Mol Membr Biol, 2006. **23**(1): p. 29-39.

66. Berland, K.M., P.T. So, and E. Gratton, *Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment.* Biophys J, 1995. **68**(2): p. 694-701.

67. Lukacs, G.L., et al., *Size-dependent DNA mobility in cytoplasm and nucleus.* J Biol Chem, 2000. **275**(3): p. 1625-9.

68. Wachsmuth, M., W. Waldeck, and J. Langowski, *Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy.* J Mol Biol, 2000. **298**(4): p. 677-89.

69. Abney, J.R., et al., *Chromatin dynamics in interphase nuclei and its implications for nuclear structure.* J Cell Biol, 1997. **137**(7): p. 1459-68.

70. Hemmerich, P., et al., *Dynamics of inner kinetochore assembly and maintenance in living cells.* J Cell Biol, 2008. **180**(6): p. 1101-14.

71. Rajagopalan, S., F. Huang, and A.R. Fersht, *Single-Molecule characterization of oligomerization kinetics and equilibria of the tumor suppressor p53.* Nucleic Acids Res, 2011. **39**(6): p. 2294-303.

72. Krouglova, T., J. Vercammen, and Y. Engelborghs, *Correct diffusion coefficients of proteins in fluorescence correlation spectroscopy. Application to tubulin oligomers induced by Mg2+ and Paclitaxel.* Biophys J, 2004. **87**(4): p. 2635-46.

73. Salvucci, M.E., *Potential for interactions between the carboxy- and amino-termini of Rubisco activase subunits.* FEBS Lett, 2004. **560**(1-3): p. 205-9.

74. Gendron, P.O., F. Avaltroni, and K.J. Wilkinson, *Diffusion coefficients of several rhodamine derivatives as determined by pulsed field gradient-nuclear magnetic resonance and fluorescence correlation spectroscopy.* J Fluoresc, 2008. **18**(6): p. 1093-101.

75. Muller, C.B., A. Loman, V. Pacheco, F. Koberling, D. Willbold, W. Richtering, and J. Enderlein, *Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy.* Europhys. Lett, 2008(83): p. 46001.

76. Petrasek, Z. and P. Schwille, *Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy.* Biophys J, 2008. **94**(4): p. 1437-48.

77. Rubinstein, M., and R. H. Colby. , *Polymer Physics.* Oxford University Press., 2003.

78. Squire, P.G. and M.E. Himmel, *Hydrodynamics and protein hydration.* Arch Biochem Biophys, 1979. **196**(1): p. 165-77.

79. Chen, Y., et al., *The photon counting histogram in fluorescence fluctuation spectroscopy.* Biophys J, 1999. **77**(1): p. 553-67.

80. Perroud, T.D., et al., *Photon counting histogram for one-photon excitation.* Chemphyschem, 2003. **4**(10): p. 1121-3.

81. Nath, S., et al., *Early aggregation steps in alpha-synuclein as measured by FCS and FRET: evidence for a contagious conformational change.* Biophys J, 2010. **98**(7): p. 1302-11.

82. Sahoo, B., et al., *Protein aggregation probed by two-photon fluorescence correlation spectroscopy of native tryptophan.* J Chem Phys, 2008. **129**(7): p. 075103.

83. Yu, L., et al., *Determination of critical micelle concentrations and aggregation numbers by fluorescence correlation spectroscopy: aggregation of a lipopolysaccharide.* Anal Chim Acta, 2006. **556**(1): p. 216-25.

84. Bieniossek, C., B. Niederhauser, and U. Baumann. , *Apo-FtsH crystal structure. RCSB protein data bank.* http://www.pdb.org 2009.

85. Niwa, H., et al., *Hexameric ring structure of the ATPase domain of the membrane-integrated metalloprotease FtsH from Thermus thermophilus HB8.* Structure, 2002. **10**(10): p. 1415-23.

86. Veronese, P.K., R.P. Stafford, and A.L. Lucius, *The Escherichia coli ClpA molecular chaperone self-assembles into tetramers.* Biochemistry, 2009. **48**(39): p. 9221-33.

87. F., P., *The Brownian movement of an ellipsoid. - The dielectric dispersion of ellipsoidal molecules.* J Phys-Paris 1934. **5**: p. 497-511.

88. Squire, P.G. and M.E. Himmel, *Hydrodynamics and Protein Hydration.* Archives of Biochemistry and Biophysics, 1979. **196**(1): p. 165-177.

89. Meseth, U., et al., *Resolution of fluorescence correlation measurements.* Biophysical Journal, 1999. **76**(3): p. 1619-1631.

90. Snider, J. and W.A. Houry, *AAA+ proteins: diversity in function, similarity in structure.* Biochemical Society Transactions, 2008. **36**: p. 72-77.

91. Ogura, T. and A.J. Wilkinson, *AAA(+) superfamily ATPases: common structure-diverse function.* Genes to Cells, 2001. **6**(7): p. 575-597.

92. Babst, M., et al., *The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function.* Embo Journal, 1998. **17**(11): p. 2982-2993.

93. Fodje, M.N., et al., *Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase.* Journal of Molecular Biology, 2001. **311**(1): p. 111-122.

94. Lenzen, C.U., et al., *Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein.* Cell, 1998. **94**(4): p. 525-536.

95. Hanson, P.I. and S.W. Whiteheart, *AAA+ proteins: Have engine, will work.* Nature Reviews Molecular Cell Biology, 2005. **6**(7): p. 519-529.

96. Page, A.N., et al., *Structure and Biochemical Activities of Escherichia coli MgsA.* Journal of Biological Chemistry, 2011. **286**(14).

97. Simonetta, K.R., et al., *The Mechanism of ATP-Dependent Primer-Template Recognition by a Clamp Loader Complex.* Cell, 2009. **137**(4): p. 659-671.

98. Lee, S.Y., et al., *Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA(+) ATPase domains.* Genes & Development, 2003. **17**(20): p. 2552-2563.

99. Portis, A.R., *Rubisco activase - Rubisco's catalytic chaperone.* Photosynthesis Research, 2003. **75**(1): p. 11-27.

100. Barta, C., et al., *Structural changes associated with the acute thermal instability of Rubisco activase.* Archives of Biochemistry and Biophysics, 2010. **499**(1-2): p. 17-25.

101. Henderson, J.N., et al., *Biophysical characterization of higher plant Rubisco activase.* Biochimica Et Biophysica Acta-Proteins and Proteomics, 2013. **1834**(1): p. 87-97.

102. Blayney, M.J., S.M. Whitney, and J.L. Beck, *NanoESI Mass Spectrometry of Rubisco and Rubisco Activase Structures and Their Interactions with Nucleotides and Sugar Phosphates.* Journal of the American Society for Mass Spectrometry, 2011. **22**(9): p. 1588-1601.

103. Mueller-Cajar, O., et al., *Structure and function of the AAA(+) protein CbbX, a red-type Rubisco activase.* Nature, 2011. **479**(7372): p. 194-U66.

104. Lilley, R.M. and A.R. Portis, *ATP hydrolysis activity and polymerization state of ribulose-1,5-bisphosphate carboxylase oxygenase activase - Do the effects of Mg2+, K+, and activase concentrations indicate a functional similarity to actin?* Plant Physiology, 1997. **114**(2): p. 605-613.

105. Li, C.S., D.F. Wang, and A.R. Portis, *Identification of critical arginine residues in the functioning of Rubisco activase.* Archives of Biochemistry and Biophysics, 2006. **450**(2): p. 176-182.

106. Chakraborty, M., et al., *Protein Oligomerization Monitored by Fluorescence Fluctuation Spectroscopy: Self-Assembly of Rubisco Activase.* Biophysical Journal, 2012. **103**(5): p. 949-958.

107. Rajagopalan, S., F. Huang, and A.R. Fersht, *Single-Molecule characterization of oligomerization kinetics and equilibria of the tumor suppressor p53.* Nucleic Acids Research, 2011. **39**(6): p. 2294-2303.

108. Krouglova, T., J. Vercammen, and Y. Engelborghs, *Correct diffusion coefficients of proteins in fluorescence correlation spectroscopy. Application to tubulin oligomers induced by Mg2+ and paclitaxel.* Biophysical Journal, 2004. **87**(4): p. 2635-2646.

109. Weibezahn, J., et al., *Characterization of a trap mutant of the AAA plus chaperone ClpB.* Journal of Biological Chemistry, 2003. **278**(35): p. 32608-32617.

110. Dalal, S., et al., *Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway.* Molecular Biology of the Cell, 2004. **15**(2): p. 637-648.

111. Hersch, G.L., et al., *Asymmetric interactions of ATP with the AAA+ ClpX(6) unfoldase: Allosteric control of a protein machine.* Cell, 2005. **121**(7): p. 1017-1027.

112. Monroe, N., et al., *The Oligomeric State of the Active Vps4 AAA ATPase.* Journal of Molecular Biology, 2014. **426**(3): p. 510-525.

113. Salvucci, M.E., *Potential for interactions between the carboxy- and amino-termini of Rubisco activase subunits.* Febs Letters, 2004. **560**(1-3): p. 205-209.

114. Provencher, S.W., *Inverse Problems in Polymer Characterization - Direct Analysis of Polydispersity with Photon Correlation Spectroscopy.* Makromolekulare Chemie-Macromolecular Chemistry and Physics, 1979. **180**(1): p. 201-209.

115. Koppel, D.E., *Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy - Method of Cumulants.* Journal of Chemical Physics, 1972. **57**(11): p. 4814-&.

116. Goodsell, D.S. and A.J. Olson, *Structural symmetry and protein function.* Annual Review of Biophysics and Biomolecular Structure, 2000. **29**: p. 105-153.

117. Ali, M.H. and B. Imperiali, *Protein oligomerization: How and why.* Bioorganic & Medicinal Chemistry, 2005. **13**(17): p. 5013-5020.

118. Wang, Z.Y. and A.R. Portis, *A Fluorometric Study with 1-Anilinonaphthalene-8-Sulfonic Acid (Ans) of the Interactions of Atp and Adp with Rubisco Activase.* Biochimica Et Biophysica Acta, 1991. **1079**(3): p. 263-267.

119. van de Loo, F.J. and M.E. Salvucci, *Involvement of two aspartate residues of Rubisco activase in coordination of the ATP gamma-phosphate and subunit cooperativity.* Biochemistry, 1998. **37**(13): p. 4621-5.

120. Gulick, A.M., et al., *X-ray structures of the MgADP, MgATP gamma S, and MgAMPPNP complexes of the Dichyostelium discoideum myosin motor domain.* Biochemistry, 1997. **36**(39): p. 11619-11628.

121. Buck, M. and T.R. Hoover, *An ATPase R-Finger Leaves Its Print on Transcriptional Activation.* Structure, 2010. **18**(11): p. 1391-1392.

122. Scott, A., et al., *Structural and mechanistic studies of VPS4 proteins.* Embo Journal, 2005. **24**(20): p. 3658-3669.

123. Ishijima, S., et al., *Light-induced increase in free Mg2+ concentration in spinach chloroplasts: Measurement of free Mg2+ by using a fluorescent probe and necessity of stromal alkalinization.* Archives of Biochemistry and Biophysics, 2003. **412**(1): p. 126-132.

124. Spreitzer, R.J. and M.E. Salvucci, *Rubisco: Structure, regulatory interactions, and possibilities for a better enzyme.* Annual Review of Plant Biology, 2002. **53**: p. 449-475.

125. Dawson, R.M.C., *Data for biochemical research*. 3rd ed. 1986, Oxford: Clarendon Press. xii, 580 p.

126. Brown, P.H., A. Balbo, and P. Schuck, *Characterizing protein-protein interactions by sedimentation velocity analytical ultracentrifugation.* Curr Protoc Immunol, 2008. **Chapter 18**: p. Unit 18 15.

127. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling.* Biophys J, 2000. **78**(3): p. 1606-19.

128. Balbo, A., et al., *Measuring protein-protein interactions by equilibrium sedimentation.* Curr Protoc Immunol, 2007. **Chapter 18**: p. Unit 18 8.

129. Lebowitz, J., M.S. Lewis, and P. Schuck, *Modern analytical ultracentrifugation in protein science: a tutorial review.* Protein Sci, 2002. **11**(9): p. 2067-79.

130. Schuck, P., et al., *Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems.* Biophys J, 2002. **82**(2): p. 1096-111.

131. Dam, J. and P. Schuck, *Calculating sedimentation coefficient distributions by direct modeling of sedimentation velocity concentration profiles.* Methods Enzymol, 2004. **384**: p. 185-212.

132. Ericsson, U.B., et al., *Thermofluor-based high-throughput stability optimization of proteins for structural studies.* Anal Biochem, 2006. **357**(2): p. 289-98.

133. Wang, D. and A.R. Portis, Jr., *Two conserved tryptophan residues are responsible for intrinsic fluorescence enhancement in Rubisco activase upon ATP binding.* Photosynth Res, 2006. **88**(2): p. 185-93.

134. Story, R.M. and T.A. Steitz, *Structure of the recA protein-ADP complex.* Nature, 1992. **355**(6358): p. 374-6.

135. Weibezahn, J., et al., *Characterization of a trap mutant of the AAA+ chaperone ClpB.* J Biol Chem, 2003. **278**(35): p. 32608-17.

136. Zhang, N., P. Schurmann, and A.R. Portis, Jr., *Characterization of the regulatory function of the 46-kDa isoform of Rubisco activase from Arabidopsis.* Photosynth Res, 2001. **68**(1): p. 29-37.

137. Senior, A.E., S. Nadanaciva, and J. Weber, *Rate acceleration of ATP hydrolysis by F(1)F(o)-ATP synthase.* J Exp Biol, 2000. **203**(Pt 1): p. 35-40.

138. Frasch, W.D., *The participation of metals in the mechanism of the F(1)-ATPase.* Biochim Biophys Acta, 2000. **1458**(2-3): p. 310-25.

APPENDIX A

STATEMENT OF PERMISSION FROM CO-AUTHORS

STATEMENT OF PERMISSION FROM CO-AUTHORS

Co-authors on the previously published article “Protein Oligomerization Monitored by Fluorescence Fluctuation Spectroscopy: Self-Assembly of Rubisco Activase” and article that is in preparation: “ATP and magnesium promote rubisco activase hexamer formation at low micromolar concentrations” have granted their permission for use of the articles in this dissertation. This list of co-authors includes: Chakraborty, M., Henderson J.N., Salvucci, M. E., Wachter, R. M., Levitus, M., Hazra S., Serban A.