Effects of the Terminal Regions on Rubisco Activase Form and Function.

William Finnis.
Supervisor: Dr F. Grant Pearce.

Table of Contents.

Acknowledgements.

Firstly I would like to extend my gratitude to my supervisor, Dr F. Grant Pearce for his time and contribution educating me on the laboratory equipment and general biochemistry. Always ready to help out if need be and provide plenty of criticism that although was not always welcome, was always relevant. In the one year Dr Pearce seamlessly juggled 3 masters students, one new baby, 3 publications and building a new life in Australia. Special thanks to Dr Ren Dobson for being my co-supervisor and inviting me along to his lab meetings where he would routinely purchase me a beer.

I would like to thank my work mates in room 630; James Davies, Hamish Cleland, Cameron MacDonald and Serena Watkin. They were always there to discuss problems with and help me out when I required it. More realistically however they were to have a laugh with, generally when things weren't working out in the laboratory.

I would like to thank my girlfriend Erica Boyd for being just about the best thing any masters student could wish for, providing endless support that was especially helpful whenever I become stressed or fatigued.

Lastly I would of course like to thank the University of Canterbury for the privilege of studying as a masters student the beautiful SBS building.

Abstract.

Rubisco is the sole enzyme responsible for the incorporation of CO_2 into sugar molecules useable by all life on earth. In the vast majority of discovered autotrophic organisms their respective rubisco has been accompanied by another enzyme called rubisco activase. The catalytically inefficient nature of rubisco requires rubisco activase yet much of its important activity remains undiscovered.

Rubisco activase has at both of its terminal ends a dynamic and flexible domain that have poorly understood functional consequences. This projects aims are to gain a better understanding on how these terminal domains are required for rubisco activase.

Various rubisco activase constructs were recombinantly purified and experimented on. The experiments included various kinetic assays and thermal melt temperatures but also powerful techniques such as analytical ultracentrifugation and small angle X-ray scattering.

The results show that the rubisco activase enzyme produces dynamic changes to its assembly and function with the binding of nucleotide and the addition and subtraction of terminal components.

Abbreviations.

Rubisco - Ribulose-1,5-Bisphosphate carboxylase/oxygenase
Rca - Rubisco activase
ATP - Adenosine Triphosphate
ADP - Adenosine Diphosphate
RuBP - Ribulose-1,5-Bisphosphate
XuBP - Xylulose 1,5-Bisphosphate
Ca1P - 2-Carboxy-D-arabinitol 1-phosphate
PdBP - Pentodiulose-1,5-bisphosphate
RcbL - Rubisco large subunit
RcbS - Rubisco small subunit
T-EM - Transmission electron microscopy
C-EM - Cryogenic electron microscopy
AUC - Analytical Ultracentrifugation
SAXS - Small Angle X-ray Scattering
DSF - Differential Scanning Fluorometry
SDS-PAGE - Sodium dodecyl sulfate polyacrylamide gel electrophoresis.
IEX - Anion exchange chromatography
SEC - Size exclusion chromatography
SA - Specific activity.
NADH - Nicotinamide adenine dinucleotide hydride
LDH - Lactate dehydrogenase
PK - Pyruvate Kinase

Regularly used in figures to denote Rca constructs.
T-RV - Tobacco R294V
T-RV nt - Tobacco R294V N-terminal truncated
Tα - Tobacco alpha extension
Tob - Native tobacco Rca

Chapter 1: Introduction.

1.1.a Overview

Photosynthesis is critical to virtually all life on earth as it is the capture of the sun's energy for use in the biochemical processes that constitute living. The photosynthetic process, from light and carbon dioxide (CO_2) to sugar and oxygen (O_2), occurs in many discrete metabolic steps. The first two divisions that can be made is the light harvesting component that uses chlorophyll and membrane proteins to capture lights energy in molecular form, aptly named the light dependant stage. Secondly is the light independent stage where the accumulation of atmospheric (CO_2) into sugar molecules takes place through a cyclic series of enzymatic reactions called the Calvin-Benson-Bassham cycle. At the core of the CO_2 fixing operation is Ribulose-1,5-Bisphosphate carboxylase/oxygenase (rubisco).

Rubisco, at its central position of plant food metabolism, is seemingly unproficient at its job. Besides a slow catalytic rate, it has a habit of inhibiting itself with the creation of misfire products [1,2] and undergoing photorespiration [3,4]. Photorespiration is accidental binding of O_2 instead of CO_2 in the active site for the reaction with RuBP. The result is the creation of the toxic compound, phosphoglycolate which require energy to be transported between the cytosol, peroxisome and mitochondrion to be recycled, creating reactive hydrogen peroxide in the process. This activity, along with rubiscos other catalytic problems, is widely thought to be a result of rubisco having evolved into a “dead end” during the early mileina where there was very little oxygen in the atmosphere. Due to its unproductive catalytic rate and important central position in plant metabolism, rubisco is often produced in relatively massive amounts to other proteins and can constitute up to 50% of the soluble protein of a leaf [5]. Because of this and the abundance of photosynthetic plants, algae and other organisms, rubisco is widely considered the most produced protein on the planet [5].

Rubisco requires a cohort of ancillary proteins for formation and upkeep, one of these is Rubisco Activase (Rca). Rca is required for the prolonged integrity of the rubisco enzyme, as the name implies it activates rubisco from its inhibited states. In plants modified to express lower amounts of Rca there was a parallel decrease in the plants carboxylase activity [6].

1.1.b Carbon dioxide metabolism.

There exists a diverse array of plant life that have some differences in the way CO_2 is metabolized. These differences have been classified as C3, C4 and CAM carbon fixation. C3 plants make up the majority of the world's plant biomass and require the diffusion of CO_2 into the cell. C4 plants instead actively pump CO_2 into the photosynthesizing cells by binding it onto phosphoenolpyruvate to create the 4 carbon molecule (hence the name) oxaloacetic acid. Oxaloacetic acid is then converted into malate or aspartate which are transported to the chloroplast where CO_2 can be enzymatically released from these compounds creating a high local CO_2 concentration. All C4 plants are angiosperms and include many species of grass and the crops maize and sugarcane. CAM (crassulacean acid metabolism) plants occupy arid regions where more water saving methods have evolved. One such way is to close the stomata meaning that more moisture is retained, yet carbon dioxide cannot diffuse in. To get around this, CAM plants open their stomata at night when it is cool and store uptaken CO2 as malate in vacuoles. During daytime, the malate is transported to the chloroplast where it can release CO_2. Once again the high local CO_2 concentration leads to a greater photosynthetic efficiency. However across all this diversity, Rca is required for the continued function of rubisco and therefore the plant. Improved Rca and rubisco provides a window to universally increase a plant's photosynthetic ability and CO_2 consumption.

1.1.c Improving rubisco for improved crops.

Practical applications of studying the interaction between rubisco and Rca involves increasing the productive ability of crops due to increasing the catalytic ability of rubisco. The carbohydrates created by the calvin cycle are used for energy and biomass by the plant and it is hypothesised that increasing the catalytic ability of rubisco will have positive effects on the plants growth due to the slow rate of rubisco being a large limiting step to growth. The perceivably flawed rubisco has many attempts at improving its catalytic ability with differing levels of success [7]. One of the most viable ways to improve its ability is to place different species rubisco into plants. A genetically modified plant that had a rubisco from a cyanobacterial shown to have faster CO_2 fixation demonstrating that the technology exists to improve biomass accumulation [8]. In relation to Rca, its job maintaining rubisco at 100% efficiency is thought to be improvable and research has been conducted on that premise [9]. It has been observed that the carboxylase activity of plants reduces at higher temperature, the reason for this is debated but many have pointed out the low denaturation temperature of Rca (as low as 25°C) as the culprit [10]. This has made

thermostability one sort after feature of engineered or transgenic Rca. It also has been shown that some mutations to Rca actually increase the rate of activase activity [11,12] leading researchers to believe that a plant with improved Rca activity is a viable creation.

1.2 Rubisco.

1.2.a Structure and function.

Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) fixes atmospheric CO_2 onto ribulose-1,5- bisphosphate (RuBP) in the light independent phase of photosynthesis. Rubisco inhabits the stroma of chloroplasts and does not directly interact with the light harvesting process occurring on the thylakoid membrane. The ATP and NADH created in the light dependent phase are used to to cyclically regenerate RuBP required for Rubiscos carbon fixation catalysis. The 5 carbon RuBP and CO_2 is converted into an unstable 6 carbon sugar that rapidly decays into two 3 carbon sugars of 3-phosphoglycerate (3-PG). From this point, 3-PG can be transferred into the glycolysis metabolic pathway or cyclically metabolised to form RuBP again in the calvin cycle [13]. Rubisco exhibits little diversity for the large span of organisms that use it. Demonstrating either that it has evolved into a dead end early on or that it exhibits the catalytic limits of enzymes. All rubisco is broken down into two forms and numerous types. The structure of form 1 type B rubisco used by terrestrial land plants has what is generally described as a barrel structure composed of eight large subunits (RbcL,~50KD) arranged antiparallel relative to their neighbour. The small subunits (RbcS,~15KD) are assembled into two tetramers that cap the ends of the barrel as shown in figure 1. Form 2 found in dinoflagellates and some bacteria consists only of RcbL in either dimeric or oligomeric forms.

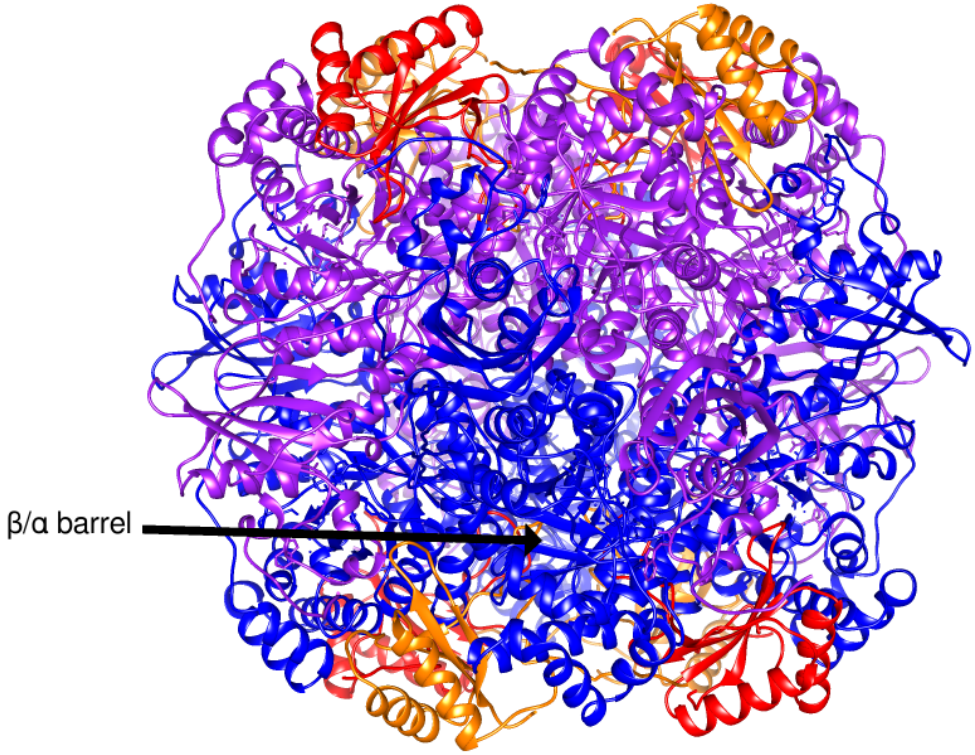


Figure 1: Structure of spinach rubisco (1RCX) large subunits in purple and blue, small subunits in red and orange. The position of the β/α (TIM) barrel is indicated.

RcbL consists of N-terminal α+β domain of ~150 amino acids continuing onto the ~320 residue triosephosphate isomerase (TIM or β/α) barrel comprising the C-terminal end. The active site resides at to top of the TIM barrel where upon RuBP binding to active rubisco, loop 6 of the barrel folds over and locks the active site. The loop is then further cemented in place by the C-terminal end which binds to the folded loop locking out all solvent access [14]. Rubisco has eight active sites that occur at the interface of two antiparallel subunits where 'complementary' N and C-domain residues reside. In all eukaryotes the rbcL gene is located on the plastid genome and the rbcS gene is almost always located on the genome with the exception of non-green algae where it is also plastid encoded [15].

1.2.b Activation.

Rubisco has a convoluted activation process. Positioned at the active site is a lysine residue that blocks substrate binding and activity of the enzyme. This highly conserved lysine (201 on RcbL in spinach) residue has to be carbamylated on its terminal amine by non-substrate CO_2 to allow structural changes to the enzyme. Carbamylation allows the introduction of the Mg^{2+} coenzyme to bind in the active site

to the carbamylated lysine, once bound the enzyme is active [14]. At multiple points along the enzymes activation sequence it can be inhibited. There is terminology used to to describe the interactions of rubisco with its coenzymes and inhibitors. E for the rubisco enzyme, C for carbamoylated lysine, M for bound metal ion and I for inhibitor as demonstrated in figure 2.

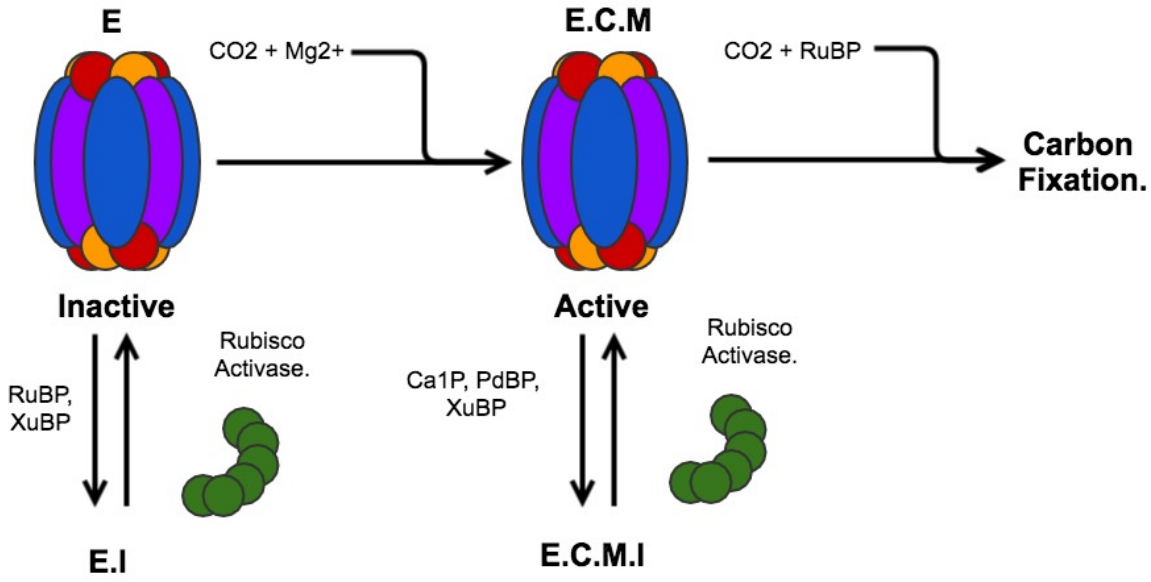


Figure 2, Rubisco activation pathway and the role of rubisco activase in it.

The catalytic activity of rubisco creates multiple reactive intermediates during the many steps it undertakes. It is common for rubisco to create misfire products due to the protonation and oxygenation of these steps [2]. When this occurs, rubisco forms a closed state that does not reopen or reopens very slowly and so enters the need for Rca which is required for the continued function of rubisco.

1.3 Rubisco activase.

1.3.a An introduction.

Rubisco activase (Rca) was first discovered by examining a mutant *Arabidopsis* plant that could only grow in conditions with high levels of CO_2 [16]. It was found that Rca would stop rubisco “fallover” the term for the gradual reduction of catalytic rate over time [17]. Sequence studies revealed that it had a fold classifying it as an ATPase Associated with diverse cellular Activities (AAA+). The AAA+ fold consists of 2 central domains of ~230 residues referred to the αβ (large) and α (small) regions

corresponding to the secondary structure make up of the domain exhibited in figure 3. The larger αβ domain contains the canonical walker motifs that are associated

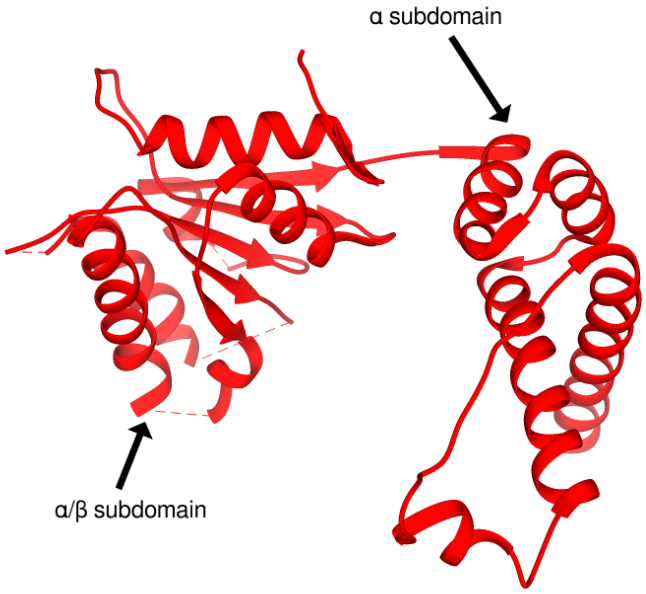


Figure 3: Structure of tobacco Rca (3ZW6) demonstrating the domains. The C and N-Terminal domains are not present in the structure.

with ATP binding. Between the two domains exists a short flexible linker allowing for the two subdomains to maneuver and reposition in relation to each other. External to the core AAA+ domain are the terminal domains which have, in Rca, proven difficult to image, generally they perform tasks that complement the activity of the AAA+ domain [18]. The α domain has a sensor motif consisting of an important α helix called helix 9 that has proven important for Rca interactions with rubisco [19]. As the name implies, AAA+ enzymes are found in a diverse array of processes and require ATP for their function. This family includes clpX and FtsH which unfold proteins for degradation and helicase that parts complementary DNA strands for replication replication. The general description of the activity of AAA+ proteins is the unfolding or the rearrangement of molecules for degradation, maturation, transport and replication [18]. In Rca, nucleotide binding effects subunit assembly and when ATP is bound it exhibits a basal rate of ATP hydrolysis. The rate of ATP hydrolysis is allosterically regulated by the adjacent protein, increasing with protein concentration and quaternary structure [20,21]. It was thought a conserved arginine 293 (294 in tobacco) residue on the small domain enters the neighbouring subunit and

hydrolyses the terminal phosphate of bound ATP. This was show to be false and this invasive residue allosterically modulates the surrounding structure to perform the hydrolysis [19]. The conformational changes caused by ATP hydrolysis are thought to be utilized for the mechanical action of the protein, similar to other AAA+ proteins. Rubisco activase mechanism for the removal of binded inhibited activation of rubisco is still considered to be unknown.

1.3.b Quaternary Structure.

The quaternary structure of AAA+ proteins are generally either hexameric, forming a ring with a pore through the middle, or form spiraling oligomers, the structure of *in vivo* active Rca is contested between these two. The first observed Rca with the whole AAA+ domain intact with the N and C terminal truncated was tobacco Rca. This crystallized as a hexamer though in an open helical ring rather than a closed one, known as the $P6_5$ space group [12]. Despite this it was still generally believed that Rca had a hexameric quaternary structure for activity due to other evidence (mainly the previously discovered cbbX covered in next segment). Figure 4 is the crystallized and imaged tobacco Rca modeled as a hexamer. The hexameric structure does not account for the observed results where Rca can form structures of up to 16 monomers and that the rubisco reactivation activity can peak at 1 uM corresponded to assembly of 3-5 Rca monomers [21,22]. Rca is enigmatic in that the quaternary structure can differ between species and analogues within that species and mutants that exist as hexamers after one mutation. Tobacco Rca forms oligomers but the mutant R294A was found to form stable hexamers with a large reduction to its ATPase rate and unable to reactivate inhibited rubisco. The construct R249V however retained the ATPase and activase activity while forming

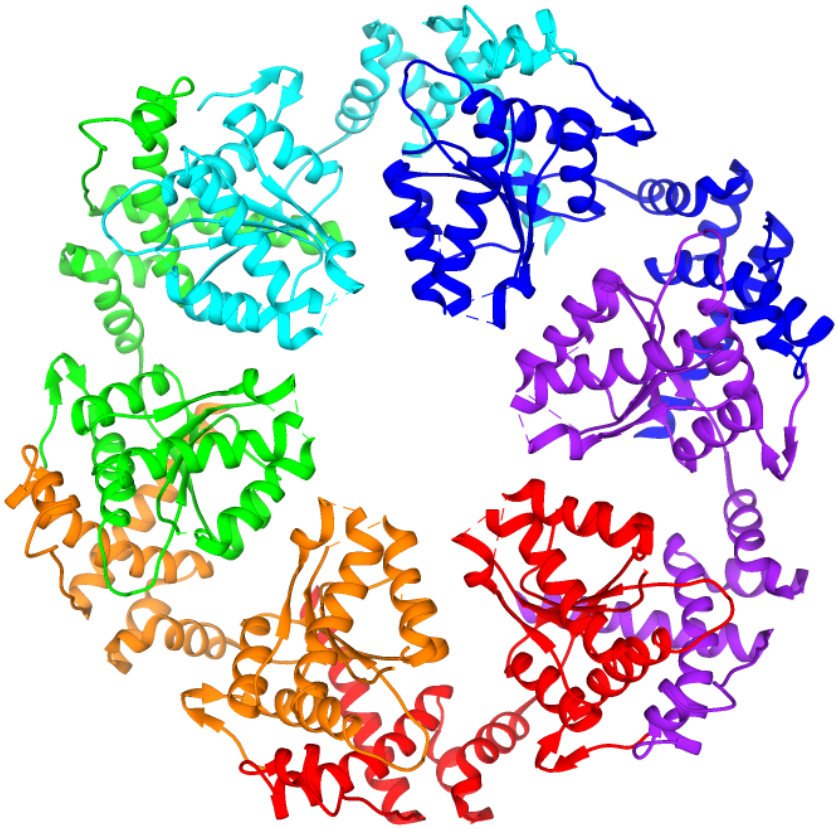


Figure 4: Tobacco Rca (3ZW6) shown as a hexamer, again the C and N terminal regions are not shown. The Rca was modeled into a hexameric formation after crystallizing as an open helix.

stable hexamers [12]. It wasn't until 2014 that a wild type activase was observed to exist as a hexamer in solution in spinach long form and more transiently in cotton short form [23,24]. The spinach alpha or long form that on its own and up to one half mixture long/short formed hexamers in the presence of ATPγS, the spinach short form was observed as a range of oligomers. Even as hexameric forming constructs, the Rca enzymes of tobacco R294V and spinach long form had ATPase and reactivation activity in what was shown to be averaging between dimers to trimers [23]. Interestingly, the the concentrations at which Rca is shown to be at full ATPase and activase activity at around 2 μM are vastly under the observed concentrations that exist in the stroma of the chloroplast at 170 - 500 μM [25]. It is unknown what the quaternary structure or Rca is at its high native concentration, as well as other aspects of its function.

1.3.c Rubisco activation.

At the core of most Rca research is a desire to reveal how Rca interacts with and activates its substrate, rubisco. A large advancement was made in 2011 when it was demonstrated that an Rca analogue from red algae called cbbX likely activated red algae rubisco through local unfolding to release inhibitors [26]. The red algae RcbL has C-terminal “tails” that when removed, reduced cbbX’s ability to reactivate the rubisco. The conclusion was that cbbX activates the red rubisco by streaming these “tails” through the hexamers central pore causing unfolding the release of inhibitors. Rca is thought to act in a similar way though specific details are missing. It is currently believed that Rca activates rubisco through local unfolding and the release of inhibitors via pulling external regions through the central pore. While the points of contact between the Rca and rubisco are unknown, some vital residues have been shown to be important but a picture of the residue-residue interface is lacking (see Species recognition). It is debated what part of the rubisco enzyme it translocates through the pore if this is the case. The candidates are the C-Terminal end of the large subunit which folds over into the active site during the reaction, or one of the αβ loops belonging to the TIM barrel that contains the active site. While the C-terminal end was the target for the red algae cbbX, green-type RbcL from terrestrial plants lack the extended C-terminal “tail.” [27]. Rca problematically lacks the pore motifs seen on other protein threading AAA+, showing that the mechanistic action it undertakes for reactivation could be wildly different however mutations to aromatic pore residues had drastic negative effects on activase activity [12]. Which face Rca engages rubisco is also unknown although residues on helix 9 of the α subunit have been shown to be important for rubisco reactivation and one face of Rca has shown to be more conserved through all species than the other face [12,19].

1.3.d Motor protein.

AAA+ are known for their mechanistic action and are known as molecular motors [18]. Currently there is no reason not to discard a hypothesis that Rca also functions as a motor protein and the discovery of the mechanistic action of cbbX only reinforces this. A single molecule experiment [28] on the AAA+ protein threading enzyme clpX showed that it bound to the target protein and translocated it through the pore after several ATP hydrolysis events. ATP hydrolysis was required for local denaturation, most ATP hydrolysis events were not successful. Upon success and local unfolding the translocation was instantaneous not requiring ATP hydrolysis. The paper concluded that although brownian motion could not be ruled out entirely contributing for translocation action, they concluded that it is highly likely the result of

mechanical motion due to the fact it worked against an applied force in the opposite direction and that clpX protein was in close contact with the target protein during the experiment.

1.3.e Crystallization and imaging.

Attempts to crystallize and diffract Rca have only yielded the central AAA+ domain to be captured. The dynamic N and C-terminal ends appear to disrupt crystal formation therefore the most complete structures we have of Rca have the N and C-terminal domain truncated which causally effects on the packing and arrangement of the crystal structure. Rca generally crystallizes as an open hexamer in the $P6_5$ space group. A 97 residue long section of *Larrea tridentata* (creosote bush) Rca was the first crystallized with the intent of understanding resists temperature denaturation due to the plants natural habit [19]. The structure showed an elongated C-domain (α-subdomain) about 50 Å long, which when compared to the AAA+ protein FtsH was shown to extend outwards longer than the classical AAA+. Due to its high resolution of 1.9Å, the paper goes onto describe the amino acids that interact creating the α-helix heavy structure. The selective residues lysine-313 and valine 309 were shown to be on the end of this elongated paddle and hypothesised that 313 does not make contact with rubisco due to a disruption of an internal salt bridge. It was in this study that it was concluded that Arg-293 does not interact with the bond nucleotide due to its distant location from the active site. A mostly complete cbbX from *R. sphaeroides* was the next structure to be published. It managed a resolution of 3.1 Å with the omission of 8 residues at the C-terminal, helped by the fact that it does not have an N-terminal domain [26]. The RuBP binding site was discovered distal to the ATP binding site due to bound sulphate ions inhabiting a cleft on the α subdomain. Transmission electron microscopy images revealed a hexameric structure in the presence of RuBP and ATP but long fibrillar structure with just ATP. The structure of cbbX, which was shown to stream a polypeptide through it, had a central pore 25 Å across. The crystal structure of tobacco Rca was published soon after with a truncated N-terminal domain and the the C-terminal extension was missing density [12]. The structure had a resolution of 2.95 Å which was commented on being due to the low thermal stability of Rca demonstrated by a high Wilson B factor of 113 $Å^2$. The structure crystallized in the $P6_5$ space group but due to other experimental evidence and the active hexameric state of cbbX they concluded that is exists as a hexamer for activity. Tobacco R294V was discovered to have similar ATPase rates and activase activity and form stable hexamers imaged using a transmission electron microscope (TEM). The crystal model (made hexamer) and TEM were overlaid and

generally confirmed that they were similar, except for a empty electron density on one face of the the TEM model which was concluded to be the missing N-terminal region not on the crystal structure. This electron density was on the same face of the Rca as the important helix 9 containing the species selective residues. Both models had a pore size ~36 Å across and was hypothesized that it was larger than the singular threading cbbX pore (of ~25 Å) and possibly threaded the βα loop of RcbL through the pore rather than the C-terminal end. It is likely the protein had adopted an non-native conformation in the crystallization process due to the truncations, mutations and modeling. Small angle X-ray scattering (SAXS) uses similar principles to X-ray crystallography the major difference is that crystal is not needed and the protein remains solubilized in water and does not have to be truncated to form crystals therefore can be considered to be closer to its native environment. SAXS was performed on tobacco Rca and generated models showed a pore size dissimilar to the one shown in the crystal tobacco Rca and closer in size to the cbbX enzyme which threads the C-terminal end [21]. Rca from *Arabidopsis thaliana* (Thale cress) was later crystallized and imaged [29]. It also crystallized in the $P6_5$ space group and required the truncation of 9 residues at the C-terminal end. Interestingly the the N-terminal domain also crystallized but was to dynamic to be captured by the diffraction pattern, it was however shown to be present in the crystal by mass spectrometry. This Rca from *Arabidopsis,* in the same space group as the tobacco Rca had differences in their crystal packing. The length of the sixfold screw axis differed majorly between them with the tobacco Rca packing into about half that of the Thale cress Rca.

1.3.f Rca Sequence and Isoforms.

Two isoforms of Rca are present in the stroma in most plants, The are termed the alpha (α) or long isoform and beta (β) or short isoform. The differences between them is that the α isoform has an extra ~30 residues called the C-terminal extension on the C-terminal domain shown on the domain layout in figure 5. This usually occurs due to alternative splicing of a single gene after transcription [30] but also due to the presence of two or more Rca genes [31]. Some species produce only the β isoform, such as Tobacco which interestingly has 3 separate Rca genes [32]. The stoichiometry of α and β activase is generally in favour of the β -isoform but 1:1 ratios are also observed [25]. Rubisco activase is regulated by ATP/ADP ratio inside the cell. During the day/night cycle the levels of ATP/ADP ascend and descend due to the production by the light dependent phase of photosynthesis. The high level of

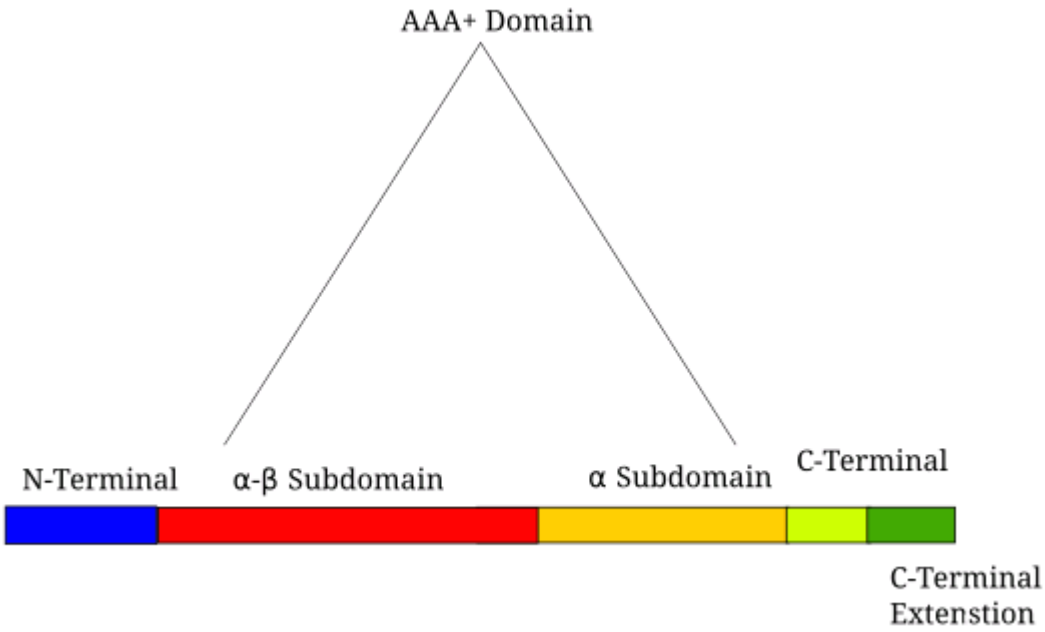


Figure 5: The domain layout of rubisco activase.

ADP slows down the whole calvin cycle due to lack ATP and also due to the inhibiting ability ADP has on Rca and especially the oxidized form of α activase [27] (see C-terminal).

1.3.g Species recognition.

Some rubisco activase can activate rubiscos from other species but usually at a decreased rate however there exists a divide between plants from the family solanaceae and plants that are non-solanaceae. Isolating the residues involved with species diversity recognition would allow a picture of how the Rca interacts with rubisco. An important step was the discovery of how to interconvert the species selectivity between the two [33]. Solanaceae is an important family of angiosperms that vary largely from trees to vines. Many important crops are in this family including tomatoes, potatoes and tobacco. It was discovered by mutational analysis that residues in positions 89 and 94 (residue diversity amongst species) are important for for Rca to rubisco interaction. Mutations to *Chlamydomonas reinhardtii* RbcL in these positions (P89R and D94K) made the rubisco susceptible to activation by solanaceae Rca. These residues reside distant to the active site by 30 Å and hypothesized that conformational changes to this region determined the selectivity. On the Rca protein itself, it was revealed that residues in position 313 and 316 of helix 9 of the α subdomain were important for rubisco recognition. The residues at 313 where a positively charged lysine in spinach (non-solanaceae) and a negatively charged

aspartic acid in tobacco (solanaceae) while they both had nonpolar residues in position 316, valine in spinach and leucine in tobacco. The small helix 9 is shown in Figure 6 positioned above the neighbouring subunit along with the highly conserved αβ subdomain lining to the central pore. The α subdomain is located behind the neighboring subunit containing multiple instances of sequence diversity. The residues at position 313 were demonstrated to be the primary species selectivity pivots while swapping the 216 residues merely improves upon the intra-species activation of Rca when the 313 mutation exists [34]. The high resolution crystal structure concluded, due to structural limitations it is likely the 313 mutations cause conformational adjustments that confer to selectivity and recognition rather than residue interactions with rubisco [19].

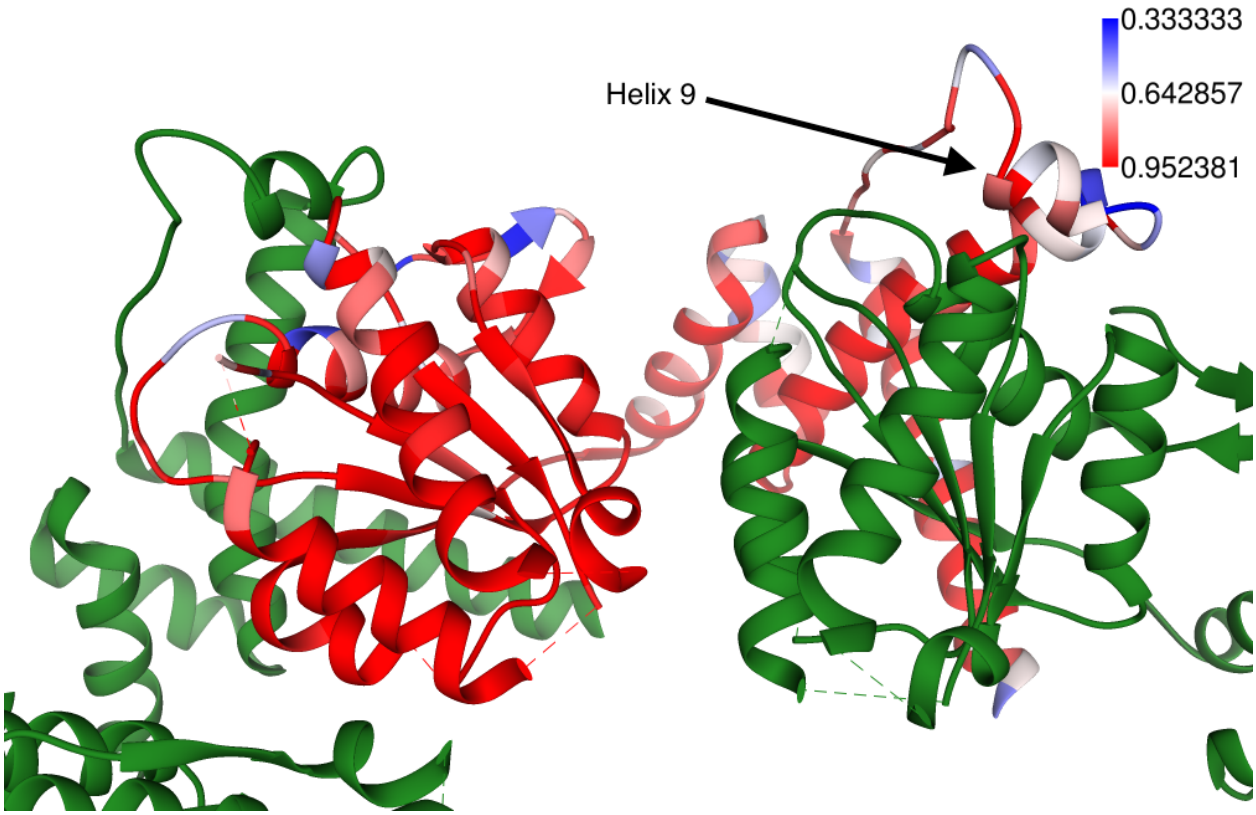


Figure 6: a structural overlay of the AAA+ domain displaying the conserved residues of 19 Rca proteins, Red representing low diversity to blue representing diverse regions. The important Helix 9 is indicated.

1.3.h Rca analogues.

All autotrophic organism found on earth require carbon based sugars and therefore, rubisco. Other proteins have been discovered to activate rubisco similar to Rca and all have the same AAA+ protein fold. It is hypothesized that this a result of convergent evolution, rather than divergent. Rubisco from red algae and bacteria were shown to not undergo fallover as green rubisco from terrestrial plants and green algae, suggesting that they didn't require a rubisco activase [2]. It was shown in

Rhodobacter sphaeroides however that a gene labeled cbbX encoded a red algae rubisco activase [26]. CbbX was shown to have a AAA+ fold like green-Rca yet differed in the terminal domains having a shortened N-terminal extension unique to cbbX and no C-terminal domain. Cbbx did not exhibit basal ATPase activity with just ATP but required the addition of RuBP binding to a site distant from the nucleotide. Using TEM it was shown to have a dynamic quaternary structure, exhibiting oligomeric chains in the presence of ATP and closed hexamers with ATP and RuBP. It lacked any discernable quaternary structure in just RuBP or neither. The RuBP binding site was shown to be located in a conserved cleft on the α subdomain. A third separately evolved rubisco activase class was then discovered in *Acidithiobacillus ferrooxidans* [35]. This autotroph would obtain energy from oxidizing inorganic compounds such as sulfur and ion as its "light dependent phase" then use the energy obtained to fix carbon which therefore requires a rubisco enzyme. *A. ferrooxidans* has two rubisco isoforms within its genome corresponding to form I and form II rubiscos. Both have two open reading frames downstream of them called cbbO and cbbQ. CbbQ encodes a small AAA+ protein comprising of just the AAA+ domain with no notable extensions at the terminal regions. It displayed a hexameric quaternary structure. CbbO is a new addition to the rubisco auxiliary proteins, demonstrated to singularly bind to the cbbQ hexamer to form the $(cbbQ)_6(cbbO)$ holoenzyme which acted as a rubisco activase for *A. ferrooxidans* rubisco. CbbO appears to help the hexameric AAA+ $cbbQ_6$ adhere to rubisco.

1.4 Rubisco activase Terminal Domains.

1.4.a N-terminal domain.

The N-terminal domain is ~70 residues in size and has been shown to be very dynamic. In crystallography it is either truncated for crystal forming purposes or not captured in the diffraction pattern. The domain sticks radial outwards of the center of symmetry but due to their dynamic nature have not been imaged to a good degree. This domain movement is only slightly characterized, only descriptive but not quantitative conclusions can be drawn. In 2011 when [12] overlaid the T-EM structure of tobacco R294V and the N and C-terminal truncated crystal structure they noticed an area of electron density missing in the crystal structure that was present in the T-EM density map. This missing density was shown to be on one face of the Rca hexamaer and protuding out the side. In [23] using rigid body modeling software, CORAL, the N-terminal was demonstrated to extent out in front and sideways of the central domains in the same fashion as the T-EM structure.

The N-terminal domain has been proven to be vital for the interactions between rubisco and Rca. Mutations in the region of the residue tryptophan 16 heavily reduces the rubisco reactivation rate [11,36]. Truncations of the N-terminal domain rarely have any effect on ATPase rate but have catastrophic effects on the activase rate when deleting much more than ten residues [11,12]. Otherwise the role of the N-terminal domain remains a mystery, it has been termed a “molecular lasso” [23] where its highly dynamic nature could allow it to recruit rubisco and adhere to it while it performs the structure unraveling similar to the recently discovered cbbO.

1.4.b C-terminal Domain.

The C-terminal domain is present in all Rca of green plants and algae. Most plant species have two analogues of Rca termed alpha or long form and Beta or short form where the long form has an extra 22-39 Residues at the C-terminal end termed the C-terminal extension. As covered before, The α and β isoforms generally are the result of alternative splicing of the same gene, others however have 2 genes encoding the two isoforms [30,31]. This C-terminal extension is redox sensitive due to two conserved cysteine residues (392 and 411 in spinach/*arabidopsis*) demonstrated to sensitise Rca to oxidizing conditions for regulation. The oxidation of these two cysteine residues has been shown to increase the Rca susceptibility to be inhibited by ADP, resulting a loss of ATPase and activase activity for regulation [37]. Some plants like tobacco, tomato and some algae produce no alpha isoform instead their beta isoform is more sensitive to inhibition by ADP than species that produces both [38]. A model has been constructed that suggested that the negatively charged C-terminal region once oxidized strucraly chnages to form a loop. This loop is able to bind to a positively charged site near the active site, producing allosteric effects to the activity and possibly other aspects [39]. Interestingly one of the only native plant Rca shown to exhibit a closed hexameric quaternary structure is the Spinach alpha isoform in the presence of ATP [23]. A study was conducted where deletions were made to the C-terminal end of spinach rubisco short form and showed that the ATPase rate was not affected and the rubisco activation rate actually doubled [11]. The General consensus is that the C-terminal domain is used in regulation. Though more is known of the function of the C-terminal than the N-terminal it is still elusive to image and it is not known for sure how it functions to regulate Rca.

1.5 Aims

1.5.a Hypothesis.

The goal of this research is understand how the terminal regions of rubisco activase affect the protein's interaction with itself to perform its biological function. This follows on from a line of research with the ultimate goal of understanding how Rca interacts with itself and rubisco to perform its biological function. The hypotheses entering into this research is changes to the N and C-terminal will have no effect on the ATPase rate and quaternary structure of the Rca construct.

1.5.b Approaches.

Many of the constructs tested in this research are very similar or identical to constructs previously studied. ATPase rate and quaternary structure have proven to be variable to even within species, such as hexameric spinach alpha and oligomeric spinach beta [23]. Therefore some of the constructs studied may actually have surprisingly different effects on the activity of the enzyme. Much of the research done also did not use or have some of the techniques that will be used in this research including; Small Angle X-ray Scattering (SAXS), Analytical Ultra-Centrufugation (AUC) and before long, Cryogenic Electron Microscopy (C-EM). These techniques allow for new information to be obtained from these Rca constructs. AUC allows the molecular mass to be obtain for a species in solution allowing an understanding of the quaternary structure under different conditions. SAXS and C-EM also provide information on the quaternary structure but can also be used to help understand the positioning of the subunits as these techniques actually generate images to a limited resolution. Although not actually used in this study C-EM is a upcoming powerful way of imaging proteins. This study has an additional focus on stabilizing the dynamic nature of Rca in both secondary structure and quaternary structure for imaging purposes with C-EM. This is the reason why the R294V mutation is used with the N-terminal truncated construct to observe if it still forms stable hexamers. The constructs used in this study will be constructs of tobacco as tobacco Rca was more reliable to produce than spinach (explained in results). The before described tobacco R294V will be one construct along with tobacco R294V N-terminal truncated which has the first 70 Residues deleted. The C-terminal experimental construct will be the

addition of the redox sensitive C-terminal extension from spinach onto (non-R294V) tobacco Rca which does not occur naturally.

Chapter 2: Methods.

2.1.a Genes.

The Rca genes were bought from Epoch Life Sciences encoded into pET-11a plasmids which are then transformed into BL21 competent *E. coli* with the pLysS plasmid. The Rca containing *E. coli* were stored in LB/glycerol stocks at -80C°.

2.1.b Purification.

Rca constructs were purified according to [21] or, when possible, substituting poly-His tag affinity chromatography for anion exchange chromatography (IEX) described [40]. In our case the plasmid encoding Tobacco R294V had a N-terminal polyhistidine tag and therefore could use affinity chromatography. The process involves transferring the frozen plasmid containing *E.coli* into an autoclaved 10 ml starter culture of LB containing the correct antibiotics at 1mM concentration (Ampicillin for pET-11a and chloramphenicol for PlysS.) The starter culture is left overnight incubation at 37°C upon a shaker at 180 rpm. The starter culture is then transferred to autoclaved M9ZB media which is used due to the fact it can support a higher concentration of bacteria.

Table 1: M9ZB media

NH_4Cl	KH_2PO_4	$NaHPO_4$	NaCl	Bactotryptone	Yeast-Extract
1 g/L	4 g/L	6 g/L	5 g/L	10 g/L	5 g/L

The M9ZB media is made in 600mL batches in 2L baffled conical flasks and autoclaved. At the time of transfer, a 40% glucose solution is added to a final concentration of 0.4M glucose and a 1M MgSO4 solution is added for a final concentration of 1mM. The correct antibiotics are also added to 1 mM concentration. The starter culture is added at a volume of 1-0.5% of the volume of the media and placed on a shaker at 37°C and allowed to grow. For induction, the OD600 can be measured and induction done at a OD of 0.6 but generally the method involved induction after 4 hours of growth. At this point the media is transferred to a shaker (180 rpm) and incubated at 27°C for an hour to cool. Isopropyl β-D-1 thiogalactopyranoside (IPTG) is then introduced to a concentration of 0.4 mM and the media is left until purification begins the next day (or 16 hours).

Table 2: Ion exchange chromatography buffer.

Hepes (pH 7.2 - NaOH)	$MgCl_2$	EDTA	KCl (elution)
25 mM	5 mM	1 mM	1M

Table 3: Poly-histidine tag affinity chromatography buffer.

Na_2HPO_4 (pH 7.2 - NaOH)	KCl	Imidazole (binding)	Imidazole (elution)
20 mM	500 mM	30 mM	300 mM

All purifications steps are undertaken at 4°C. The bacteria laden media was centrifuged at 8000 rpm for 10 minutes to pelleting the cells. The binding buffer of the chromatography method is used creating a suspension buffer by adding ATP, DDT and PMSF up to 2 mM, 5 mM and 1 mM respectively [40]. The cell pellets are resuspended in the buffer where a cell pellet from 600 ml of media is resuspended in 20 ml of buffer. The cells are then lysed open via sonication. The solution is then centrifuged at 18000 rpm for 15 minutes where insoluble cellular detritus is pelleted. The supernatant is removed and run over a GE Healthcare HiTrap Q FF or GE Healthcare HisTrap HP 5 mL column. In IEX, salt gradient is increased to 50% elution buffer over 20 minutes, activase usually elutes between 30 and 40 mS/m, the fractions can be examined by SDS-PAGE. When the 50% elution buffer is met, 100% elution buffer is then run over the column for two column volumes. The fractions containing Rca are then pooled together and a saturated ammonium sulfate solution (adjusted to pH 7.0) is added up to a final concentration of 37.5%. Use the equation (Volume of Supernatant/0.625-volume of supernatant) The solution is let to sit for 10 minutes while the protein precipitates out. This is then centrifuged at 12000 rpm for 10 minutes pelleting the precipitated protein. The supernatant is discarded and the pellet resuspended in BTP buffer.

Table 4: BTP buffer.

BTP (pH 8.0 - NaOH)	KCl	EDTA
20 mM	20 mM	0.2 mM

At this point glycerol can be added and the sample flash frozen in liquid nitrogen for use later. R294V has a polyhistidine tag with TEV cleavage site therefore purified

R294V samples are incubated with TEV protease before Size Exclusion chromatography (SEC). The sample then undergoes the third and final purification step through the size exclusion column (200 increase). If the sample is relatively pure, two peaks will elute off the column, the first contains rubisco activase (excluding aggregate peak).

2.1.c SDS-PAGE.

SDS-PAGE (Sodium Dodecyl Sulfate-PolyAcryliamide Gel Electrophoresis.) was carried out numerous times to assay the purity of samples. The process was carried out as per the manufacturer's instructions. Alongside, 5 µl of Novex® Sharp Pre-Stained Protein Standard was run as a ladder.

2.1.d Protein parameters.

Protein parameters used were generated by ExPASy ProtParam tool.

Table 5: protein parameters.

protein	residues	MW (D)	Ext. coefficient*	Abs 0.1%*
T-RV	384	42773.6	39880	0.932
T-RV nt	314	34893.1	29910	0.857
Tα	421	46758.0	42860	0.917
Tob	384	42830.72	39880	0.931

*Assuming all cysteine residues are reduced.

Protein concentrations were checked before each experiment by UV-vis 280 nm after centrifugation at 4°C on the Nanodrop 1000 spectrophotometer.

2.1.e ATPase activity.

ATPase activity assays are conducted by measuring the oxidation of NADH by a lactate dehydrogenase and pyruvate kinase coupled assay as described [21]. The reaction mixture is made to these specifications.

Table 6: ATPase assay components.

Compound.	Concentration.	mg/mL
Lactate dehydrogenase	900-1400 u/ml	-
Pyruvate kinase	600-1000 u/ml	-
ATP	124 mM	68.8
NADH	8 mM	5.8
phosphopyruvate	50 mM	10.3

The assay is conducted at 25 °C where sample mixtures a left in the Varian Cary 100 Bio UV-Vis Spec to equilibrate for 10 minutes. It is conducted in Tricine buffer which consists of 100 mM tricine, 10 mM $MgCl_2$, 20 mM KCl adjusted to pH 8.0 with KOH. The enzymes were sourced from Sigma Aldrich as a mixture of both.

Table 7: reaction mixture.

Compound.	Amount transferred to make 1mL	Final Conc
Tricine buffer.	925uL	-
LDH + PK solution.	15uL	12+ u/mL
ATP.	20uL	2.5 mM
NADH.	20uL	0.16 mM
phosphopyruvate.	20uL	10 mM

The spectrometer is blanked by a solution of tricine buffer at 25°C. The assay components are mixed and left to incubate for 10 minutes. The spectrometer is activated and begins to take the absorption at 340 nm. After a baseline has been

established (~10 seconds), the spec is opened and the volume of Rca solution is added and thoroughly mixed using a large pipette. The spec is then closed and the assay let to progress for 4 minutes. The rate is taken from 2-3 minutes after mixing to allow for the Rca's lag in activity. Rca stored in BTP buffer does not need to be buffer exchanged into tricine buffer.

2.1.f Differential Scanning Fluorometry.

Differential Scanning Fluorometry (DSF) is conducted using a protein concentration of 0.25 mg/mL in the BTP storage buffer in an Applied Biosystems Quantstudio 3. The mixture is made to 100ul of the correct concentrations and 25 µL is then pipetted into 4 wells. In all cases the wells contained 0.2 mM nucleotide and 5 mM Mg^{2+}. To each, a small sample of diluted Sypro-Orange is added. Sypro-orange dye was used at a 10X concentration, starting from 5000X concentrate it was diluted 1 in 10 then 2 µL is added to the 100 µL sample for a final concentration of 10X. The samples Started at 10 °C were heated at a constant up to 90°C. The melting temperature was obtained from the midpoint of the melt peak called the thermal midpoint.

2.1.g Analytical Ultracentrifugation.

Analytical Ultracentrifugation (AUC) sedimentation velocity experiments were conducted on a Beckman-Coulter ProteomeLab XL-1 described in [21]. The experiments were done at varied concentrations shown on the key of the data. $MgCl_2$ is added to 5 mM with nucleotide containing experiments. The obtained data was processed in SEDFIT where the frictional ratio used was the value published in [21]. AUC can and was conducted at different temperatures however all values are transformed to S20w values. This was achieved by using the program SEDNTERP to get a modifier to fix obtained values.

2.1.h Small Angle X-ray Scattering.

Small Angle X-ray Scattering (SAXS) experiments were undertaken at the Australian Synchrotron in Melbourne. The samples where run in BTP buffer containing nucleotide specified to the experiment undertaken with a constant Mg^{2+} conc. There were two techniques used, plate and in line column. In line column used a 3 mL SEC column to separate out the sample and place in buffer. Buffer matching occurs by subtracting the pure buffer from the scattered sample using the software ScatterBrain. Using the 96 well plate, the sample is taken up into a capillary in line with the beam. Buffer matching is conducted by using two blank samples containing no protein. The in line column method the sample is run down a size exclusion

column where the buffer matching occurs as the buffer elutes first. In all cases, detector and beam setup the same as described in [21].
Buffer matching was conducted using the scatterBrain developed at the Australian Synchrotron. Data analysis was conducted with ATSAS software developed by EMBL Hamburg. Values obtained were from the ATSAS data analysis software Primus. Models generated from SAX data were generated in the ATSAS's GASBOR using the number of points equal to the number of residues of the protein (R294V=384, R294V N-terminal truncated=314) and then modeled in PyMol setting the sphere size to 1.5.

Chapter 3: Results

3.1.a Purification.

With the described methods ample amounts of Tobacco rubisco activase was purified. With IEX, rubisco activase would consistently elute at the described mS/m and can be pooled for further purification. The amount of purified protein almost exactly matches up with the observed thermal stability at low levels of nucleotide.

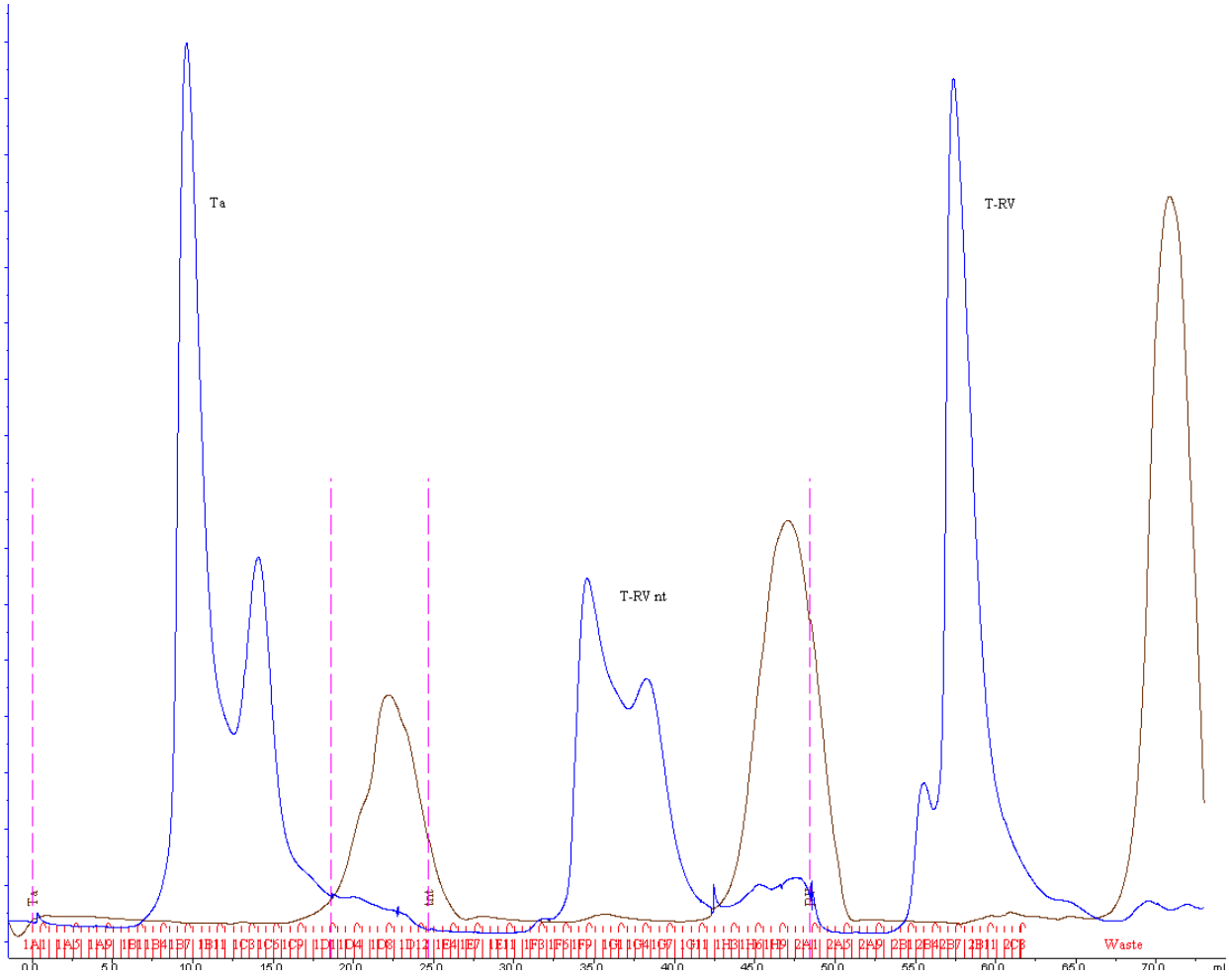
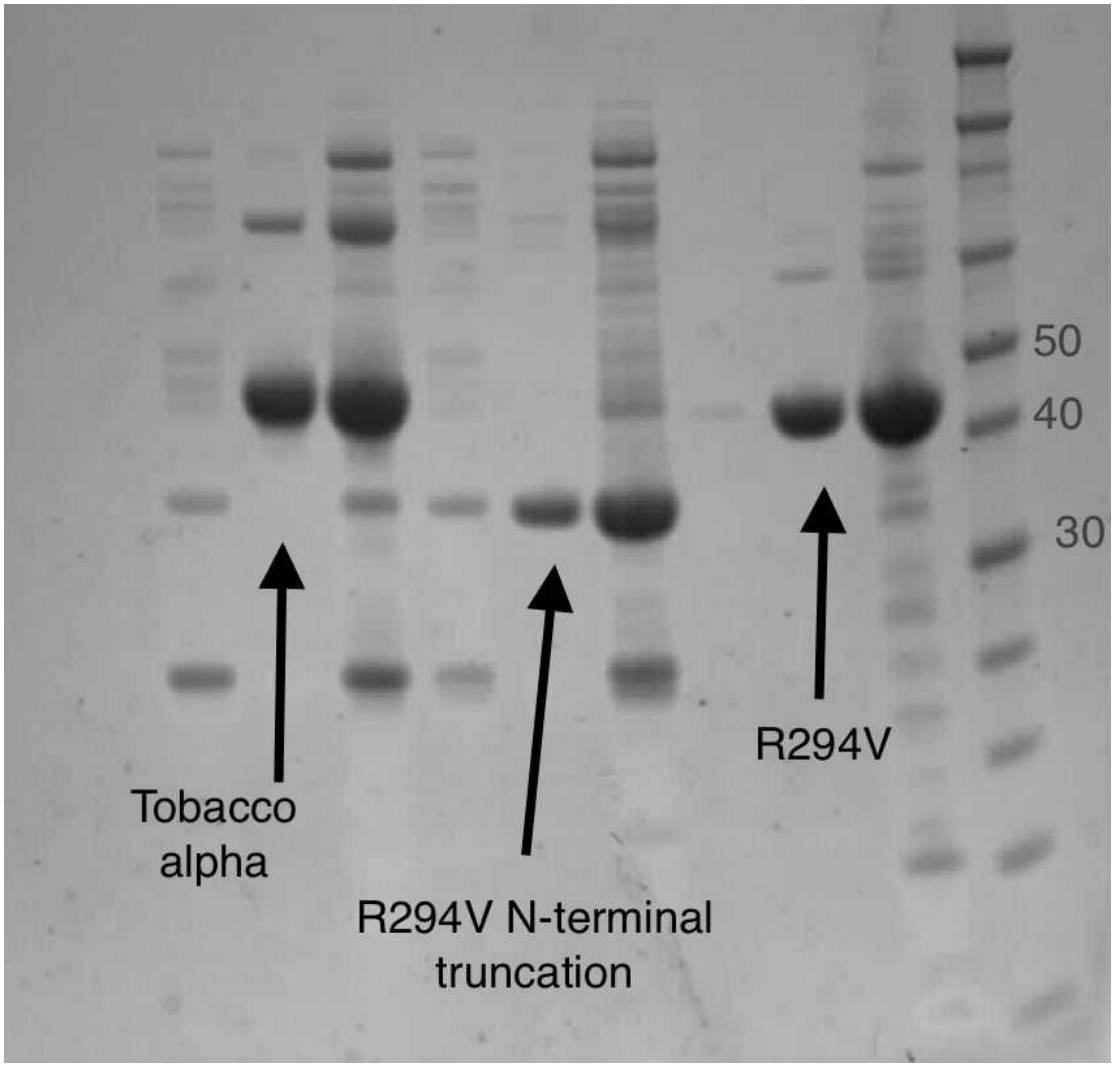


Figure 7: Akta read out during SEC purification process.
SEC was conducted with the samples as the final step in the purification procedure. Two peaks were regularly spotted and in all cases that were purified by IEX, the first peak to elute (the larger sized protein) contained rubisco activase.

A typical UV light and conductivity output with Rca purification. In this case the purified constructs are Tobacco alpha (T α), Tobacco R294V N-terminal truncated (T-RV nt) and then tobacco R294V (T-RV). Generally the first peak contains the rubisco activase in this case however T-RV had aggregate that eluted first and lacked the second peak due to it being purified by His-tag chromatography. Peaks were pooled together and tested on a SDS-PAGE for identification and purity. The Gel Pictured below contains the tobacco constructs after purification by SEC.



3.1.b Analytical Ultracentrifugation.

Sedimentation velocity experiments in the presence and absence of nucleotide.

Figure 8: Tobacco R294V and Tobacco R294V N-terminal truncated at 0.25 mg/mL S20w plot without nucleotide.

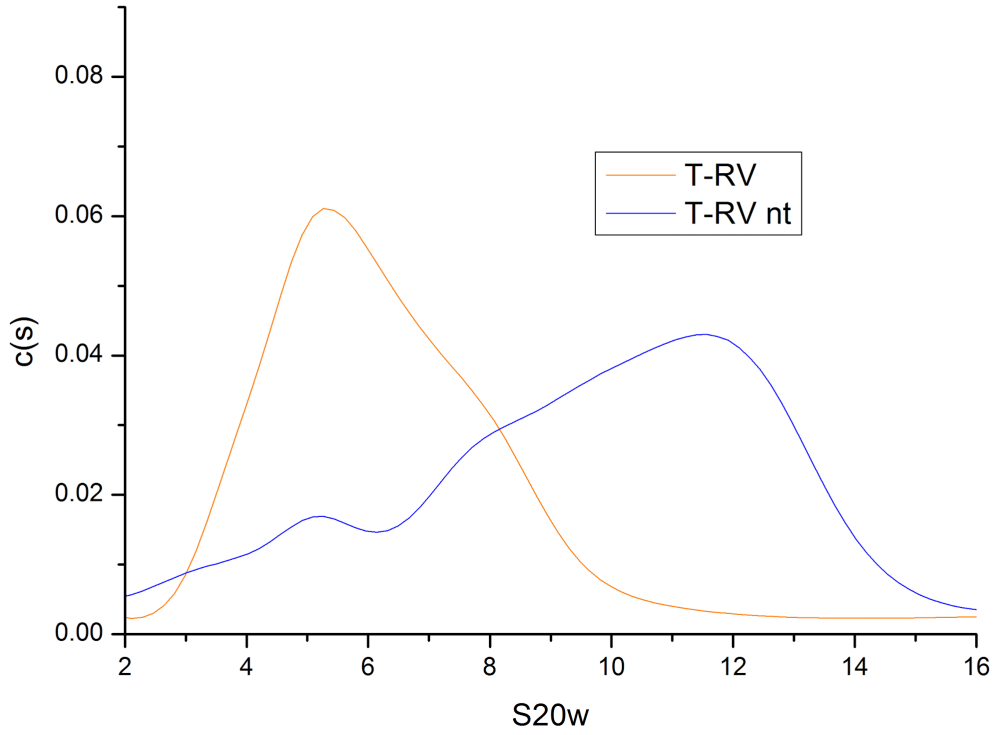


Figure 9: Tobacco R294V and tobacco R294V N-terminal truncated in the presence of ATPYS.

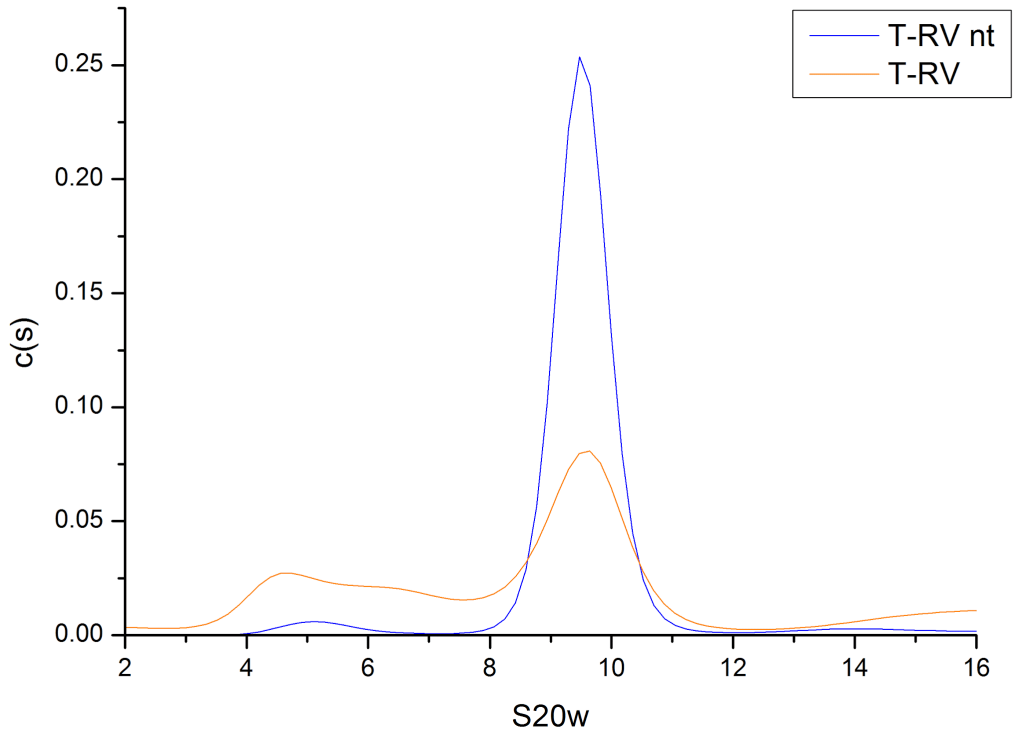
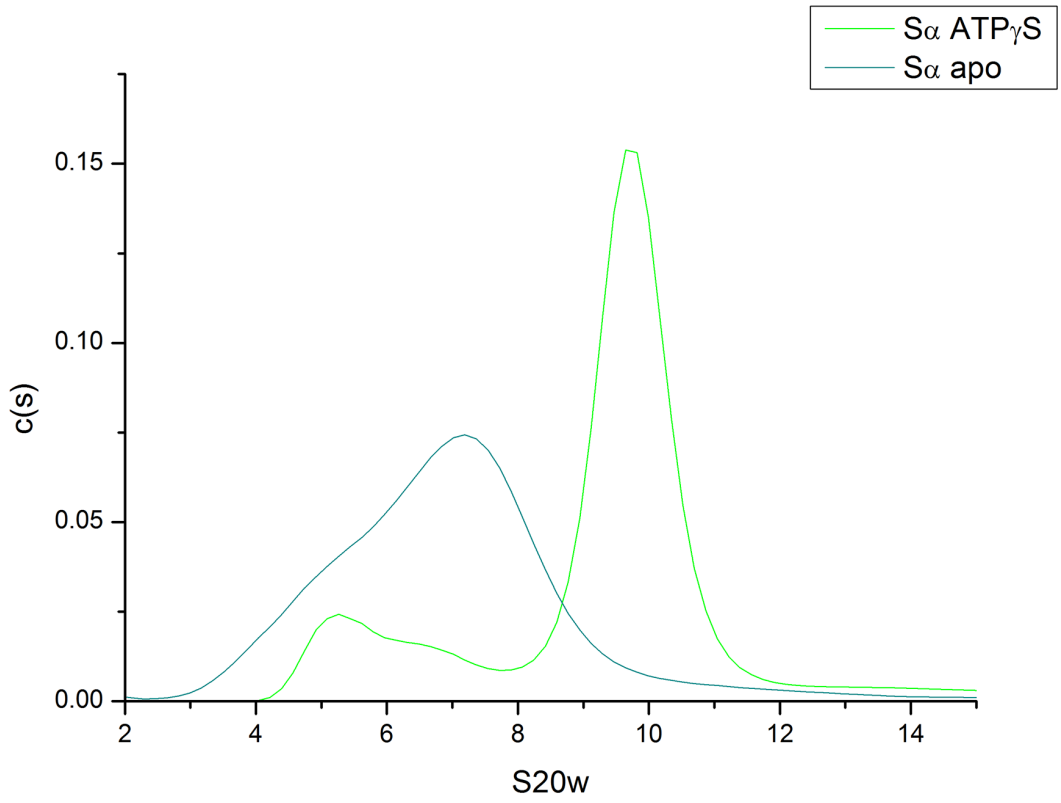


Figure 10: Spinach alpha isoform in the presence and absence of ATPγS



Figures 8-10 were carried out at 12°C in BTP buffer and the obtained values were later converted to S20w values obtained via SEDNTERP. These 3 graphs show the effects of ATPγS on the quaternary structure of two of the tobacco constructs, T-RV and T-RV nt and effect on the Spinach long or alpha isoform Rca. Comparing Figures 1 and 2 show a homogenising of the species formed from a dispersed peak containing shoulders to a singular peak at ~ S = 9.5 There is however observed species between S = 4-7, this however is contaminant. The developers of SEDFIT recommend a 95% purity for publication, this is not meet by the T-RV experiment but satisfied in the T-RV nt experiment. T-RV S values have been published before and so were not followed up. Figure 3, demonstrates that spinach Rca alpha isoform forms a singular species in the presence of ATP γ S, once again there is contamination in the sample.

Figure 11: Tobacco alpha S20w graph with increased concentration and ATPYS.

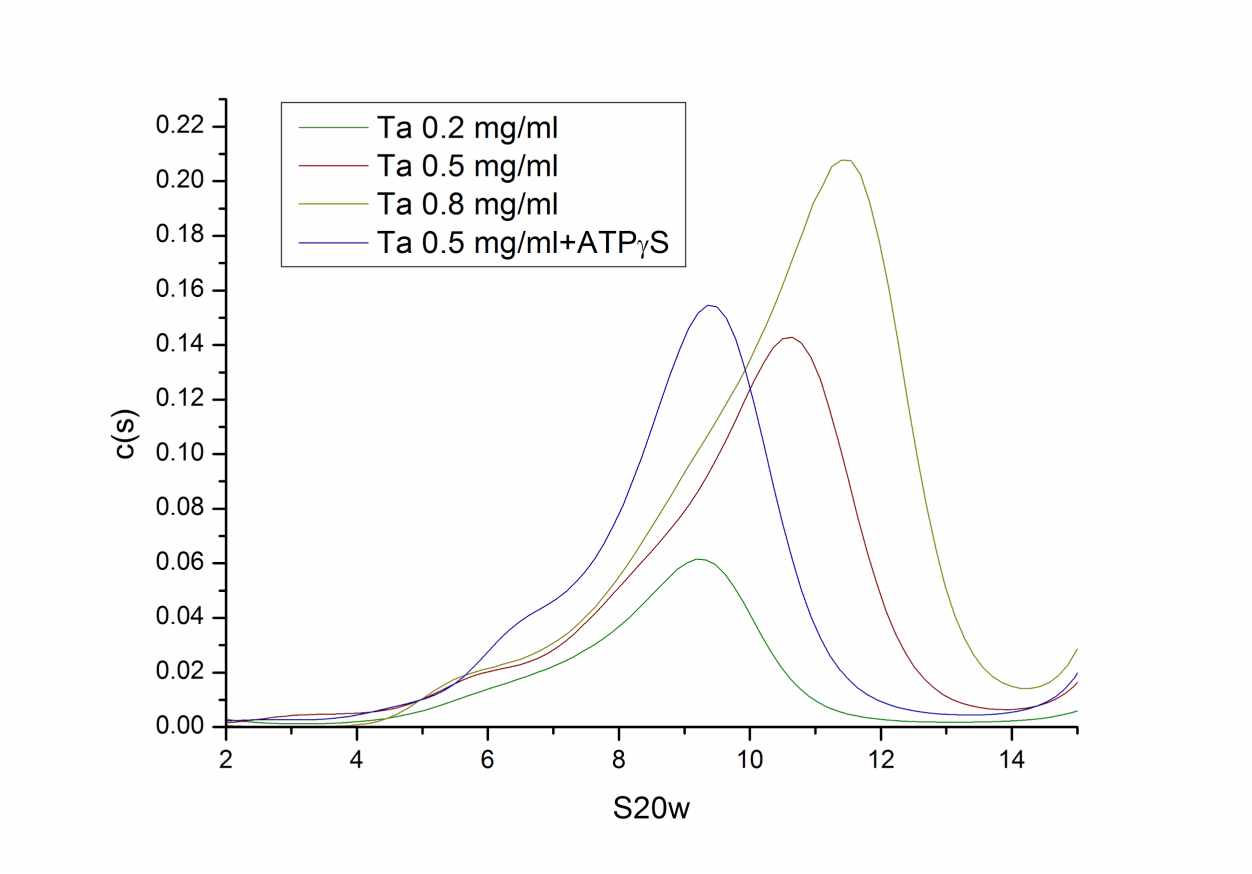


Figure 11: was conducted at 20°C in BTP and shows the sedimentation coefficients of the tobacco alpha construct and importantly how it does not form a singular peak with the presence of ATPYS.

Figure 12: Tobacco R294V N-terminal truncated S20w plot with increasing concentration in the absence of nucleotide.

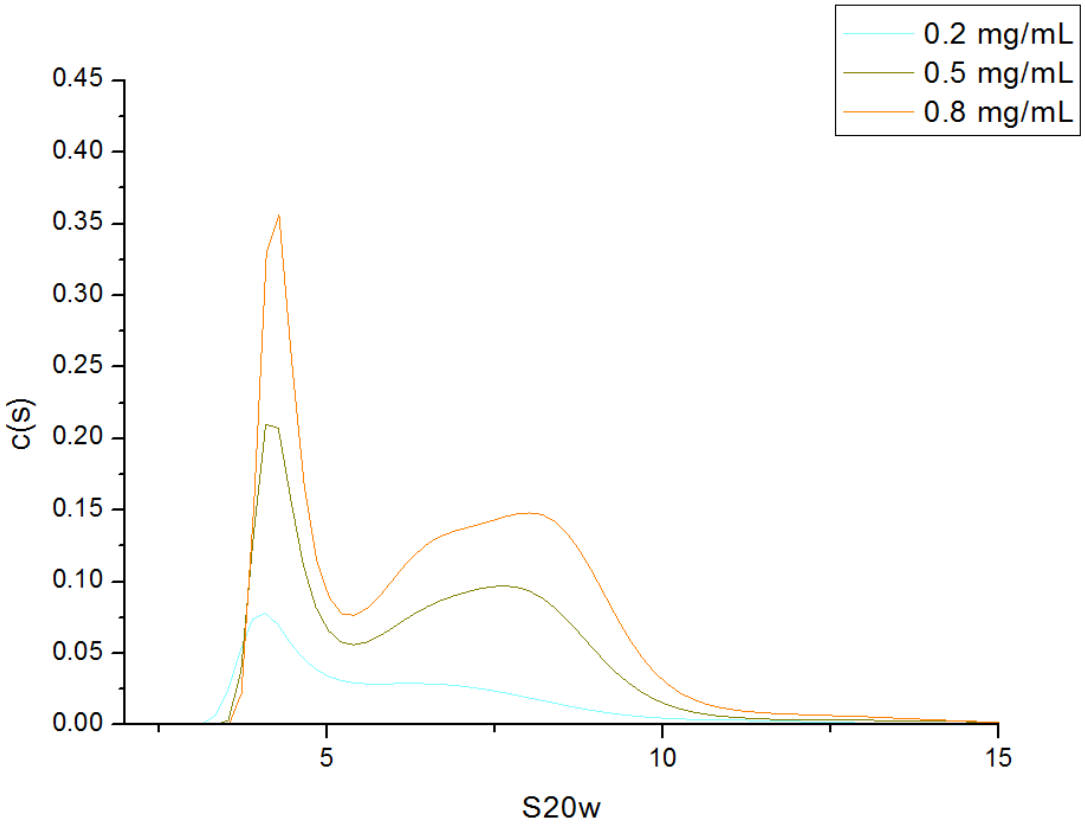


Figure 13: Tobacco R294V N-terminal truncated at 0.5 mg/mL in the presence and absence of ADP.

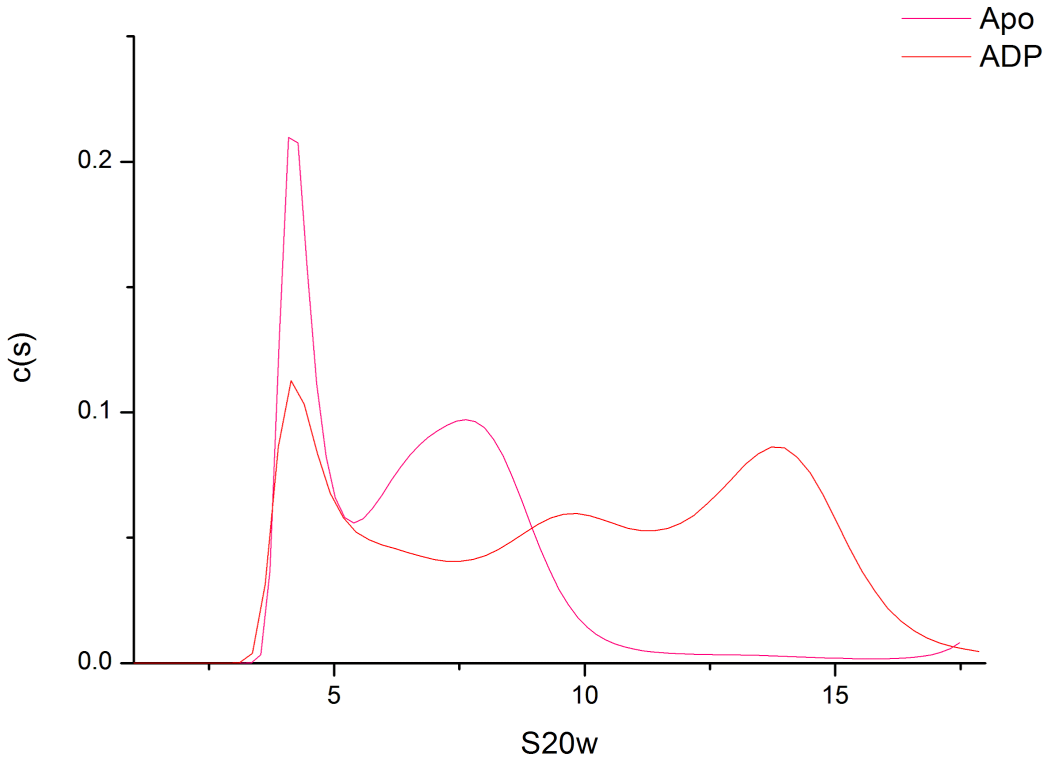
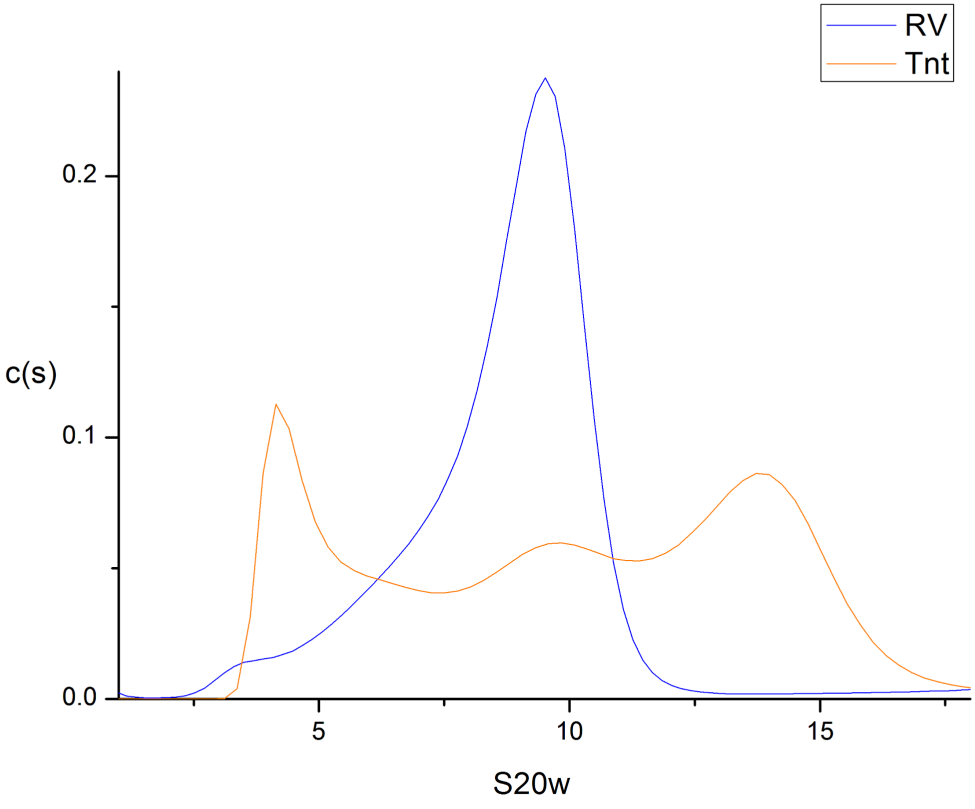


Figure 14: tobacco R294V and tobacco R294V N-terminal truncated at 0.5 mg/mL with ADP.



Figures 12-14 were conducted at 12°C in BTP Buffer. It shows the T-RV nt construct with increasing concentration and in the presence of ADP compared to T-RV.

3.1.c Kinetics.

Figure 15: Specific ATPase rates of native tobacco Rca and Rca constructs.

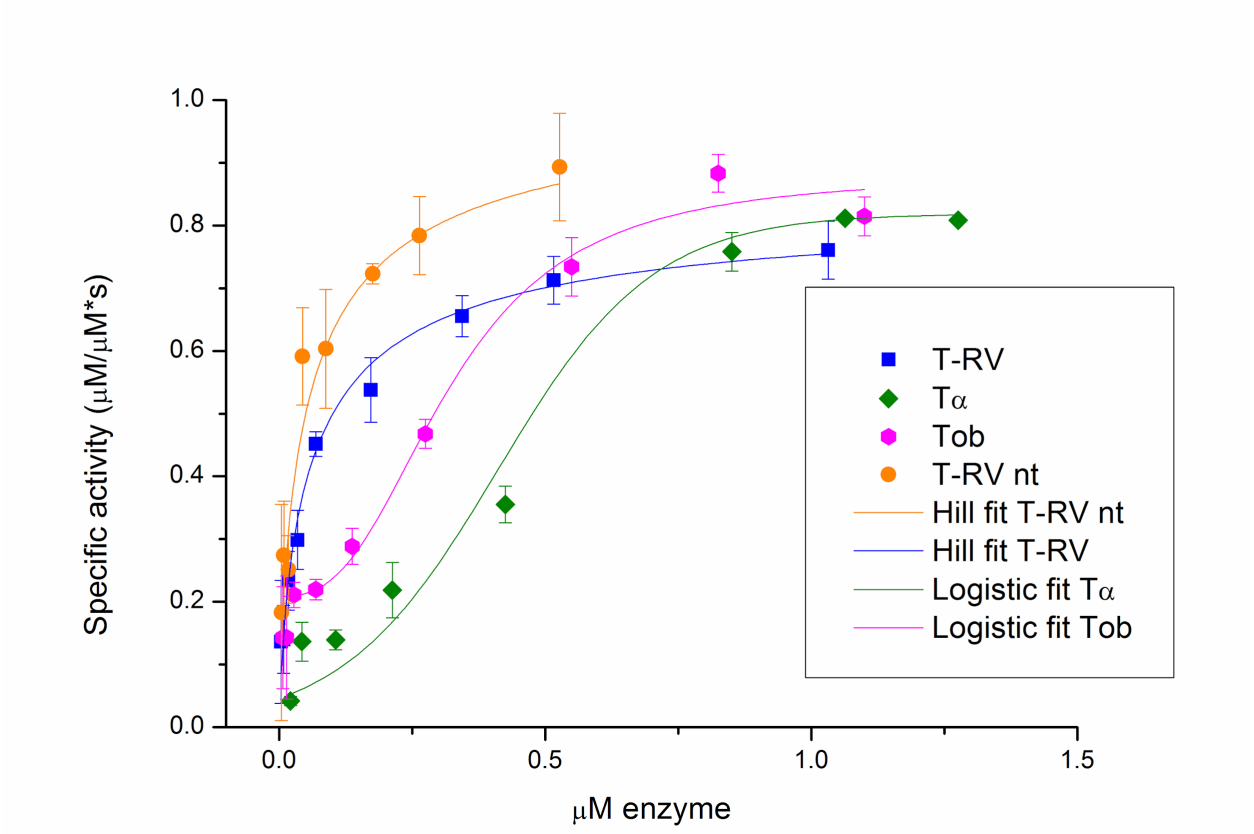


Figure 8 shows the obtained specific ATPase rates of the 3 tobacco Rca constructs and native tobacco Rca. The specific activity here is expressed as μM of product per μM of enzyme per second. An equation has also been generated and the line shown. A Hill equation was used to fit the theoretically very tight sigmoid shape formed by

the two R294V constructs. A logistic equation has been used to fit the less tight sigmoidal shape of the non-R294V Tα and native tobacco Rca (Tob).
The maximal specific activity can be calculated from the fitted functions.

Table 8: Maximum specific activities (SA) of tobacco Rca and constructs.

Construct	T-RV	T-RV nt	Tα	Tob
Max SA	0.761	1.0327	0.817	0.88156

3.1.d Differential scanning fluorometry.
Figure 16: Thermal midpoints in the presence of ATPYS.

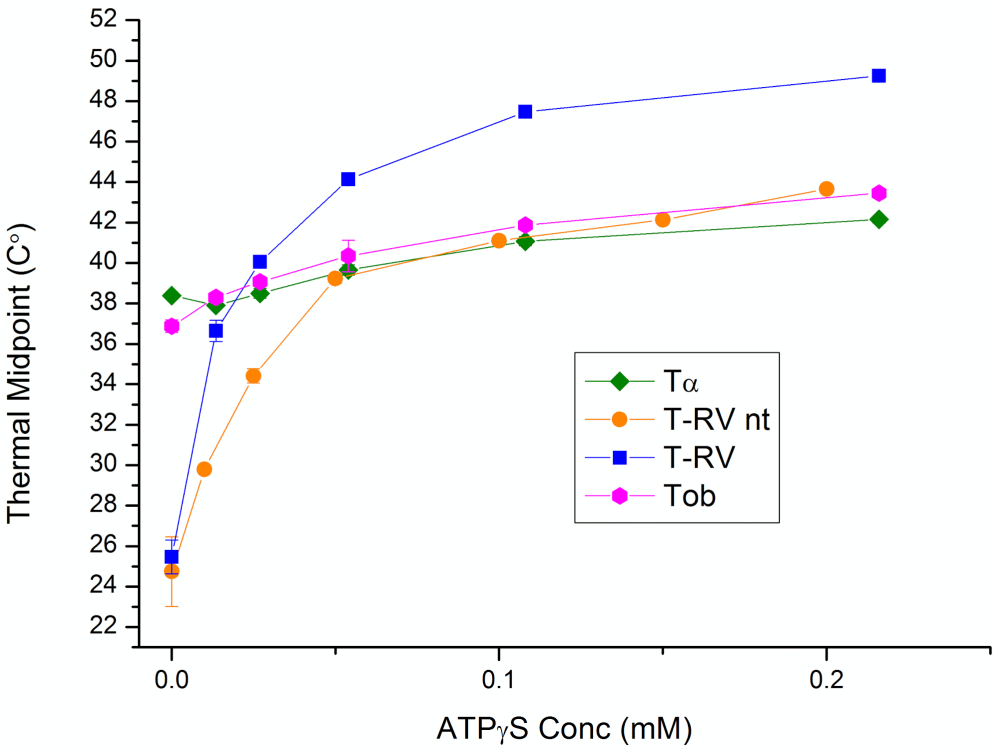
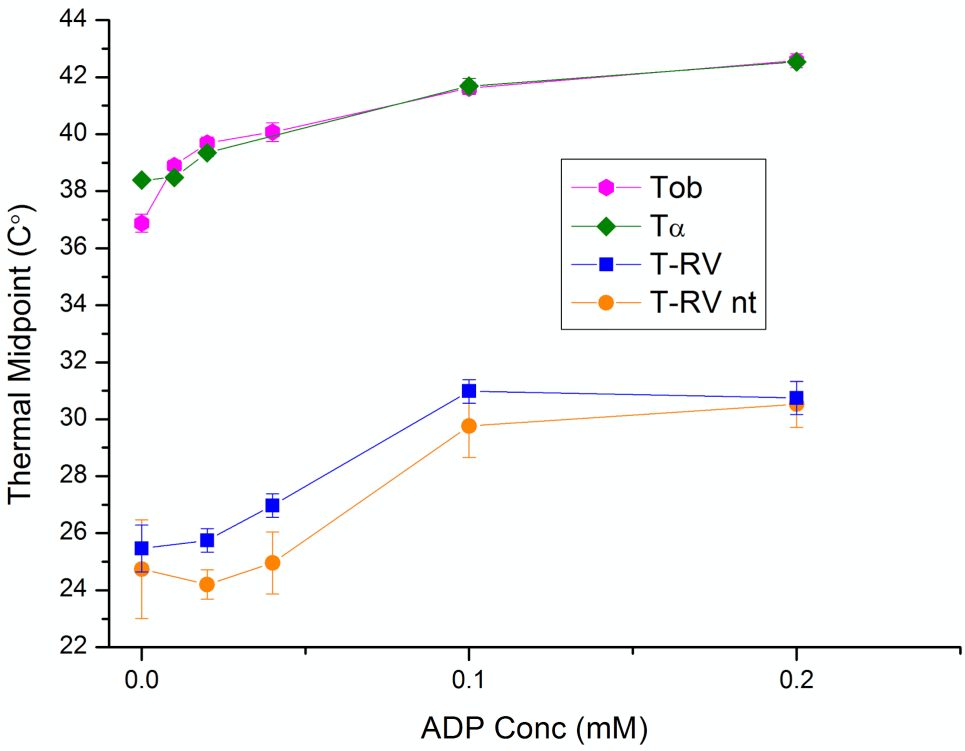


Figure 17: Thermal midpoints in the presence of ADP.

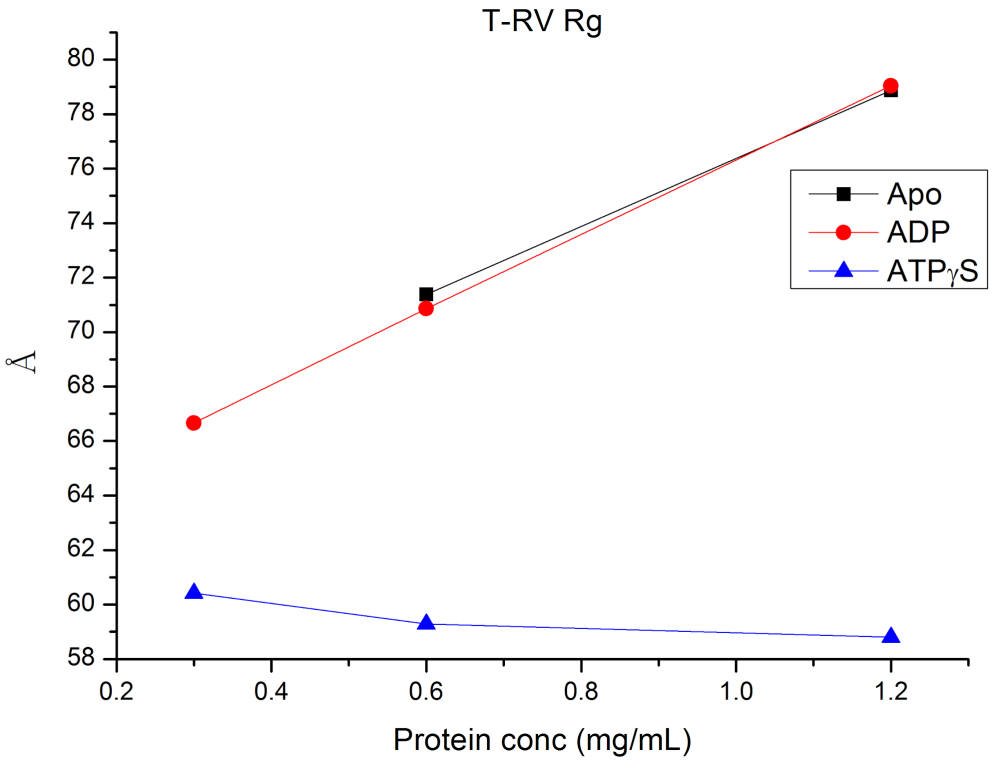


Figures 16 and 17 are thermal midpoints obtained for the Tobacco constructs and native tobacco Rca. During experimentation the DSF machine was updated and so all were done on Applied Biosystems Quantstudio 3 however the tobacco R294 N-terminal truncated ATPYS and Tobacco alpha ADP experiments were done on the Biorad IQ5. The results displayed for no nucleotide have the same concentration of 5 mM Mg^{2+} as the experiments with nucleotide.

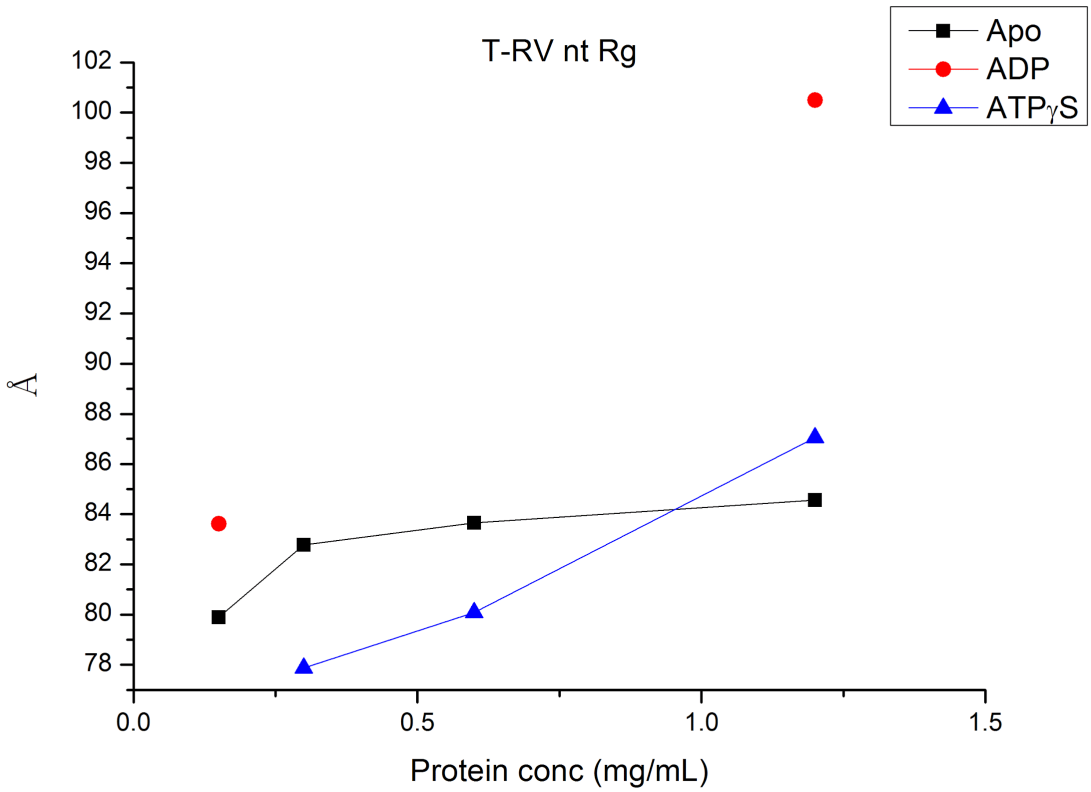
3.1.e Small-Angle X-ray Scattering.

Figures 18-20: Tobacco Rca constructs radius of gyration with increasing concentrations in the presence of ATPγS, ADP and Apo.

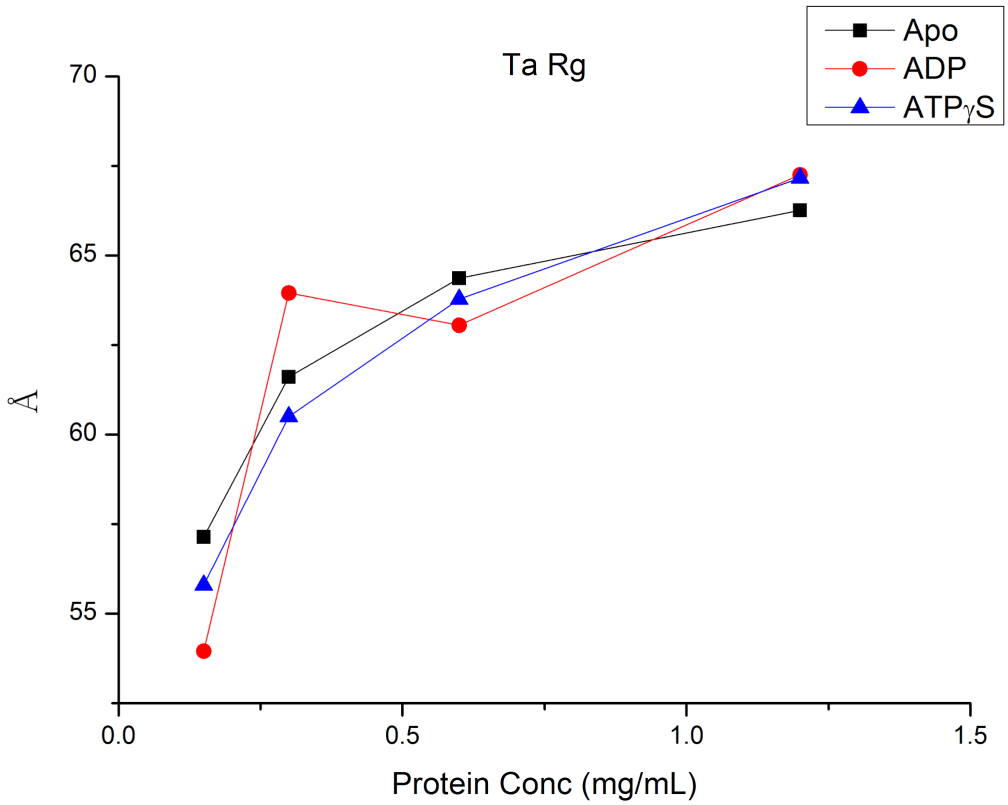
18.



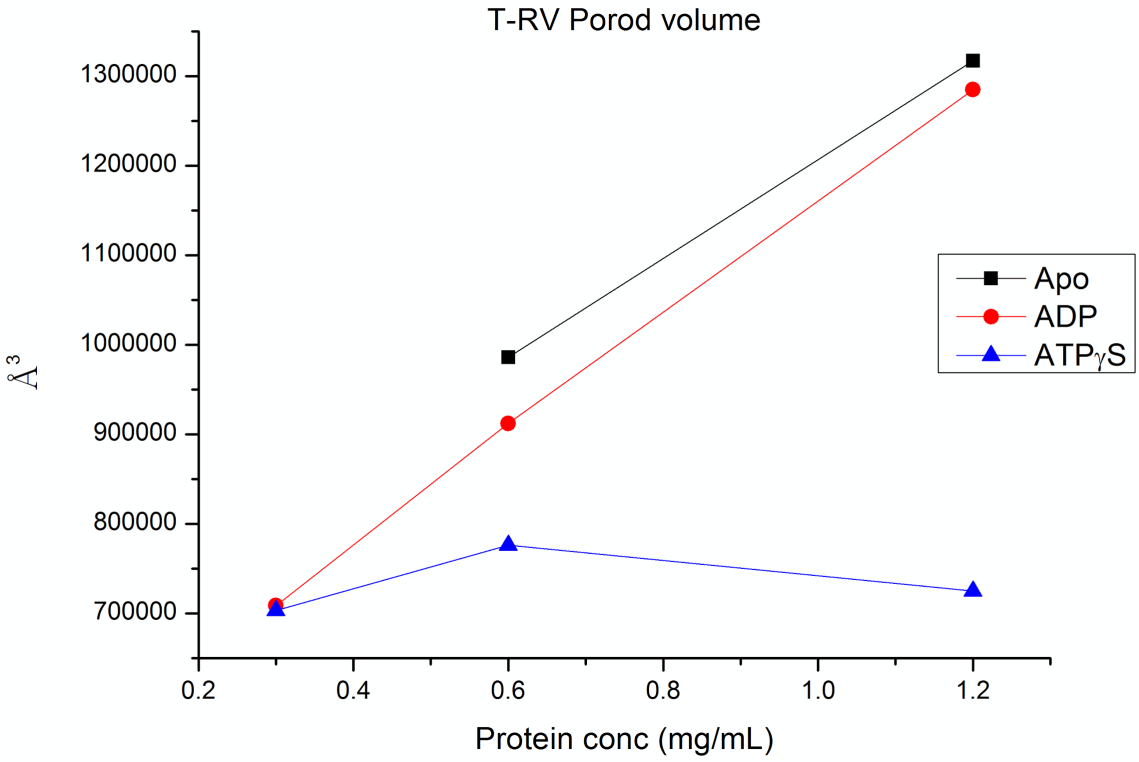
19.


T-RV nt Rg
Apo
ADP
ATPγS
Å
102
100
98
96
94
92
90
88
86
84
82
80
78
0.0
0.5
1.0
1.5
Protein conc (mg/mL)

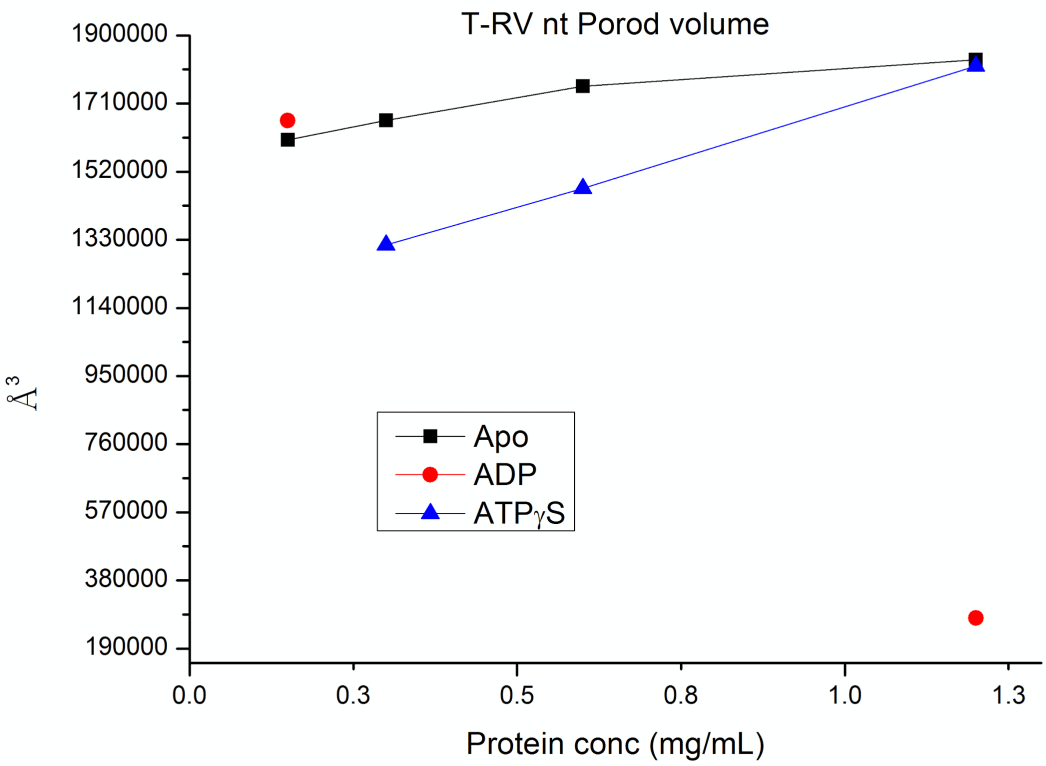
20.



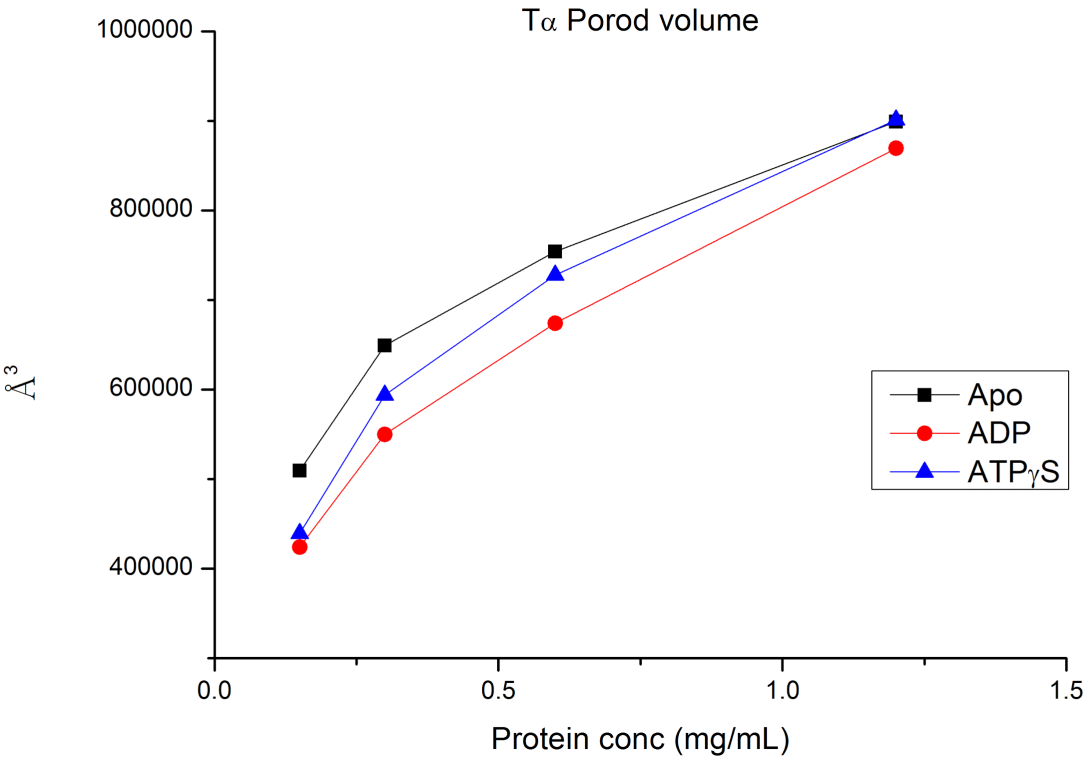
Figures 21-23: Tobacco Rca constructs Porod volume with increasing concentrations in the presence of ATPYS, ADP and Apo.
21.



22.


T-RV nt Porod volume
1900000
1710000
1520000
1330000
1140000
950000
760000
570000
380000
190000
Å3
0.0
0.3
0.5
0.8
1.0
1.3
Protein conc (mg/mL)
Apo
ADP
ATPγS

23.



Figures 18-23 demonstrate data obtained by Australian synchrotron in Melbourne using the plate method in november. It is important to note the T-RV nt could have been compromised at this time.

3.1.f Kratky plots.

Kratky plots obtained from Primus SAX data analysis, in each case the sample with ATPYS is in red, ADP in green and Apo in blue.

Figure 24: tobacco R294V

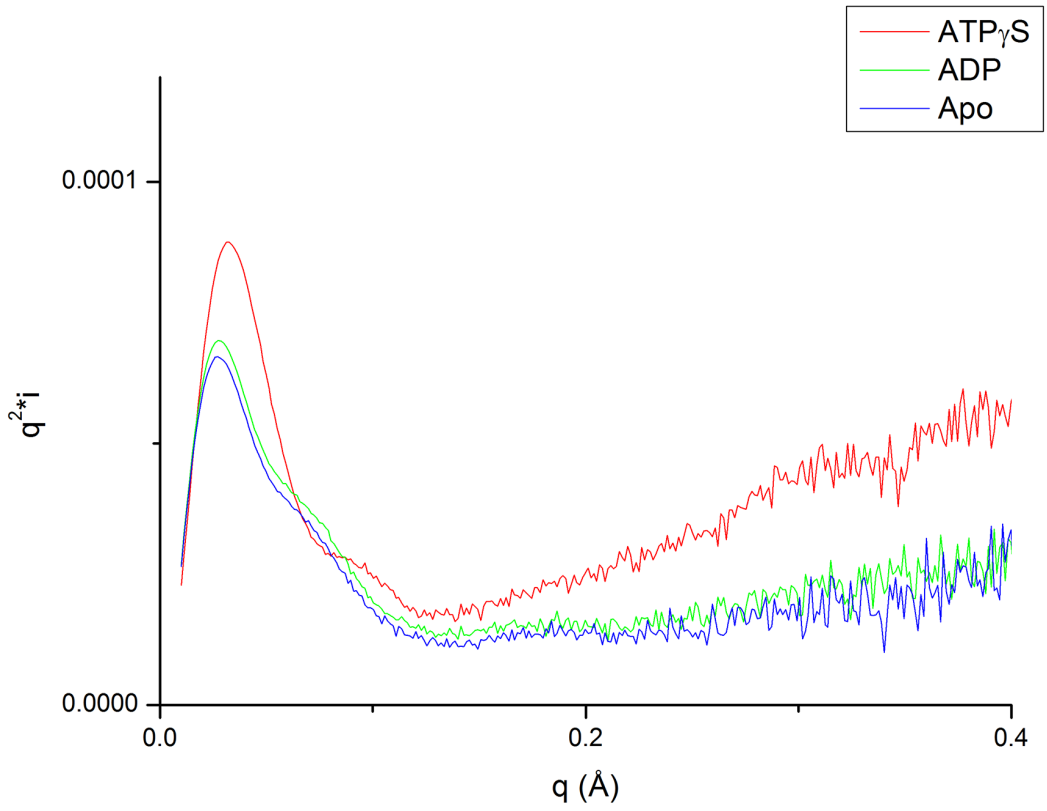


Figure 25: tobacco R294V N-terminal truncated

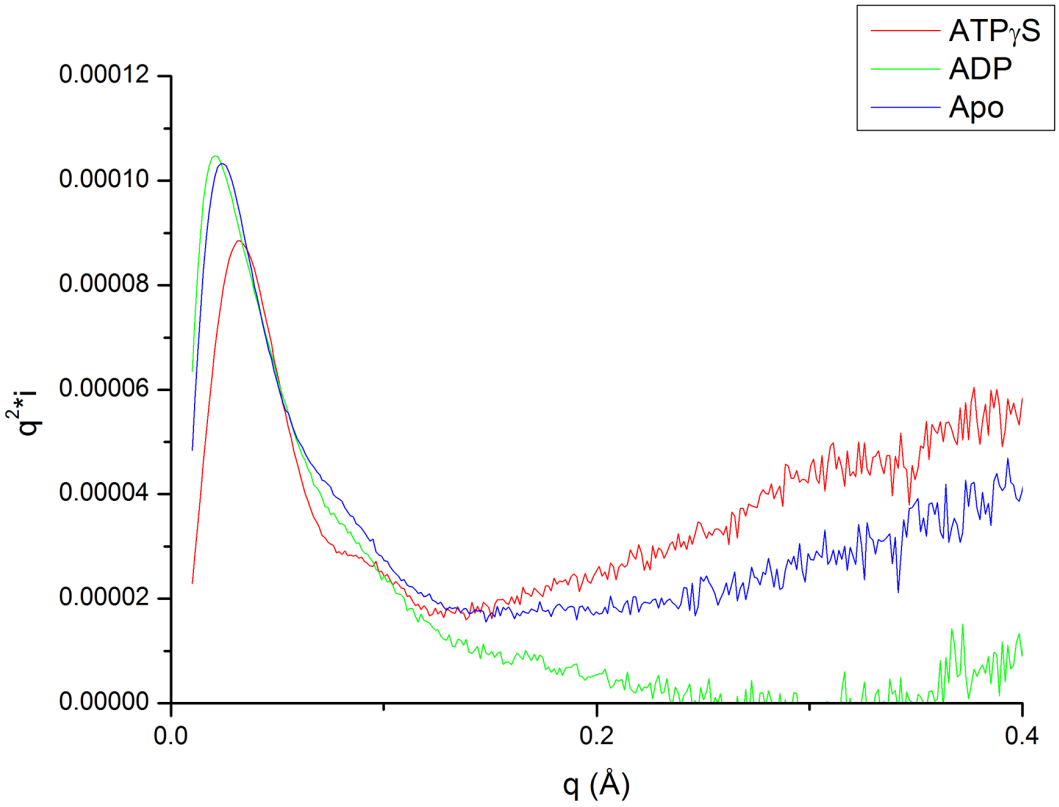
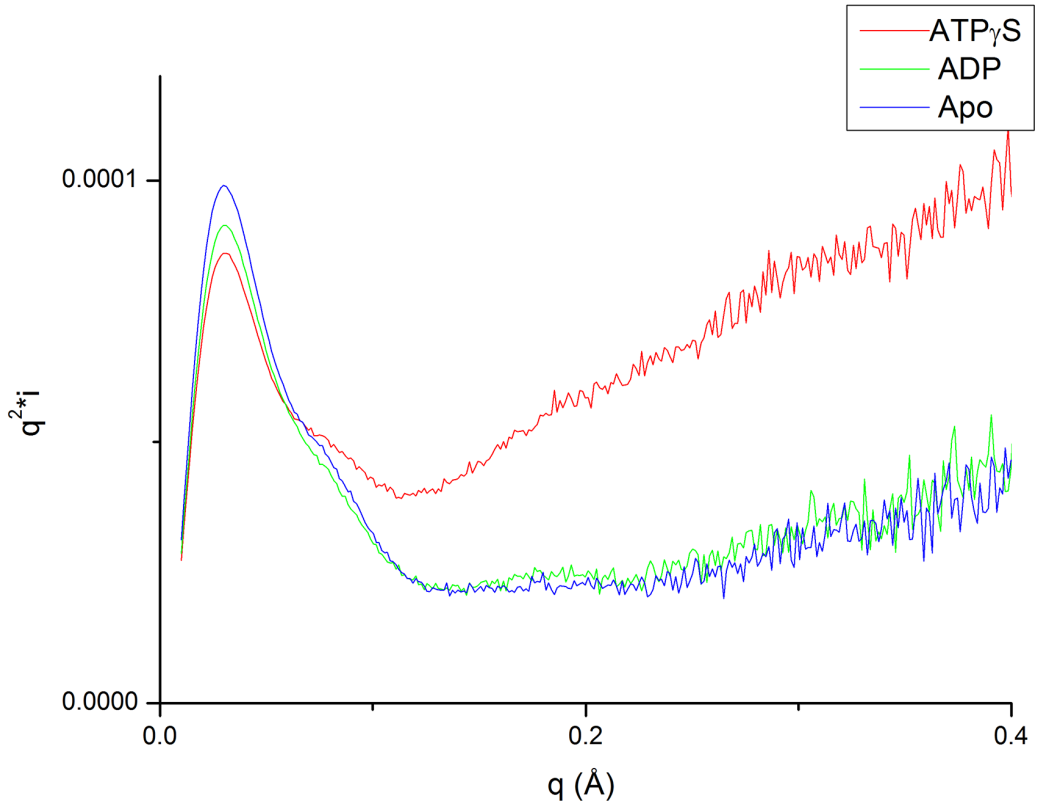


Figure 26: tobacco alpha.



Figures 24-26 are Kratky plots generated by the constructs scattering at 1.2 mg/mL protein concentration under conditions containing ATPYS, ADP and Apo. In each case the construct in the presence of ATPYS shows a higher plateau which qualitatively suggests noncompactness.

3.1.g GASBOR generated models.

Figures 27-30 are GASBOR generated images from the March SAX data collection in Melbourne using the in line column.

Figure 27: tobacco R294V pore view.

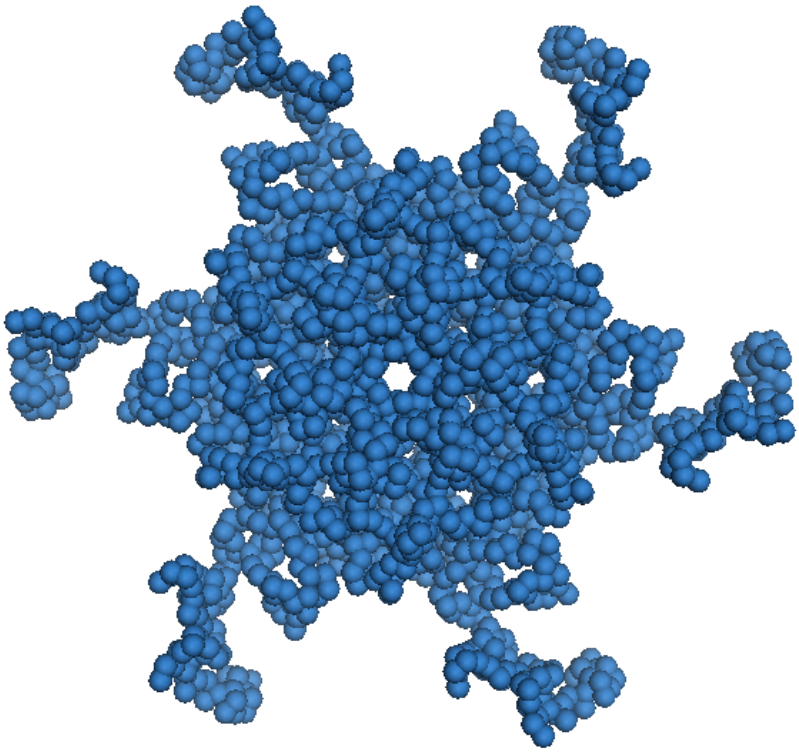


Figure 28: tobacco R294V side view.

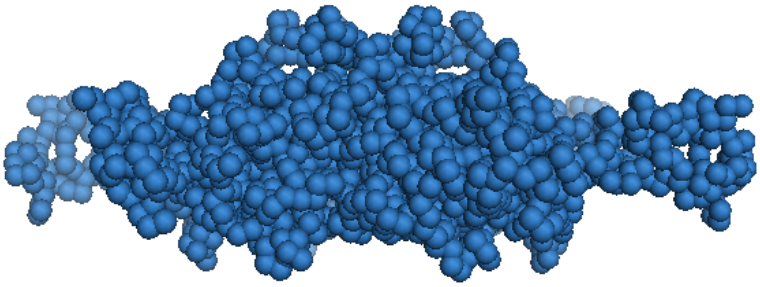


Figure 29: tobacco R294V N-terminal truncated pore view.

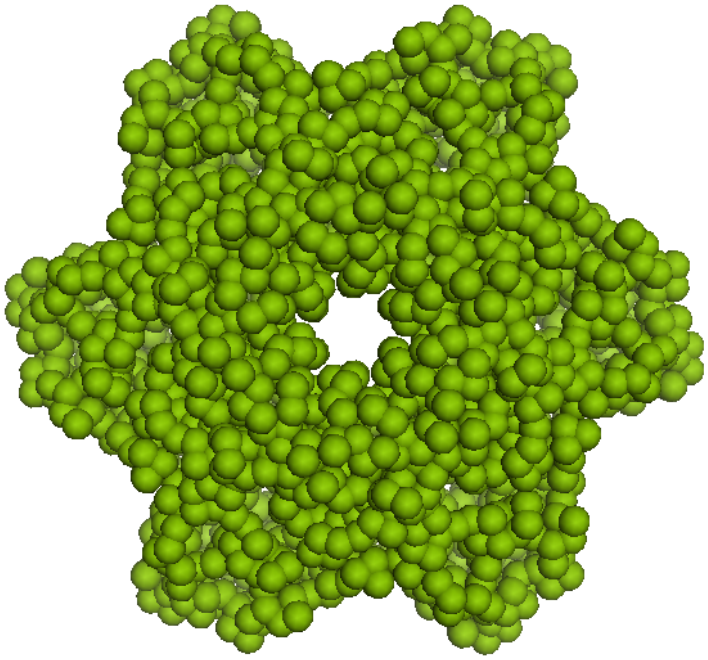


Figure 30: tobacco R294V N-terminal truncated side view.

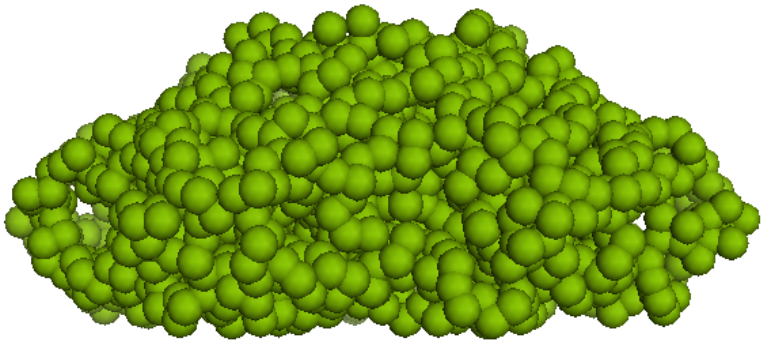


Table 9: Porod and Rg values from data used to generate images.

In line SAX data	Porod	Rg
T-RV	566798	56.06
T-RV nt	327797	42.53

Chapter 4: Discussion

4.1.a Purification.

The initial focus of my thesis was to assay the diversity of cbbX (red algae Rca) across multiple species. Both anion exchange and polyhistidine tag affinity chromatography were utilized, with both the yield was similar for the species tried. The yield for all the cbbX proteins were very low, the highest yield was repeatedly discovered to be from *R. sphaeroides* which is the species the original research on cbbX had used [26]. Compounding the complications brought by the low yields was the fact that the extinction coefficient of cbbX was very low at 16390 $M^{-1}sm^{-1}$ meaning that many UV based experiments became more challenging due to requiring more protein. For example the AUC uses a UV light system for detecting the protein.

Onto the terrestrial plant Rca which had consistently higher yields and therefore the thesis was changed to studying the Rca constructs. Once again both anion exchange and polyhistidine tag chromatography were utilized. The Spinach constructs regularly purified but bizarrely would later appear to be aggregate. The soluble protein would not pellet under centrifugation but would in DSF appear to max out the fluorescence suggesting the protein was completely aggregate. On AUC the protein would show a peak at ~5 S which would not change with the presence of ATPYS.

In the end it was the tobacco Rca constructs the purified reliably and so were used in this thesis.

4.2 N-terminal.

4.2.a Truncation of the N-terminal domain enhances subunit assembly.

The tobacco R294V N-terminal truncated construct appears supercharged to other studied hexameric Rca as that it forms much larger sized species in the absence of nucleotide. First demonstrated by the AUC data (figure 8), where the $S20_w$ peak of 11.57 is over double that of apo tobacco R294V peak at 5.26. The species formed by the N-terminal truncated construct spread out from 8 to 13 compared to the S values from non-truncated of 4-8.

Even with the disappointment of the faulty protein in figures 12-14 the S values for those peaks are higher than what is observed by Apo R294V the peak at 0.5 mg/mL being 7.62. In the presence of ATPYS in figure 9 both constructs produced a peak of

around 9.5S constituting a hexamer. The increase in oligomeric state can also be observed in the SAXS data where the Porod volume of the R294V N-terminal truncated (+1710000 is considerably larger than T-RV (1,300,000). Important to note is that the R294V N-terminal truncated may have been compromised at this point which may help to explain the bizarre behavior of the ADP experiments. Unfortunately due to the compromised protein in the AUC and SAXS experiments, nothing conclusive can be resolved about the N-terminal truncated Rca in the presence of ADP.

One hypothesis created to explain this is the removal of the N-terminal domains lowers the steric hindrance between the subunits, causing stronger interactions leading to oligomerization. Another hypothesis is that the truncation of the N-terminal is causing changes to the central AAA+ domain causing the enhanced assembly. The presence of various species with a higher molecular weight than a hexamer demonstrates that without nucleotide the R294V mutation creates small spiral oligomers rather than small pre-hexameric assemblies. With the addition of nucleotide however the N-terminal truncated construct forms a stable hexamer. The R294V mutation only causes hexameric assembly when nucleotide is bound due to structural changes to the AAA+ domain affecting subunit interactions. The observed crystal packing structures of *Arabidopsis thaliana* and *Nicotiana tabacum* [12,29] had the N-terminal present *Arabidopsis* Rca form a more drawn out spiral while the N-terminal truncated *Nicotina* formed a flatter spiral. These differences could be put down to species differences but was also hypothesised in the *Arabidopsis* paper that the N-terminal has influenced the packing. They based this on overlaying the *Arabidopsis* Rca with that of the AAA+ protein p97 [41] and found good overlap and structural homology. It is important to note the difference between my results and these two structures is the R294V mutation present in my Rca, which could be aiding in the quaternary assembly in an unforeseen way. In any case the highly dynamic nature of Rca's quaternary structure remains cryptic. Future experiments using C-EM may shed light on this issue as individual proteins can be observed (though not in as high resolution as a hexameric structure) and the helical arrangement could be elucidated and in different conditions such as ATP or ADP such as [12,26].

4.2.b Truncation of the N-terminal domain causes a loss of thermal stability.

With the loss of the N-terminal domain, the truncated construct has a lower thermal melting point than the non-truncated R294V mutant. This is evident along with the lower levels of purification and difficulty in storage and experimentation. In the presence of ATPYS the R294V mutation causes a sizeable increase in

thermostability as observed in this research and [23] with the loss of the N-terminal the stability drops down to that of the non-R294V mutants. The shape of the thermal midpoint data still however displays the rapid increase seen in R294V though at a lower rate and the maximum is lower. In the presence of ADP the Thermostability of both R294V mutants is far below that of the non-R294V. With ADP the difference between R294V and R294V N-terminal truncated is not as pronounce and may be due to the truncated construct forming larger complexes as demonstrated in the AUC data.

4.2.c Truncation of the N-terminal domain increases rate of ATPase activity.

Although the error margins on the ATPase rate of T-RV nt are relatively large the rate of the N-terminal truncated domain increases faster and is higher than that of T-RV. The max SA of R294V being 0.761 and 1.027 μM/μM*s for the truncated R294V. A minimally higher rate of ATPase activity (also within error) for an N-terminally truncated tobacco Rca has been recorded in a similar mutant in [12] which lacked the R294V mutation. This observation supports the previous claim that the N-terminal truncated R294V construct forms large complexes at lower concentrations due to the fact that the ATPase rate is allosterically modulated by neighbouring subunits. Further evidence for the Truncated construct degrading while in storage is that one of the later attempts at finalizing the obtained data the rate of N-terminal truncated R294V was reduced and had a max SA of ~0.6 μM/μM*s (not shown).

4.2.d Modeling.

The R294V mutation is important for the imaging process as it creates a singular species that has six fold symmetry. Images of R294V and R294V N-terminal truncated were generated from SAX data. The resulting models showed a dispersed R294V with the N-terminal radiating outward while the truncated version was tightly packed. The model of R294V demonstrated density on one face of the Rca hexamer that was in the same position as shown in [12], this electron density was missing in the truncated model. With the N-terminal gone the image generated of the central AAA+ domain could perhaps be better resolved. Qualitative evidence for this includes the the central pore is much better resolved in the N-terminal truncated model and that the non-truncated model has several gaps appearing on the model making it look porus.

4.3 C-terminal.

4.3.a The C-terminal extension does not produce hexamer formation in tobacco Rca.

The tobacco alpha construct was analysed by AUC to discern its quaternary structure with increasing concentration and in the presence of nucleotide. The morphology of the peak includes a shoulder which is seen in previous oligomeric Rca AUC experiments. With increasing concentration, the position of the major peak at S= 9.18, 10.64 and 11.4 demonstrates to a molecular weight shifting upwards from 215177 to 297775 Da assuming a frictional ratio of 1.4. This corresponds to species between 4 and 7 monomers forming. The addition of nucleotide caused a decrease in the sedimentation coefficient of from 10.64 to 9.35 suggesting the oligomer has shortened by at least one monomer. The SAX data obtained on Tobacco alpha supports the observed drop in molecular weight with the addition of ATP□S. The data shows the porod volume and Rg of tobacco alpha are smaller than with ATP□S than with ADP or Apo. The observed results for tobacco Rca from [23] suggest that the addition of nucleotide does not inhibit the size of the quaternary assembly (until higher protein concentrations). Meaning that the addition of the extension onto tobacco Rca is having a slight detrimental effect on the ability to assemble into oligomers. The C-terminal extension on Spinach Rca which occurs in nature causes hexamer formation in the presence of ATPYS (figure 10), the peak tightens and changes upwards in S value from 7.19 to 9.64 Corresponding to a hexamer, the same as what is observed in R294V. The failure to produce hexamer formation in tobacco shows that the C-terminal extension is interacting with another part of the protein.

4.3.b Addition of the C-terminal extension slightly lowers ATPase SA rate of increase.

The C-terminal extension does not occur on tobacco Rca *in planta* this construct was created to gauge the effects of the C-terminal on Rca. It affects the ATPase rate by reducing the rate at which maximum ATPase is reached. The tobacco alpha SA displays a similar sigmoidal shape and maximal rate to the native tobacco but

extended out suggesting that it is probably interrupting the quaternary structure assembly similar to the way that removal of the N-terminal can increase the ATPase SA due to increasing the ability to assemble. This in stark contrast to the effect of the extension on Spinach Rca in [23], which the spinach alpha Rca has an accelerated rate of SA increase and raised the maximal SA compared to spinach beta Rca. This is interesting due to the difference in quaternary structure form by these two. It seems as if a hexameric Rca structure could be predicted by the rate of SA increase with concentration if the shape of the SA curve has no observable “lag phase” at the lower concentrations.

4.3.c The C-terminal increases thermostability with no nucleotide.

The Tobacco alpha Rca construct demonstrated a base higher level of thermal stability than both the R294V mutants and native tobacco under Apo conditions in figures 16 and 17. It was immediately evident that the protein has been stabilized as it consistently produced higher yields during purification compared to the R294 constructs. It however, shows very little change in thermal midpoint with the addition of nucleotide. In the presence of ATPYS the thermal midpoint changes very little, rising ~5°C over the concentrations shown. This is in stark contrast to the thermal midpoints observed by Spinach alpha [23] which rises 20+°C in the presences of ATPYS similar to that observed by tobacco R294V. This demonstrates that that like the ATPase SA rate the thermal Midpoint in the presence of ATPYS could be used to predict hexameric Rca. Interestingly in the presence of ADP, Tobacco alpha has a higher thermal midpoint of ~40°C than the R294V mutants who like the Spinach alpha have a midpoint of ~30°C.

This demonstrates that Rca that are hexameric with ATP are super sensitive and tight binding to ATP but desensitised to ADP which incites speculation. In future research on an Rca from another species that requires a hexameric conformation for image, an assay using DSF in the presence of ADP and ATPYS could be a way to screen for hexamer forming mutants. This would be much cheaper than trying to find their molecular weights in the presence of ATPYS although less reliable.

4.4 AA+ domain.

4.4.a ATP bind causes a partially unfolded motif on a kratky plot.

Displaying the Kratky plots of each of the constructs with or without nucleotide shows that in the presence of ATPYS the post curve plateau created is higher than in ADP or Apo conditions. This motif is generally known to mean that there is a higher level

of flexibility in the protein being imaged. It is hypothesised rubisco activase undergoes structural changes with the binding of ATP and the N-terminal domain could activate by becoming the “molecular lasso” for rubisco binding and adhesion. This observation follows this theory except for the fact the the N-terminal truncated R294V also exhibit this risen plateau demonstrating that the changes are probably occurring to the central AAA+ domain. This result is also at odds with the data showing that with the binding of ATPYS the thermal stability increases, especially with the R294V mutants. To understand more about this phenomenon crystal structures could be gained with Rca in the presence of ADP and ATPYS and the two compared. Future C-EM experiments could also elucidate more on this. For a better image of the AAA+ domain an Rca that forms stable hexamers in the absence of nucleotide or with ADP would be ideal for imaging purposes.

Chapter 5: Conclusion.

5.1.a The conclusion.

The hypothesis was that the addition and subtraction at the terminal regions would have no effect on the quaternary structure and ATPase rate of Rca. The results show that while changes did occur the overall result was that these modifications were not catastrophic to these aspects of Rca. From the results then it is concluded that large changes such as these to the N and C-terminal regions modulate the ATPase rate and quaternary assembly but do not completely render the central domain ineffectual in these aspects.

5.1.b future imaging experiments.

The results show that with the admission of the dynamic terminal regions especially the N-terminal region the core αβ and α domains are stable enough to imaged as shown by previous crystallization studies and our new SAX data. The N-terminal truncated construct still forms quaternary structure of the same arrangement as the original Rca. The R294V mutation causes a hexameric structure that is vital for imaging due to forming uniform quaternary states that don't vary in size and the hexamer formation is symmetrical increasing the power of imaging techniques. This provides a unique opportunity to obtain better images of the central domain that will be vital in future research utilizing Cryogenic electron microscopy. The terminal regions remain dynamic and hard to image but there is still much to understand concerning the the central AAA+ domain as demonstrated by the possible dynamic structural change shown in the kratky plots. Much future research on Rca may focus on mass imaging where Rca is captured in various different states and different forms/mutations and then these structures are compared to obtain an understanding of how Rca functions.

Chapter 6: References.

1) Edmondson, D., Badger, M. and Andrews, T. (1990) A Kinetic Characterization of Slow Inactivation of Ribulosebisphosphate Carboxylase during Catalysis. PLANT PHYSIOLOGY **93**, 1376-1382.

2) Pearce, F. (2006) Catalytic by-product formation and ligand binding by ribulose bisphosphate carboxylases from different phylogenies. *Biochem. J.* **399,** 525-534

3) Bauwe, H., Hagemann, M. and Fernie, A. (2010) Photorespiration: players, partners and origin. Trends in Plant Science **15**, 330-336.

4) Tcherkez, G. (2015) The mechanism of Rubisco-catalysed oxygenation. Plant, Cell Environment **39**, 983-997.

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

6) Salvucci, M., DeRidder, B. and Portis, A. (2006) Effect of activase level and isoform on the thermotolerance of photosynthesis in Arabidopsis. Journal of Experimental Botany **57**, 3793-3799.

7) Whitney, S., Houtz, R. and Alonso, H. (2010) Advancing Our Understanding and Capacity to Engineer Nature's CO2-Sequestering Enzyme, Rubisco. PLANT PHYSIOLOGY **155**, 27-35.

8) Lin, M., Occhialini, A., Andralojc, P., Parry, M. and Hanson, M. (2014) A faster Rubisco with potential to increase photosynthesis in crops. Nature **513**, 547-550.

9) Mueller-Cajar, O., Stotz, M. and Bracher, A. (2013) Maintaining photosynthetic CO2 fixation via protein remodelling: the Rubisco activases. Photosynthesis Research **119**, 191-201.

10)Salvucci, M., Osteryoung, K., Crafts-Brandner, S. and Vierling, E. (2001) Exceptional Sensitivity of Rubisco Activase to Thermal Denaturation in Vitro and in Vivo. PLANT PHYSIOLOGY **127**, 1053-1064.

11) Esau, B., Snyder, G. and Portis, Jr., A. (1996) Differential Effects of N- and C-Terminal Deletions on the Two Activities of Rubisco Activase. Archives of Biochemistry and Biophysics **326**, 100-105.

12) Stotz, M., Mueller-Cajar, O., Ciniawsky, S., Wendler, P., Hartl, F., Bracher, A. and Hayer-Hartl, M. (2011) Structure of green-type Rubisco activase from tobacco. Nature Structural Molecular Biology **18**, 1366-1370.

13) Miziorko, H. and Lorimer, G. (1983) Ribulose-1,5-bisphosphate carboxylase-oxygenase. Annual Review of Biochemistry **52**, 507-535.

14) Stec, B. (2012) Structural mechanism of RuBisCO activation by carbamylation of the active site lysine. *Proceedings of the National Academy of Sciences* **109,** 18785-18790

15) Tabita, F. (1999) Microbial Ribulose 1,5-bisphosphate carboxylase/oxygenase: a different perspective. photosynthesis research **60,** 1-28.

16) Salvucci, M., Portis, A. and Ogren, W. (1985) A soluble chloroplast protein catalyzes ribulosebisphosphate carboxylase/oxygenase activation in vivo. Photosynthesis Research **7**, 193-201.

17) Robinson, S. and Portis, A. (1989) Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Activase Protein Prevents the in Vitro Decline in Activity of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase. PLANT PHYSIOLOGY **90**, 968-971.

18) Snider, J., Thibault, G., and Houry, W. (2008) The AAA+ superfamily of functionally diverse proteins. *Genome Biol* **9,** 216

19) Henderson, J., Kuriata, A., Fromme, R., Salvucci, M. and Wachter, R. (2011) Atomic Resolution X-ray Structure of the Substrate Recognition Domain of Higher Plant Ribulose-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase. Journal of Biological Chemistry **286**, 35683-35688.

20) Salvucci, M. (1992) Subunit interactions of Rubisco activase: Polyethylene glycol promotes self-association, stimulates ATPase and activation activities, and enhances interactions with Rubisco. Archives of Biochemistry and Biophysics **298**, 688-696.

21) Keown, J., Griffin, M., Mertens, H., and Pearce, F. (2013) Small Oligomers of Ribulose-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase Are Required for Biological Activity. *Journal of Biological Chemistry* **288,** 20607-20615

22) Chakraborty, M., Kuriata, A., Nathan Henderson, J., Salvucci, M., Wachter, R. and Levitus, M. (2012) Protein Oligomerization Monitored by Fluorescence Fluctuation Spectroscopy: Self-Assembly of Rubisco Activase. Biophysical Journal **103**, 949-958.

23) Keown, J., and Pearce, F. (2014) Characterization of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase activase isoforms reveals hexameric assemblies with increased thermal stability. *Biochem. J.* **464,** 413-423

24) Kuriata, A., Chakraborty, M., Henderson, J., Hazra, S., Serban, A., Pham, T., Levitus, M. and Wachter, R. (2014) ATP and Magnesium Promote Cotton Short-Form Ribulose-1,5-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase Hexamer Formation at Low Micromolar Concentrations. Biochemistry **53**, 7232-7246.

25) Portis, A. (2003) Rubisco activase - Rubisco's catalytic chaperone. Photosynthesis Research. **75**, 11-27.

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

27) Portis, A., Li, C., Wang, D. and Salvucci, M. (2008) Regulation of Rubisco activase and its interaction with Rubisco. Journal of Experimental Botany **59**, 1597-1604.

28) Aubin-Tam, M., Olivares, A., Sauer, R., Baker, T. and Lang, M. (2011) Single-Molecule Protein Unfolding and Translocation by an ATP-Fueled Proteolytic Machine. Cell **145**, 257-267.

29) Hasse, D., Larsson, A. and Andersson, I. (2015) Structure of Arabidopsis thaliana Rubisco activase. International Union of Crystallography **71**, 800-808.

30) Werneke, J., Chatfield, J. and Ogren, W. (1989) Alternative mRNA Splicing Generates the Two Ribulosebisphosphate Carboxylase/Oxygenase Activase Polypeptides in Spinach and Arabidopsis. The Plant Cell **1**, 815.

31) Salvucci, M. and van de Loo, F. (2003) Two isoforms of Rubisco activase in cotton, the products of separate genes not alternative splicing. Planta **216**, 736-744.

32) Qian, J. and Rodermel, S. (1993) Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Activase cDNAS from Nicotiana tabacum. Plant Physiology **102**, 683-684.

33) Ott, C., Smith, B., Portis, A. and Spreitzer, R. (2000) Activase Region on Chloroplast Ribulose-1,5-bisphosphate Carboxylase/Oxygenase. Journal of Biological Chemistry **275**, 26241-26244.

34) Li, C., Salvucci, M. and Portis, A. (2005) Two Residues of Rubisco Activase Involved in Recognition of the Rubisco Substrate. Journal of Biological Chemistry **280**, 24864-24869.

35) Tsai, Y., Lapina, M., Bhushan, S. and Mueller-Cajar, O. (2015) Identification and characterization of multiple rubisco activases in chemoautotrophic bacteria. Nature Communications **6**, 8883.

36) van de Loo, F. and Salvucci, M. (1996) Activation of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase (Rubisco) Involves Rubisco Activase Trp16. *Biochemistry* **35**, 8143-8148

37) Zhang, N. and Portis, A. (1999) Mechanism of light regulation of Rubisco: A specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f. Proceedings of the National Academy of Sciences **96**, 9438-9443.

38) Carmo-Silva, A. and Salvucci, M. (2013) The Regulatory Properties of Rubisco Activase Differ among Species and Affect Photosynthetic Induction during Light Transitions. PLANT PHYSIOLOGY **161**, 1645-1655.

39) Zhang, N., Schurmann, P. and Portis, A. (2001) Characterization of the regulatory function of the 46-kDa isoform of Rubisco activase from Arabidopsis. Photosynthesis Research **68**, 29-37.

40) Barta, C., Carmo-Silva, A. and Salvucci, M. (2011) Purification of Rubisco Activase from Leaves or after Expression in Escherichia coli. Methods in Molecular Biology 684, 363-374.

41) Tang, W., Li, D., Li, C., Esser, L., Dai, R., Guo, L. and Xia, D. (2010) A novel ATP-dependent conformation in p97 N–D1 fragment revealed by crystal structures of disease-related mutants. The EMBO Journal **29**, 2217-2229.