Exploring triosephosphate isomerase diversity in photosynthetic organisms

Gareth Jones

2022

School of Biological Sciences

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Master of Science

Abstract

Triosephosphate isomerase (TPI) is one of the most well characterised enzymes in the BRaunschweig ENzyme DAtabase (BRENDA). TPI is widely investigated in the context of glycolysis/gluconeogenesis, however, in photosynthetic organisms TPI is also involved in the regeneration phase of the Calvin-Benson-Bassham (CBB) cycle. Cyanobacteria possess a single TPI involved in both pathways, however, plants and many algae possess two TPI isoforms, a cytoplasmic isoform and a chloroplastic isoform, that are involved in glycolysis/gluconeogenesis and the CBB cycle respectively. This TPI organisation provides an opportunity to investigate how photosynthesis and adaptation to the chloroplast has shaped the evolution of TPI. For instance, the TPI inhibitor, 2-phosphoglycolate (2PG) is produced in the chloroplast due to oxygenation activity of ribulose 1,5 bisphosphate carboxylase/oxygenase (rubisco).

This project gains insight into TPI evolution by investigating the TPIs of three photosynthetic, and one derived non-photosynthetic, species. The organisms investigated were; the model cyanobacterium *Synechocystis* sp. PCC6803, the model plant *Arabidopsis thaliana*, the red alga *Porphyra umbilicalis*, and the non-photosynthetic parasitic plant *Cuscuta australis*. The TPIs from these organisms were characterised using structure/function and kinetic paradigms.
The characterisation of *Synechocystis* and *A. thaliana* TPIs were used to establish workflow and to generate a baseline to compare the underexplored TPIs of *C. australis* and *P. umbilicalis*.
The *C. australis* TPIs were investigated in the context of relaxed purifying selection resulting from adaptation of *C. australis* to parasitism and subsequent loss of photosynthesis. I reported that relaxed purifying selection has eroded the activity of *C. australis* TPIs compared to *A. thalaina* orthologues. Additionally, the structures of both *C. australis* cytoplasmic and chloroplast TPIs were solved by X-ray crystallography to identify where in the structure relaxed purifying selection has acted.
The *P. umbilicalis* TPIs were investigated in the context of 2PG inhibition as the rubisco of red alga are infamous for their low rate of rubisco oxygenation activity. I have demonstrated that there is no difference in 2PG inhibition between *P. umbilicalis* and *A. thaliana* TPIs. However, unexpectedly I identified that the chloroplast TPI of *P. umbilicalis* is exquisitely sensitive to oxidation.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Wayne Patrick. Your guidance, support and expertise has been invaluable. It has been a privilege to work in your lab and learn from you.
To Dr. Chelsea Vickers, thank you for your guidance, support and patience over the past year.
To the rest of the MME lab thank you all for your positivity and assistance, you have made this lab a wonderful place to study.
Thank you to my friends and flatmates for your emotional support through the ups and downs of the year.
Finally, thank you to my parents, without your emotional and financial support I could not have pursued my masters and made it to where I am today. I am eternally grateful for all that you have done for me.

Table of Contents

List of abbreviations used in text

Standard abbreviations used in this thesis are written according to the nomenclature in the *Journal of Biological Chemistry*; other non-standard abbreviations are listed below.

2PG	2-phosphoglycolate
3PGA	3-phosphoglycerate
BRENDA	BRaunschweig ENzyme DAtabase
Cary	Cary series UV-Vis spectrophotometer
c*Ath*TPI	cytoplasmic TPI of *Arabidopsis thaliana*
CBB cycle	Calvin-Benson-Bassham cycle
c*Cau*TPI	cytoplasmic TPI of *Cuscuta australis*
c*Psa*TPI	cytoplasmic TPI of *Pisum sativum*
c*Pum*TPI	cytoplasmic TPI of *Porphyra umbilicalis*
CV	column volume
DHAP	dihydroxyacetone phosphate
FTS	fluorescence thermal shift
GAP	D-glyceraldehyde 3-phosphate
*Gga*TPI	*Gallus gallus* TPI
GPDH	α-glycerophosphate dehydrogenase
His_6	N-terminal hexahistidine
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
k_{cat}/K_M	catalytic efficiency
K_{cat}	turnover number
K_i	inhibition constant
K_M	Michaelis-Menton constant
LB	Luria-Bertani
MSA	multiple sequence alignment
p*Ath*TPI	chloroplast TPI of *Arabidopsis thaliana*
p*Cau*TPI	chloroplast TPI of *Cuscuta australis*

pLDDT	predicted local distance difference test
p*Psa*TPI	chloroplast TPI of *Pisum sativum*
p*Pum*TPI	chloroplast TPI of *Porphyra umbilicalis*
RuBP	ribulose 1,5-bisphosphate
Rubisco	ribulose 1,5-bisphosphate carboxylase/oxygenase
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SEM	standard error of the mean
*Spc*TPI	TPI of *Synechocystis* sp. PCC6803
TB	terrific-broth
TEA	triethylamine
T_m	melting temperature
TPI	triosephosphate isomerase
V_{max}	maximum velocity
w/v	weight/volume

Chapter 1. Introduction; exploring triosephosphate isomerase diversity in photosynthetic organisms

1.1 The current state of enzymology

The field of enzymology was founded 120 years ago and since its founding enzymologists have generated vast quantities of information on the structures, functions and mechanisms of thousands of enzymes (Vickers *et al.,* 2021). However, our current knowledge in enzymology tends not to reflect the astounding diversity of life. As with many other fields of research, enzymology is biased towards a few biochemically well characterised organisms. For instance, five mammals, including *Homo sapiens*, contribute to 15 % of entries in the BRaunschweig ENzyme DAtabase (BRENDA) despite mammals comprising ~0.0006% of species on Earth (Vickers *et al.,* 2021). Other well characterised species include the model bacterium *Escherichia coli*, the baker's yeast *Saccharomyces cerevisiae* and the model plant *Arabidopsis thaliana*. However, advances in gene sequencing and synthesis technology means that now almost any enzyme from any organism can be recombinantly expressed and studied. The Patrick lab is leveraging this wealth of genomic information to investigate and compare the biochemistries of enzymes across vast evolutionary distances and begin to generate a more expansive view of biochemical diversity present in the natural world (Ferla *et al.,* 2017).

1.2 Triosephosphate isomerase as a model system

Triosephosphate isomerase (TPI) is a highly conserved glycolytic/gluconeogenic enzyme, which due to the centrality of carbon metabolism, lies at the heart of metabolic networks. TPI is often touted by biochemists as a 'perfect' enzyme where its catalytic rate is limited by the rate of diffusion (Knowles and Albery, 1977, Knowles, 1991). Since the first structure of a TPI was solved (Banner *et al.,*1975) key features which underpin the exquisite catalytic power of TPI have been identified, primarily using TPIs from model organisms and medically relevant human parasites (Wierenga *et al.,* 2010). Consequently, TPI is frequently used as a model system to understand enzymatic catalysis.

1.2.1 The mechanism of triosephosphate isomerases exquisite power

TPI catalyses the interconversion of two triose phosphates, dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP, Figure 1.1). TPIs are generally homodimeric with each subunit composed of and eightfold alternating pattern of α-helices and β-sheets which form a TIM-barrel fold (Banner *et al.,* 1975, Figure 1.2). The catalytic site is at the centre of the TIM-barrel with the catalytic residues present in loop 1 (N10, K12), loop 4 (H96) and loop 6 (E166). Residues are numbered here according to position in the cytoplasmic TPI of *A. thaliana*. The geometry of the active site is further shaped by loop 3 of the other subunit, in particular T76 of loop 3 is required to stabilise N10, K12 and H96 in the orientation necessary for catalysis (Figure 1.2). E166 is the key catalytic residue responsible for initial proton extraction from the substrate, while N10, K12 and H96 are necessary for the electrostatic stabilisation of the negatively charged reaction intermediates. In addition to key residues, known conformational changes are required for TPI catalytic activity, including the 'swinging in' of the catalytic glutamic acid, into a conformation poised for catalysis, and the ~ 7 Å movement of loop 6 (Figure 1.2). The closure of loop 6 'grips' the phosphate group of the substrate molecule and is required to exclude solvent molecules from the active site (Reviewed by Wierenga *et al.,* 2010). This exquisite molecular detail has primarily been determined through thorough investigations on a relatively small number of TPIs, particularly those from the chicken, *Gallus gallus* (Herlihy *et al.,* 1976, Knowles and Albery, 1977, Zhang *et al.,* 1996) and the parasite *Trypanosoma brucei* (Wierenga, 1991, Go *et al.,* 2010, Liao *et al.,* 2018).

This figure is unavailable.

Figure 1.1, Structure of the TPI substrates and inhibitor. TPI catalyses the interconversion of dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP). The transition state analogue 2-phosphoglycolate (2PG) is a potent inhibitor of TPI. Figure taken from Casteleijn *et al.* 2006.

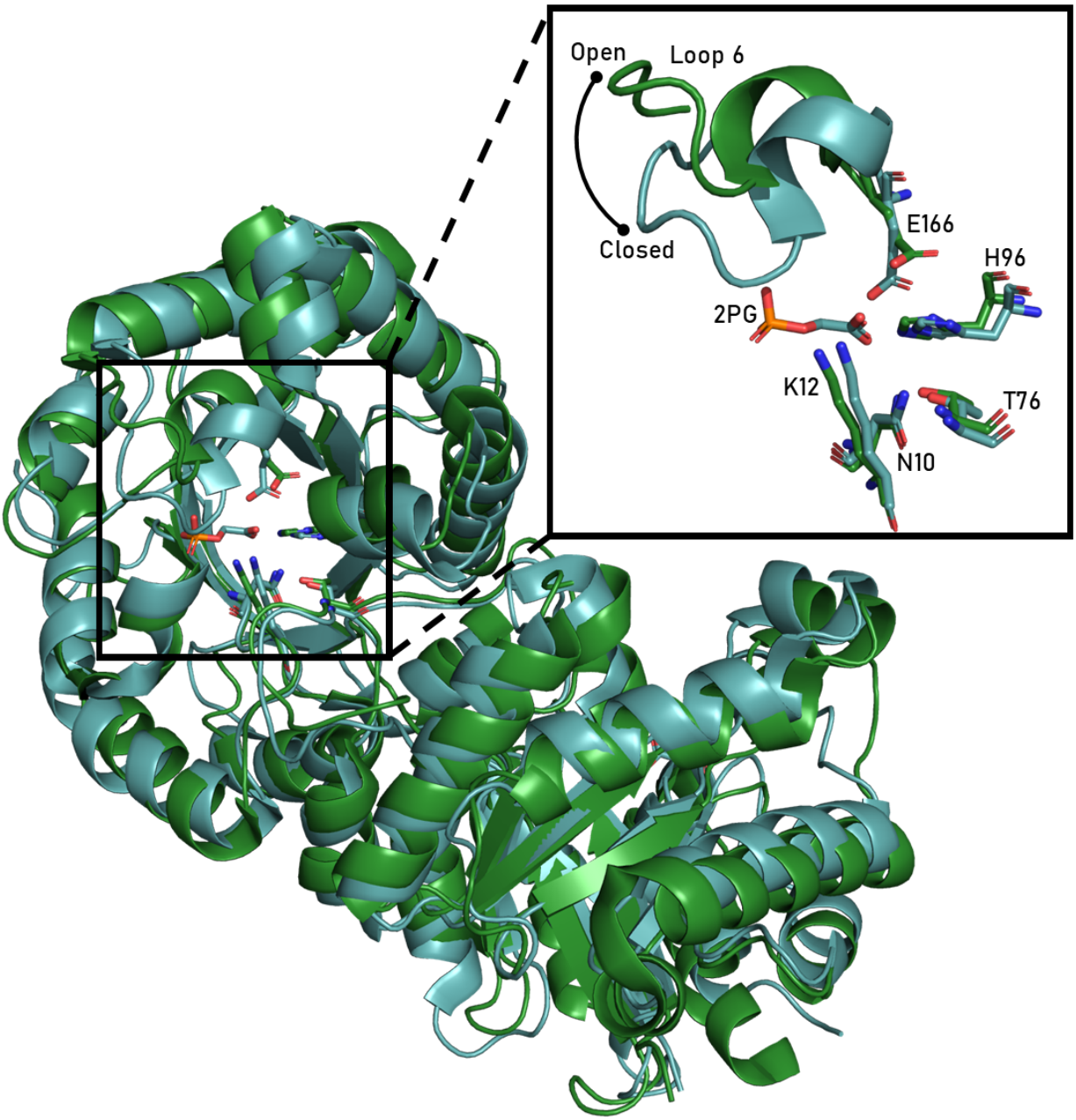


Figure 1.2, Structure of TPI. Overlay of *Synechocystis* sp. PCC6803 TPI (PDB ID: 6BVE) in teal and the cytoplasmic TPI of *Arabidopsis thaliana* (PDB ID: 4OBT) in dark green. The key catalytic residues are displayed as sticks and numbered according to their position in the *A. thaliana* cytoplasmic TPI. The TPI inhibitor 2-phosphoglycolate (2PG) is present in the active site of *Synechocystis* TPI.

1.3 Triosephosphate isomerase in the context of glycolysis/gluconeogenesis

TPI is encoded by the gene *tpiA* which is near universally conserved across the tree of life, where it lies at the heart of the glycolysis and gluconeogenesis. It is in this context that TPI has been most widely studied.

In glycolysis, the TPI catalysed conversion of DHAP to GAP is required for a net gain of two equivalents of ATP, while in gluconeogenesis TPI is required to generate both GAP and DHAP required as substrates for the fructose 1,6-bisphosphatase catalysed reaction (Figure 1.3). In both pathways TPI is unregulated. As TPI is orders of magnitude more efficient than any other enzyme in either pathway it is able to rapidly adjust the pools of DHAP and GAP to reach equilibrium. Consequently, TPI is capable of rapidly adjusting the GAP and DHAP concentrations in response to the activity of other enzymes in the pathways. Additionally, the production of GAP and DHAP facilitates the interconnectivity of glycolysis/gluconeogenesis with other metabolic pathways including lipid metabolism and the pentosephosphate pathway (Richard, 2008).

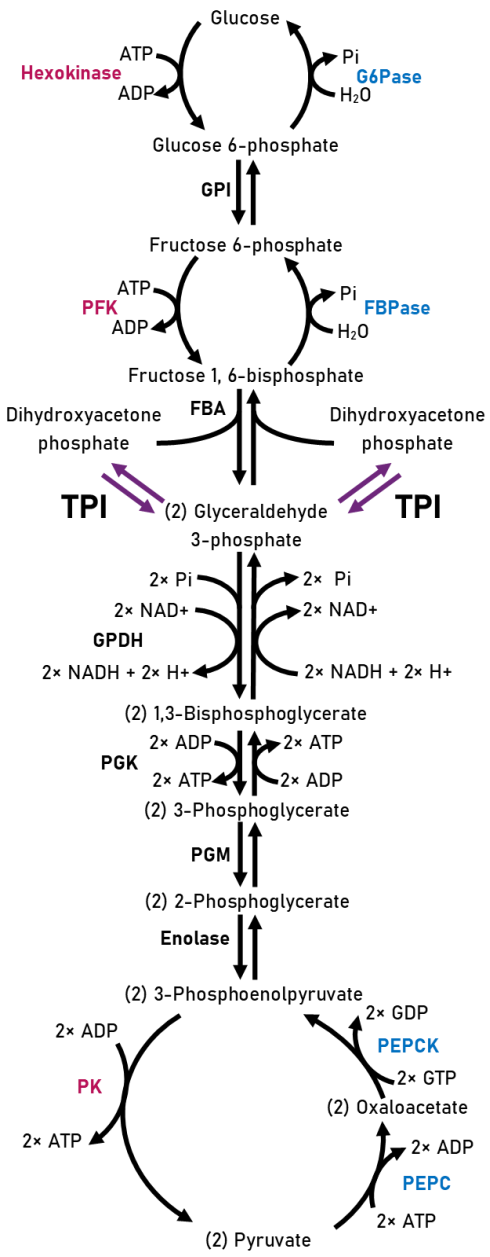


Figure 1.3, Glycolysis/gluconeogenesis. Glycolysis specific enzymes are in red while gluconeogenesis specific enzymes are in blue. The triosephosphate isomerase catalysed reaction is shown with purple arrows. G6Pase, glucose 6-phosphatase. GPI, glucosephosphate isomerase. PFK, phosphofructokinase. FBPase, fructose 1,6-bisphosphatase. FBA, fructose 1,6-bisphosphate aldolase. TPI, triosephosphate isomerase. GPDH, α-glycerophosphate dehydrogenase. PGK, phosphoglycerate kinase. PGM, phosphoglycerate mutase. PK, pyruvate kinase. PEPCK, phosphoenolpyruvate carboxykinase. PEPC, phosphoenolpyruvate carboxylase.

1.4 Triosephosphate isomerase in the context of photosynthesis

TPI is required in the regeneration phase of the Calvin-Benson-Bassham (CBB) cycle. Oxygenic photosynthesis can be divided into two sets of reactions. The first set of reactions, the light reactions, use light energy to split water into its constituent atoms. The resulting oxygen is released as a waste product in the form of O_2 while the protons are used to generate ATP and the cofactor NADPH, which are required for the second set of reactions, the CBB cycle.

The CBB cycle fixes CO_2 as sugars and can be split into three phases, the carbon fixation phase, the reduction phase, and the regeneration phase (Figure 1.4). In the fixation phase CO_2 is fixed to ribulose 1,5-bisphosphate (RuBP) by the enzyme ribulose bisphosphate carboxylase/oxygenase (rubisco) producing two 3-phosphoglycerate (3PGA). In the reduction phase 3PGA is reduced to GAP which is then fed into central carbon metabolism or used in the subsequent regeneration phase. The regeneration phase consists of a complex series of reactions whereby RuBP is regenerated from GAP. In this phase the interconversion of GAP and DHAP by TPI is required to drive subsequent reactions catalysed by transketolase or fructose 1-6, bisphosphate aldolase. The interplay of these reactions is complex and often the distinction between GAP and DHAP is not made, being frequently simplified to 'triose phosphates'(McClain & Sharkey, 2019, Schreier and Hibberd, 2019), consequently the critical function of TPI in the CBB cycle is often overlooked.

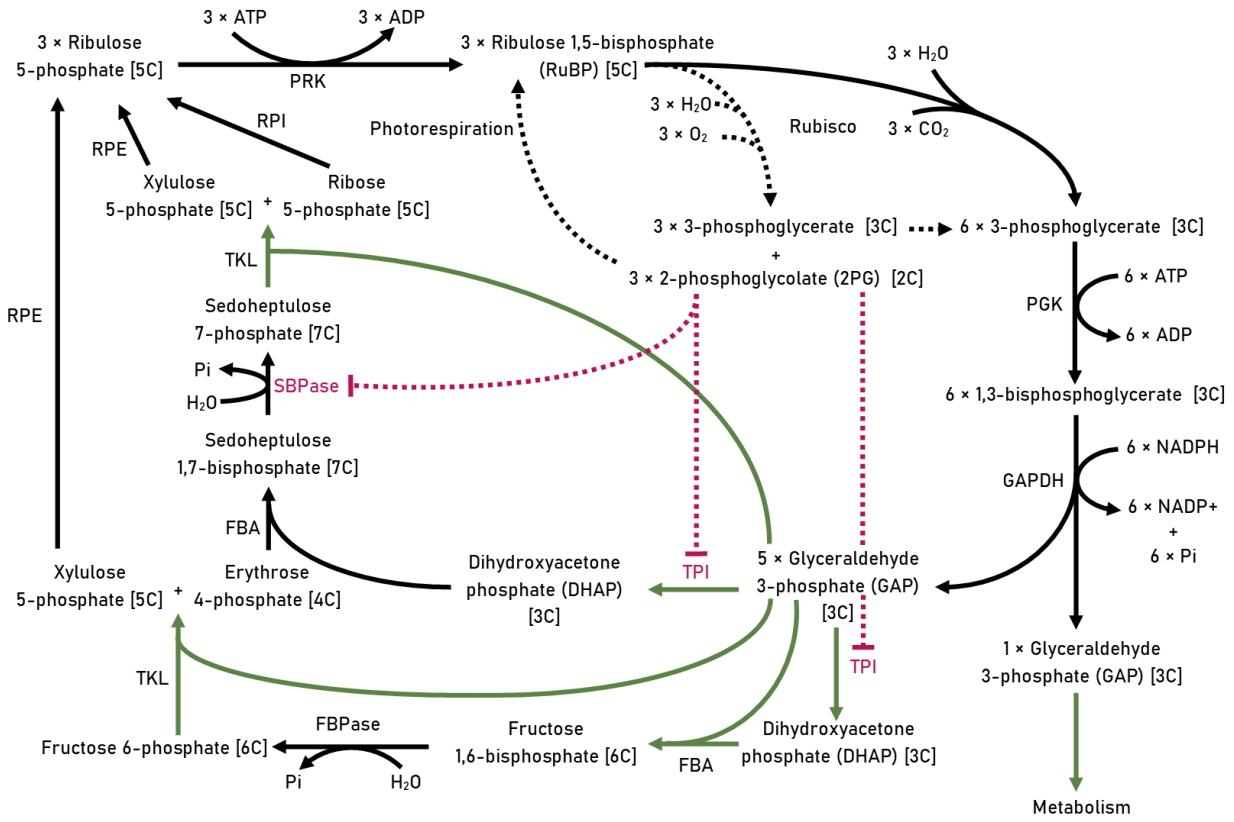


Figure 1.4, The Calvin-Benson-Bassham cycle and the impacts of 2-phosphoglycolate. Rubisco catalysed oxygenation and the resulting downstream reactions are shown by dashed arrows. For simplicity the steps of photorespiration are not shown. Enzyme inhibition by 2PG is shown in red. The fate of GAP is shown in green. PRK, phosphoribulose kinase. Rubisco, ribulose 1,5-bisphosphate carboxylase/oxygenase. PGK, 3-phosphoglycerate kinase. GAPDH, glyceraldehyde 3-phosphate dehydrogenase. TPI, triosephosphate isomerase. FBA, fructose 1,6-bisphosphate aldolase. FBPase, fructose 1,6-bisphosphatase. TKL, transketolase. SBPase, sedoheptulose 1,7-bisphosphatase. RPE, ribulose 3-phosphate epimerase. RPI, ribose 5-phosphate isomerase.

1.4.1 2-phosphoglycolate, a potent triosephosphate isomerase inhibitor produced during the CBB cycle

In addition to catalysing the carboxylation of RuBP, rubisco can also catalyse the oxygenation of RuBP, resulting in the production of 2-phosphoglycolate (2PG, Figure 1.4). The ratio of rubisco carboxylation to oxygenation activity is between 3:1 and 4:1 under atmospheric conditions (Johnson, 2016). Thus, the efficiency of photosynthesis is far less than the stoichiometry implied by Figure 1.4. 2PG is a transition state analogue and differs from GAP and DHAP by one carbon

(Lolis and Petsko, 1990, Figure 1.1). 2PG is a potent inhibitor of at least two enzymes in the CBB cycle, including TPI, and the glycolytic enzyme phosphofructokinase (Figure 1.4, Anderson, 1971, Kelly and Latzko, 1976, Flügel *et al.,* 2017). The 2PG concentration in the chloroplast is estimated to reach 100 µM, well above the reported K_i for *A. thaliana* TPIs (24-36 µM, Flügel *et al.,* 2017, Li *et al.,* 2019). 2PG can be recycled back into RuBP through the photorespiration pathway, however, this pathway requires ATP and only recovers 75% of carbon with 25% being released as CO_2 (Sharkey, 1988).

As 2PG production and photorespiration decreases the overall efficiency of photosynthesis, rubisco catalysed oxygenation is widely considered undesirable. Numerous efforts have been made to improve photosynthetic efficiency by increasing rubisco's selectivity for CO_2 over O_2, and reducing the impacts of 2PG production (Busch, 2020). However, an increasing body of evidence demonstrates that photorespiration is an important pathway in the biosynthesis of amino acids (Busch, 2020). Additionally, 2PG is proposed to regulate flux through the CBB cycle. Due to the high catalytic rate of TPI, the reaction it catalyses is often assumed to be at equilibrium in the chloroplast. At equilibrium the concentration of DHAP should be ~20 times greater than that of GAP (Sharkey and Weise, 2012). However, recent evidence found that the DHAP concentration of leaf extracts was only ~11 times greater than GAP (Li *et al.,* 2018). This disequilibrium of TPI is attributed to the accumulation of 2PG in the chloroplast resulting in partial TPI inhibition. This may be beneficial as the higher GAP concentration at disequilibrium would drive the fructose 1,6-bisphosphate aldolase catalysed reaction (Figure 1.4), which functions optimally when the DHAP and GAP concentrations are equal, thus increasing overall flux through the CBB cycle (Sharkey, 2018, Li *et al.,* 2018).

1.5 Duplication of the triosephosphate isomerase gene in photosynthetic eukaryotes

Photosynthetic prokaryotes possess a single TPI for both glycolysis/gluconeogenesis and the CBB cycle. In contrast, most photosynthetic eukaryotes possess two TPI isoforms, a cytoplasmic and a chloroplast isoform involved in glycolysis/gluconeogenesis the CBB cycle respectively (Feierabend *et al.,* 1990). The chloroplast TPI is nuclear encoded and expressed with an N-terminal chloroplast localisation tag (Henze *et al.,* 1994). This tag enables translocation into the chloroplast and is then cleaved following entry (von Heijne *et al.,* 1989). There are known

exceptions to this subcellular organisation of TPIs in eukaryotes. The green alga *Chlamydomonas reinhardtii* and the thermophilic red alga *Cyanidioschyzon merolae* possess a single TPI and have their glycolytic pathway split between the chloroplast and the cytoplasm (Johnson and Alric, 2013, Moriyama *et al.,* 2014).

Although the chloroplast is derived from the ancient endosymbiosis of a proto-cyanobacterium, the chloroplast TPI is of eukaryotic origin, derived from a duplication of the preexisting cytoplasmic TPI (Henze *et al.,* 1994, Plaxton, 1996). This duplication event occurred early in the viridiplantae lineage (Henze *et al.,* 1994). Although both the cytoplasmic and chloroplast TPIs of red algae are of eukaryotic origin (Sun *et al.,* 2008), it remains unclear if this duplication, and subsequent replacement, occurred before or after the divergence of red and green algae.

It is hypothesised that the duplicated eukaryotic TPI replaced the prokaryotic TPI due to the greater redox sensitivity of eukaryotic TPIs enabling for the modulation of TPI activity in response to the reactive oxygen species which are produced during the light reactions of photosynthesis (Dumont *et al.,* 2016, Castro-Torres *et al.,* 2018). This hypothesis is consistent with phylogenomic studies which have detected a global expansion in the redox sensitive proteome coinciding with the endosymbiotic event that gave rise to the chloroplast (Woehle *et al.,* 2017).

1.6 The photosynthetic organisms chosen for investigation

In this project I have investigated seven TPIs from four different photosynthetic organisms; the cyanobacterium *Synechocystis* sp. PCC6803, the model plant *Arabidopsis thaliana*, the non-photosynthetic parasitic plant *Cuscuta australis*, and the red alga *Porphyra umbilicalis* (Figure 1.5).

Synechocystis (Figure 1.5 A) and *A. thaliana* (Figure 1.5 B) are model organisms which have been widely used to investigate photosynthesis in prokaryotes and eukaryotes respectively. Accordingly, much of the existing knowledge on TPIs in photosynthetic organisms has come from thorough investigation of these two species (Dumont *et al.,* 2016, López-Castillo *et al.,* 2016, Castro-Torres *et al.,* 2018 and 2019).

C. australis is a member of the parasitic plant family commonly known as dodders (Figure 1.5 C). There are over 200 known species of dodder, all of which are obligate stem parasites and many of which parasitise important crop species (Dawson *et al.,* 1994). Dodders diverged from

their last common photosynthetic ancestor, the morning glories, some 33 million years ago (Sun *et al.,* 2018). Since diverging from the morning glories *Cuscuta* have undergone extensive morphological degeneration, lacking both roots and leaves and only containing trace amounts of chlorophyll. Consequently, dodders can no longer sustain themselves by photosynthesis, although photosynthetic capacity does differ depending on species (Yoshida *et al.,* 2018, Nickrent, 2020).

Porphyra are a genus of red algal seaweeds which are widely cultivated as a foodstuff, with *P. umbilicalis* (Figure 1.5 D) being harvested in the United Kingdom and Ireland to make laverbread. In addition to use as a food source, many species of red algae have been identified as sources of natural products, particularly polysaccharides, such as agar and carrageenans, which are of medical and economic importance (Minhas *et al.,* 2021). Red algae are of particular interest to molecular biologists due to the efficiency of their rubiscos, which are twice as selective for CO_2 over O_2 than those of C3 crop plants (Flamholz *et al.,* 2019).

This figure is unavailable

Figure 1.5, Organisms Investigated in this project. A. *Synechocystis* sp. PCC6803. Image taken from Touloupakis *et al.* 2016. **B.** *Arabidopsis thaliana*. Image taken from Krämer 2015. **C.** *Cuscuta australis* twining around the stem of wild tomato and *C. australis* flowers. Image modified from Sun *et al.* 2018. **D.** *Porphyra umbilicalis*. Image taken from Brawley *et al.* 2017.

1.7 Aims

In this project I have characterised seven TPIs from four different photosynthetic organisms; the cyanobacterium *Synechocystis* sp. PCC6803, the model plant *Arabidopsis thaliana*, the non-photosynthetic parasitic plant *Cuscuta australis*, and the red alga *Porphyra umbilicalis*. The TPIs were characterised using structure/function and kinetic paradigms. The comparison of cytoplasmic and chloroplast TPIs, both from a single species and between species, has enabled the inference of how photosynthesis has shaped TPI evolution.

The first specific aim was to establish protocols for producing and assaying TPIs, using the *Synechocystis* and *A. thaliana* enzymes. These enzymes were selected as there is existing data for them in the literature to benchmark this investigation against. The experiments addressing this aim are described in chapter three.

The second aim of this project was to characterise the TPIs of the parasitic plant *C. australis*. I hypothesised that, due to little to no photosynthesis occurring in the chloroplast of *C. australis*, relaxed purifying selection acting upon the chloroplast TPI would result in the accumulation of mutations which would erode its catalytic efficiency. The experiments addressing this aim are described in chapter four.

The final aim of this project was to characterise the TPIs of the red algal seaweed *P. umbilicalis*. I hypothesised that due to the increased selectivity of red algal rubiscos for CO_2 over O_2 and consequently reduced production of 2PG, the chloroplast TPI of *P. umbilicalis* would be more sensitive to 2PG inhibition. The experiments addressing this aim are described in chapter five.

Chapter 2. Materials and Methods

2.1 Bacterial expression of recombinant proteins

2.1.1 Bacterial strains

The chemically competent *E. coli* strain E. cloni 10G (Lucagen) was used as a storage strain for protein expression vectors. The chemically competent *E. coli* strain LOBSTR BL21(DE3) (Kerafast) was used to highly express recombinant TPIs for purification.

2.1.2 Designing protein expression vectors

The genes used in this project (Appendix A) were synthesised by the commercial supplier, Twist Bioscience, in pET-28a(+) expression vectors with an N-terminal hexahistidine (His_6) tag and under the control of the isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible T7 promoter. When delivered the plasmids were resuspended in nuclease free water to a final concentration of approximately 100 ng/µL before being used to transform the appropriate bacterial strains. Chloroplast localised TPIs are nuclear encoded and are expressed with an N-terminal chloroplast localisation tag which is cleaved following chloroplast entry (von Heijne *et al.,* 1989, Henze *et al.,* 1994). The programme TargetP 2.0 (Armenteros *et al.,* 2019) was used to identify the tag cleavage sites and the upstream residues were excluded from the expression constructs (Appendix B). TargetP 2.0 is not optimised for use in Rhodophyta, consequently for the chloroplast TPI of *P. umbilicalis* TargetP 2.0 identified a 36 amino acid region which contained the probable cleavage (Appendix B). Ultimately, the cleavage site was inferred based on multiple sequence alignment (MSA), with residues upstream of the first methionine in the 36 amino acid region identified by TargetP 2.0 being excluded from the final expression construct (Appendix A, B).

2.1.3 Bacterial cell culture media

E. coli were grown in both Luria-Bertani (LB) and terrific-broth (TB) media (ForMedium). Media was prepared as per manufacturer's instructions and using type 1 ultra pure water from a Barnstead system. Media was sterilised by autoclaving using a Systec DX-65 autoclave and was

then stored at room temperature until use. Solid media was prepared by adding 1.5 % agar (New Zealand Seaweeds) weight to volume (w/v) prior to autoclaving and plates being poured. Once set agar plates were stored at 4 °C.

2.1.4 Antibiotic selection

All bacterial cell culture media was supplemented with kanamycin (Kan, 30 µg/mL) to select for *E.coli* containing pET-28a(+) vectors and consequently a Kan resistance gene.

2.1.5 Transformation

50 µL aliquots of chemically competent 'Mix & Go!' (Ngaio diagnostics) *E. coli* was transformed with 50 ng of plasmid DNA. Initially the *E. coli* and plasmid DNA were mixed thoroughly, as per manufacturer's instructions, prior to plating on LB agar supplemented with Kan. However, due to persistent problems with 'Mix & Go!' cells losing competency, after mixing, 200 µL of LB was added to the cells which were then incubated at 37 °C, with constant agitation for ~1 hour. The cells were then spun down at 4 °C and ~4000g for 5 minutes and 200 µL of the culture media was removed. The pellet was then resuspended in the remaining media and plated on LB agar supplemented with Kan. The plates were then incubated at 37 °C overnight.

2.1.6 Bacterial cell culture

Single *E. coli* colonies were selected from agar plates and used to inoculate 5 mL of LB medium supplemented with Kan. The 5 mL cultures were incubated at 37 °C overnight. All liquid cultures were incubated in an Innova 44 incubator set to 200 rpm, this model does not provide information in g. The following morning the 5 mL cultures were used to generate glycerol stocks and/or to inoculate 500 mL expression cultures.

500 mL TB cultures, supplemented with 0.4 % glycerol and Kan, were used for protein expression. The TB culture was inoculated with a 5 mL overnight culture and incubated at 37 °C, 200 rpm until an O.D. 600 nm of ~0.4 was reached. The culture was then transferred to a 18 °C, 200 rpm incubator and left for ~1 hour to reach temperature. Once at 18 °C, TPI expression was induced using 0.5 mM of IPTG and the cultures were then incubated overnight (~16 hours). Cell culture was stopped by centrifugation at ~4000g for 25 minutes. The culture media was removed

and the resulting cell pellets were resuspended in the lysis buffer (Table 2.1) to a final volume of 10-20 mL. The resuspended cell pellets were then stored at -80 °C until protein purification.

2.1.7 Generating glycerol stocks

Glycerol stocks were made for both storage and protein expression *E. coli* strains. The glycerol stocks were made by combining 200 µL of autoclaved 80 % glycerol with 800 µL of a 5 mL overnight cell culture and resuspending thoroughly. Glycerol stocks were stored at -80 °C and were used to inoculate future 5 mL cultures.

2.1.8 Protein expression screen for insoluble TPIs

For TPIs which were insoluble under the standard expression conditions (Section 2.1.6) an expression trial was undertaken to try and improve solubility. A 500 mL expression culture was prepared as in section 2.6.1. 4 mL aliquots of the culture were taken at 0, 2, 4 and ~16 hours after IPTG induction. The cells were then spun out of the media at 4 °C and ~4000g for 10 minutes and the pellets were resuspended in 0.5 mL of lysis buffer (Table 2.1) and then stored at -80 °C until lysis. The *E. coli* were lysed on ice by sonication using a sonicator (Vibra-Cell) set to 20 % amplitude and cycles of 10 seconds on 30 seconds off. The ~16 hour sample was lysed using four cycles while the three other samples were lysed using two cycles.

2.1.9 Buffers

Table 2.1 lists the buffers used during the expression and purification of proteins in this project. Buffers were made using type 1 ultra pure water from a Barnstead system and then adjusted to pH 7.6 using 5 M NaOH or 12 M HCl. Prior to purification all buffers were filter sterilised using a Büchner funnel and 45 µm filter paper (Lab Supply) and then degassed for at least 15 minutes.

Table 2.1, Buffers used during TPI purification

Buffer	Composition
Lysis buffer/ IMAC buffer A	40 mM Tris pH 7.6, 300 mM NaCl, 20 mM Imidazole
IMAC buffer B	40 mM Tris pH 7.6, 300 mM NaCl, 400 mM Imidazole
SEC Buffer	40 mM Tris pH 7.6, 300 mM NaCl

2.2 Purification of recombinantly expressed proteins

2.2.1 Cell lysis

Cell pellets were thawed and then lysed by sonication (Vibra-Cell). Sonication was carried out on ice using 20 % amplitude and cycles of 30 seconds on 59 seconds off, for 3 to 6 cycles. The lysed cells were then clarified by centrifugation at 4 °C and ~24000g for 20 to 30 minutes.

2.2.2 Immobilised metal affinity chromatography

Immobilised metal affinity chromatography (IMAC) was carried out on a 5 mL HisTrap SP column (GE Healthcare) on an ÄKTA Start FPLC system. The HisTrap column was pre-equilibrated in IMAC buffer A (Table 2.1). The clarified supernatant was filtered through 45 and 22 µM filters (Interlab) prior to loading onto the column. Buffers were pushed through the column at a flow rate of 1 mL/min. The unbound fraction was eluted from the column using 25 mL of IMAC buffer A, subsequently His_6-tagged TPIs were eluted from the column in 1 mL fractions using a gradient of 0 to 100 % IMAC buffer B (Table 2.1, 20 to 400 mM imidazole) over 25 mL.

2.2.3 Regeneration of the HisTrap SP column

To avoid cross contamination between TPI purifications the HisTrap SP column was stripped and regenerated after each purification. The column was first washed with 4 column volumes (CV) of filter sterilised water, 2 CV of EDTA, 4 more CV of water, then 2 CV of 100 mM $NiCl_2$ and finally 4 more CV of water. If the column was to be used immediately for purification 4 CV of IMAC buffer A was then pushed onto the column, if not, 4 CV of 70 % ethanol was pushed onto the column which was then stored at 4 °C.

2.2.4 Size exclusion chromatography

Prior to size exclusion chromatography (SEC) the TPI containing fractions from IMAC were pooled and concentrated using a 10 kDa centrifugal concentrator (Amicon Ultra-15, Lab Supply). Spins were performed at 4 °C and ~2400 g in 10 minute intervals until a volume of approximately 0.5 mL was reached. The resulting sample was filtered using a 22 µM filter

(Interlab) and centrifuged at 4 °C and ~16000g for 10 minutes to remove any remaining aggregates. SEC was performed using an ÄKTA Pure and a S200 10/300 GL Increase column (GE healthcare) equilibrated in SEC buffer (Table 2.1). SEC was performed at a flow rate of 0.75 mL/min over 27 mL and SEC buffer was eluted from the column in 1 mL fractions. Following SEC the proteins were stored at -80 °C, unless otherwise stated, in the SEC buffer (Table 2.1).

2.2.5 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

The purity of proteins was assessed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE gels were cast in a Mini-Protean Tetra System (BioRad), using 40 % acrylamide, with an acrylamide:bisacrylamide ratio of 29:1 (BioRad) and 0.1 % of SDS. The gels were polymerised with 0.05 % ammonium persulfate and tetramethylethylenediamine.

1 µL of the protein sample was mixed with 10 µL of 2× SDS loading dye (Table 2.2) and SEC buffer (Table 2.1) to a final volume of 20 µL. Samples were then denatured by heating to 95 °C for 5 minutes prior to loading onto SDS-PAGE gel. ~15 µL of each sample were loaded onto a gel. Samples were run with 3 µL of Blue Protein Standard Broad Range (NEB) in order to approximate the molecular weight of the protein samples. Gels were run in a Mini-Protean Tetra Cell (BioRad) using 1× SDS-PAGE running buffer (Table 2.2) and run at 100 V for 30 to 40 minutes. The gels were then stained using Coomassie stain (Table 2.2) for ~30 minutes and then excess stain was removed using destain (Table 2.2) for ~16 hours. The destained gels were then imaged using a Gel Doc (UVITEC).

Table 2.2, Reagents used in SDS-PAGE

Reagent	Composition
2× SDS loading dye	150 mM Tris-Cl, pH 6.8, 6 % SDS (w/v), 30 % glycerol (v/v), 0.03 % bromophenol blue (w/v), 300 mM β-mercaptoethanol
1× SDS-PAGE running buffer	24.9 mM Trisbase, 192 mM glycine, 0.1 % SDS (w/v)
Coomassie Stain	2.5 $g.L^{-1}$ Coomassie Brilliant Blue R-250 (ThermoFisher), 50 % methanol (w/v), 10 % acetic acid (w/v)
Destain	10 % (w/v) acetic acid, 40 % (w/v) methanol

2.2.6 Determining protein concentration

For preliminary estimates of protein concentration a Nanodrop Spectrophotometer (ThermoFisher Scientific) was used. The absorbance was measured at 280 nm using 1 μL of a protein sample. The Beer-Lambert equation ($A = \varepsilon.c.l$) was used to determine protein concentration. The extinction coefficients (ε) were predicted using the ExPASy ProtParam tool (Gasteiger *et al.,* 2005). Dilutions were carried out as necessary to get an A280 of ~1.
More accurate estimates of protein concentration were done using the scanning function of a Cary Series UV-Vis Spectrophotometer (Cary, Agilent Technologies). The Cary was blanked and then baseline correction was done using the SEC buffer (Table 2.1). Protein samples were then diluted 1 in 10 and the absorbance was measured from 260 to 350 nm. Protein concentration was then calculated using the Beer-Lambert equation, as above.

2.3 Steady-state kinetic assays

2.3.1 General steady-state kinetic assay design

TPI kinetic assays were carried out using a α-glycerophosphate dehydrogenase (GPDH) coupled assay (Nickbarg and Knowles, 1988) and GAP as substrate. In this assay, TPI converts GAP to DHAP (Figure 1.1) which is then reduced to glycerol 3-phosphate by GPDH. GPDH uses NADH as a cofactor. The oxidation of NADH to NAD+ is specifically measured at 340 nm. Thus, the reduction in NADH signal is measured as a proxy for TPI activity and the slope of the linear range of the absorbance plot is used to determine the initial rate of TPI activity.
The general composition of the steady-state kinetic assays is summarised in Table 2.3. Adjustments to this assay design were made to determine GAP concentration (Section 2.3.2), TPI oxidation sensitivity (Section 2.3.4), and to determine TPI inhibition by 2PG (Section 2.3.5). All assays were carried out at 30 °C.

Table 2.3, General composition of steady-state kinetic assays

Reagent	Final Concentration
Trimethylamine (TEA, Sigma-Aldrich)	50 mM
D-glyceraldehyde 3-phosphate (GAP, Sigma-Aldrich)	0-5 mM
Nicotinamide adenine dinucleotide (NADH, Sigma-Aldrich)	0.25 mM
α-Glycerophosphate Dehydrogenase (GPDH, Sigma-Aldrich)	0.02 mg/mL
TPI	Variable

2.3.2 Enzymatic determination of GAP concentration

The substrate for kinetics assays, GAP, is sold as an 8-13 mg/mL solution which corresponds to 47-76 mM. To determine the GAP concentration more precisely the steady-state kinetic assay (Table 2.3) was modified by increasing the NADH concentration to 0.5 mM and decreasing the range of GAP from 0 to 0.4 mM, assuming a stock GAP concentration of 62 mM. The reactions were made up to 145 µL and initiated with 5 µL of undiluted TPI. The identity of the TPI was irrelevant, however, most reactions were carried out using the TPI of the thermophilic archaea *Pyrococcus furiosus*, which was left over from another project. The reactions were followed to completion and the initial and final absorbance at 340 nm was recorded using a Cary Series UV-Vis Spectrophotometer. Using a standard curve for NADH from 0 to 50 mM, diluted in 1:1 (v/v) TEA:deionised water, the initial and final NADH concentrations were calculated. As NADH and GAP are consumed in a 1:1 ratio in the coupled reaction, the change in NADH concentration is equal to the concentration of GAP added. This enabled the determination of the stock GAP concentration. The stock concentration of GAP was determined using at least two different concentrations of GAP in the assay each of which was measured in triplicate.

2.3.3 Steady-state kinetic assays for determining kinetic parameters

Assays for the determination of TPI kinetic parameters were carried out as described in Table 2.3 using 96 well U-bottom plates (Lab Supply) and a CLARIOstar Plus plate reader (BMG Labtech). Reactions were 100 µL and each GAP concentration was run in technical quadruplicate. An additional negative TPI control was run for each GAP concentration. Reactions took approximately 15-50 minutes to reach completion with data points taken every 13

to 15 seconds. Reactions were initiated by spiking 90 µL of the reagent mixture with 10 µL of TPI. The concentration of each TPI used in the assays was determined using trial assays with 0.2, 1 and 4 mM GAP. The TPI concentrations were selected so that reactions at the highest GAP concentrations reached completion in 15 to 50 minutes. The TPIs were diluted in SEC buffer (Table 2.1) to the appropriate concentration required for the assays.

For steady-state kinetic assays of the *P. umbilicalis* chloroplast TPI 10 mM DTT (Lab Supply) was added to dilutions of TPI. The final concentration of DTT in the assays was 1 mM. To minimise loss of activity the *P. umbilicalis* chloroplast TPI dilutions were made on ice immediately before initiating the assay.

The resulting initial rate data were fitted to the Michaelis-Menton equation using direct nonlinear fitting in GraphPad Prism 9 to determine kinetic parameters, v_{max}, K_M, k_{cat}, and k_{cat}/K_M. The reported kinetic parameters were determined using a minimum of three assays. For each set of assays individual TPIs were either expressed and purified once or on two independent occasions. This information is specified with the final data sets.

2.3.4 Assays assessing the oxidation sensitivity of *P. umbilicalis* TPIs

The oxidation sensitivity of *P. umbilicalis* TPIs was assessed using DTT time course assays. These assays were carried out in the Cary. Individual reactions were made to 150 µL according to table 2.3. Reactions were initiated with 15 µL of the *P. umbilicalis* TPIs. The change in NADH absorbance at 340 nm was then recorded for approximately four minutes. Immediately upon the thawing of TPI aliquots, TPI dilutions were made at room temperature using SEC buffer (Table 2.1) with or without 10 mM DTT. These dilutions were left at room temperature and aliquots were taken 0, 5, 10, 15 and 30 minutes after making up the appropriate dilution and used to initiate assays. The final concentration of DTT in assays was 1 mM. Each time point was measured in technical triplicate for each TPI for both positive and negative DTT.

2.3.5 2-phosphoglycolate inhibition assays

Inhibition assays using the TPI inhibitor 2-phosphoglycolate (2PG, Sigma-Aldrich) were carried out for the *A. thaliana* and *P. umbilicalis* TPIs. These assays were carried out as described in section 2.3.3 but with between 20 and 100 µM 2PG added to the reagent mixture. Two 2PG concentrations and a no 2PG baseline were run simultaneously with each combination of 2PG

and GAP concentration run in triplicate. The inhibition constant, K_i, was determined using direct nonlinear fitting in GraphPad Prism 9 to fit data to both the competitive model and the mixed model of enzyme inhibition. Identification of the model of inhibition which best fits the available data was done in GraphPad Prism 9 using the extra sum-of-squares F test and using the resulting F-statistic and the associated p-value to identify the best fitting model. Each reported K_i value was generated using data from a minimum of two assays. For each TPI, protein from the same purification was used in each assay.

2.4 Fluorescence thermal shift assays

Fluorescence thermal shift (FTS) assays were used to determine the melting temperature, T_m, of the *P. umbilicalis* TPIs. The FTS assays were made to 20 µL, using 10 µM of each TPI in SEC buffer (Table 2.1) and 5× Sypro Orange (Life Technologies) and were run in technical quadruplicate. Assays were run in MicroAmp™ fast optical 48-well reaction plates and sealed with MicroAmp™ optical adhesive film (ThermoFisher Scientific) in a StepOne™ real-time PCR system (Applied Biosystems). Fluorescence was monitored from 20 to 98 °C (ramp of 1 °C/min). The T_m was approximated by identifying the minima of the negative first derivative of the melt curve (-dF/dt) which was calculated in the StepOne™ software v2.1 (Applied Biosystems). In each assay conditions were run in technical quadruplicate. The T_m reported are the mean ± SEM of two biological replicates.

2.5 X-ray crystallography

2.5.1 Preparation of proteins for crystallography

Following IMAC and SEC, the TPIs were concentrated to between 15 and 40 mg/mL. The TPIs were concentrated using centrifugal concentrators (Amicon Ultra-0.5, Lab Supply) at 4 °C and ~16000g in 5-10 minute intervals. Following concentration, the protein solutions were filtered using 22 µM filters (Interlab) and then centrifuged at 4 °C and ~16000g to remove any aggregates.

2.5.2 High throughput crystal screening

Sitting-drop vapour diffusion crystallisation trials for the *C. australis* and *P. umbilicalis* TPIs were carried out using the commercial crystallisation screens, Index HT, Crystal Screen HT and PEG/Ion HT (Hampton Research) and JCSG-plus, PACT Premier and Nuc-Pro screens (Jena Bioscience). 100 µL of each screening condition was manually pipetted into 96 well Intelli-plate reservoirs (Hampton Research). Drops were then laid using a Mosquito crystal robotic system (TTP Labtech Ltd) in ratios of 1:1, 2:1 and 1:2 of purified protein (Section 2.5.1) to screening solutions. All drops used 150 nL of protein and the volume of screening solution was adjusted to make the appropriate ratio. Once the drops were laid the crystal trays were stored at 4 °C. Crystal trays were inspected manually by microscopy. Crystal trays were inspected from 1 days to up to 3 months after the tray was laid.

2.5.3 Refined crystal screening

Successful screening conditions identified in the high throughput crystal screens were then subjected to refined screens to produce more and better quality crystals. In the refined screens the initial screening conditions were adjusted slightly to optimise crystal formation (Appendix C). Refined screens were carried out using Greiner 24-well plates for hanging drops (Lab Supply). Screening conditions were made up to 500 µL and drops were set up manually using a 1:1 ratio of purified protein (Section 2.5.1) to screening solution. Once laid crystal trays were stored at 4 °C.

2.5.4 Preparation of crystals for data collection

With the assistance of Dr. Chelsea Vickers, single crystals were looped and shipped to the Australian Synchrotron for data collection. The crystals were flash frozen in liquid nitrogen with a cryoprotectant solution made with the appropriate well solution and 30 % glycerol.

2.5.5 X-ray crystallography data collection and processing

X-ray diffraction data was collected remotely using the MX2 beamline at the Australian Synchrotron (McPhillips *et al.,* 2002). Datasets were obtained for both cytoplasmic and chloroplast TPIs of *C. australis*. These datasets were collected at 100 K and a wavelength of

0.9464 Å. The raw data was auto-processed at the Australian Synchrotron returning merged mtz and log files which were used as inputs for molecular replacement. I used the existing structures of *A. thaliana* cytoplasmic and chloroplast TPIs (PDB ID: 4OBT and 4OHQ respectively) to attempt to solve the *C. australis* TPIs by molecular replacement. The template structures were prepared using the CCP4i workbench chainsaw package (Stein, 2008, Win *et al.,* 2011). Molecular replacement was then carried out using the Phenix package (Adams *et al.,* 2011) and the resulting model was manually rebuilt in Coot (Emsley and Cowtan, 2004). Subsequent structural refinement was carried out using phenix.refine of the Phenix package (Adams *et al.,* 2011). The *C. australis* chloroplast TPI structure was successfully solved at 1.90 Å. Although a 1.81 Å data set for the cytoplasmic TPI of *C. australis* was obtained I was unable to solve the structure due to the presence of non-crystallographic symmetry operators in the crystal lattice. However, Emeritus Professor Dr. Geoff Jameson (Massey University) was able to overcome this issue and solve the structure of the *C. australis* cytoplasmic TPI. At the time of writing, the *C. australis* cytoplasmic TPI structure has not yet been finalised.

Chapter 3. Generating a baseline; characterising the *Arabidopsis thaliana* and *Synechocystis* sp. PCC6803 TPIs.

3.1 Background

Arabidopsis thaliana and *Synechocystis* sp. PCC6803 are both model organisms that are widely used to study photosynthesis. Therefore, it is unsurprising that much of the existing knowledge on the TPIs of photosynthetic organisms has come from thorough investigation of their enzymes. Here, the characterisation of the TPIs from these organisms was carried out to generate a baseline for subsequent investigation of the TPIs from non-model organisms and to be able to confidently place the results from this investigation into the context of existing literature. *Synechocystis* has a single TPI (*Spc*TPI) for which both structural and kinetic data are available (Castro-Torres *et al.,* 2018). *A. thaliana* has a cytoplasmic and a chloroplast TPI (c*Ath*TPI and p*Ath*TPI, respectively) in addition to a splice variant of p*Ath*TPI which possesses a nine amino acid deletion (Chen and Thelen, 2010). Both structural and kinetic data for c*Ath*TPI and p*Ath*TPI are available (Dumont *et al.,* 2016, López-Castillo *et al.,* 2016, Castro-Torres *et al.,* 2018). However, the p*Ath*TPI splice variant lacks the catalytic residues N10 and K12 and has been previously shown to be inactive (López-Castillo *et al.,* 2016). Accordingly, the splice variant of p*Ath*TPI was not investigated in this project.

Existing investigations on these TPIs have been in the context of redox sensitivity, demonstrating that p*Ath*TPI is less susceptible to redox agents than c*Ath*TPI while *Spc*TPI is less susceptible than either *Ath*TPIs (Dumont *et al.,* 2016, López-Castillo *et al.,* 2016, Castro-Torres *et al.,* 2018). The reduced redox sensitivity of p*Ath*TPI compared to c*Ath*TPI has been hypothesised as an adaptation of p*Ath*TPI to the reactive oxygen species rich environment of the chloroplast (López-Castillo *et al.,* 2016). The greater redox sensitivity of *Ath*TPIs compared to *Spc*TPI enables greater redox modulation of enzyme activity (Dumont *et al.,* 2016, Castro-Torres *et al.,* 2018). Curiously, despite thorough redox characterisation of these three TPIs there have been no comprehensive studies on their inhibition and regulation by 2PG (Section 1.3.2).

3.2 Expression and purification of c*Ath*TPI, p*Ath*TPI and *Spc*TPI

First, the TPIs from *A. thaliana* and *Synechocystis* were recombinantly expressed in, and then purified from the *E. coli* strain LOBSTR. All three of c*Ath*TPI, p*Ath*TPI and *Spc*TPI were solubly expressed and purified as described in section 2.2. The TPIs began elution from the IMAC column between 50 and 60 % IMAC buffer B, corresponding to between 200 and 240 mM imidazole (Figure 3.1, A, C, E). Slight differences were observed between the TPIs. p*Ath*TPI begins to elute from the HisTrap column slightly earlier than c*Ath*TPI and *Spc*TPI, while the yield of *Spc*TPI is lower and does not saturate the detector as observed for c*Ath*TPI and p*Ath*TPI. Representative fractions from each elution peak were then subject to SDS-PAGE analysis. In this analysis a relatively pure band was observed between 26 and 34 kDa for all three TPIs which is consistent with the predicted molecular weight of 29.2 kDa for both His_6-c*Ath*TPI and His_6-p*Ath*TPI and 31.7 kDa for His_6-*Spc*TPI (Figure 3.1, A, C, E). Each set of TPI-containing fractions was then pooled and concentrated prior to further purification by SEC (Figure 3.1, B, D, F). All three TPIs began elution from the SEC column at approximately 10 mL, although *Spc*TPI began elution slightly later than c*Ath*TPI and p*Ath*TPI. All three of these TPIs are known to be dimers (López-Castillo *et al.,* 2016, Castro-Torres *et al.,* 2018), this enabled for the estimation of the oligomeric state for subsequent TPIs based on when they eluted from the SEC column. SDS-PAGE of the resulting SEC fractions demonstrated that both c*Ath*TPI and p*Ath*TPI were sufficiently pure for further characterisation, however a contaminant was present in *Spc*TPI between 55 and 72 kDa (Figure 3.1, B, D, F). A representative *Spc*TPI sample was rerun and no contaminant was present, indicating contamination during gel loading (Figure 3.1 F).

The *Spc*TPI fractions were then pooled for further analysis. Only the central fraction of the c*Ath*TPI and p*Ath*TPI elution peaks were used for further analysis to minimise the presence of any potential degradation products which are observed in the slight tails of the SEC peaks (Figure 3.1 B, D).

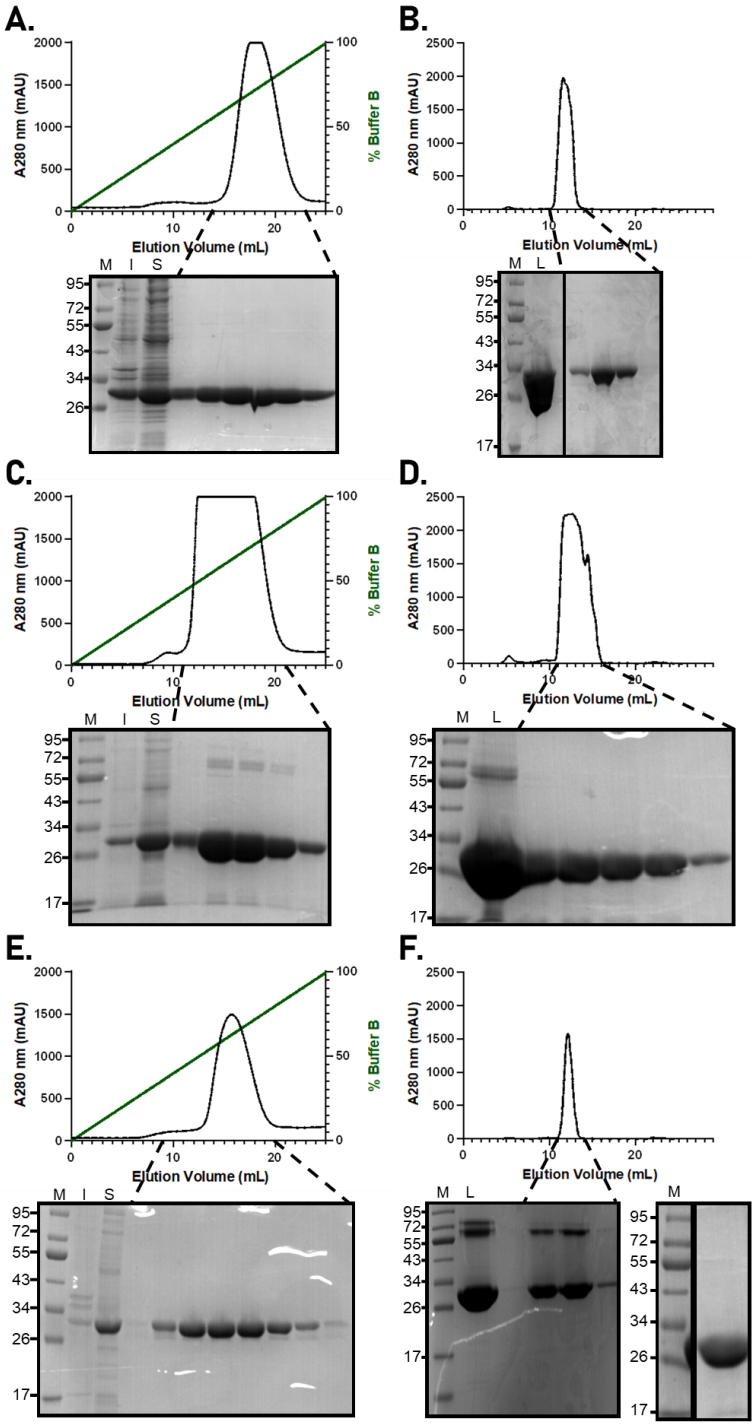


Figure 3.1, Purification of c*Ath*TPI, p*Ath*TPI and *Spc*TPI. A, IMAC trace of c*Ath*TPI. **B,** SEC trace of c*Ath*TPI. **C,** SEC trace of p*Ath*TPI. **D,** SEC trace of p*Ath*TPI. **E,** SEC trace of *Spc*TPI. **F,** SEC trace of *Spc*TPI. SDS-PAGE gels beneath traces show representative fractions from the absorbance peaks indicated. In **F,** the second gel was run to check for contamination. M, molecular weight standards (kDa). I, insoluble fraction, and S, soluble fraction, of whole cell lysis by sonication. L, load fraction; the pooled and concentrated IMAC fractions prior to SEC.

3.3 Steady-state kinetic characterisation of c*Ath*TPI, p*Ath*TPI and *Spc*TPI

3.3.1 Initial kinetic trials using c*Ath*TPI

The c*Ath*TPI was used in initial activity assays to optimise and validate the methodology outlined in section 2.3. Firstly, a storage temperature trial was performed. Aliquots of c*Ath*TPI were stored at room temperature, 4 ℃, -20 ℃ and -80 ℃ and, after one month, were assayed (Figure 3.2 A). c*Ath*TPI remained active at all temperatures after one month. Subsequent purified TPIs were stored at -80 ℃. The second trial was carried out in the presence and absence of the reducing agent DTT (Figure 3.2 B). This trial was carried out based on previous research by Dumont *et al.* (2016) who investigated the redox sensitivity of c*Ath*TPI. c*Ath*TPI was incubated at room temperature for 30 minutes with or without 5 mM DTT and then assayed. After incubation little difference in c*Ath*TPI activity was observed, with the untreated c*Ath*TPI having a slightly greater initial rate than DTT treated (Figure 3.2 B). Based on this result full scale steady-state kinetic analysis was carried out in the absence of DTT.

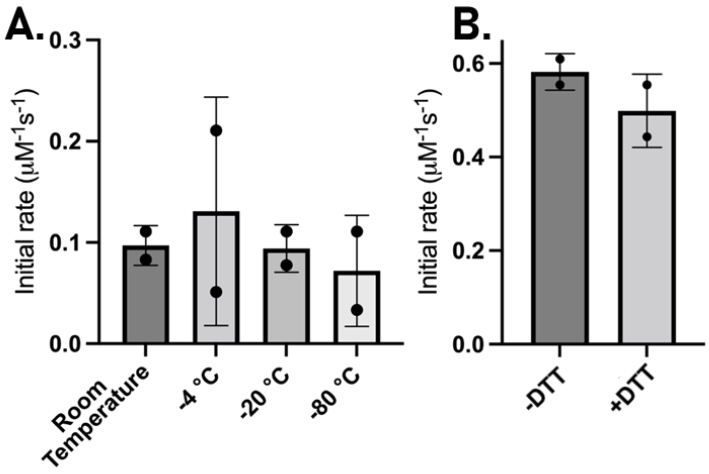


Figure 3.2, Initial kinetic trials of c*Ath*TPI. A. c*Ath*TPI storage temperature trial. Trial was conducted one month after purification and using 0.2 mM GAP. **B.** c*Ath*TPI DTT sensitivity trial. c*Ath*TPI was incubated at room temperature for 30 minutes in the presence or absence of 5 mM DTT prior to assaying. Assay was carried out using 0.5 mM GAP. For both assays, 85 pM c*Ath*TPI was used and each measurement was performed in duplicate and normalised to 0 mM GAP. Bars represent mean ± SEM with dots representing individual measurements.

3.3.2 Enzymatic determination of GAP concentration

The substrate for kinetics assays, GAP, is sold as a 8-13 mg/mL solution which corresponds to 47-76 mM. Initially it was assumed that the GAP concentration was 62 mM, the midpoint of the reported range. In order to make accurate estimates of kinetic parameters it was necessary to determine the GAP concentration more precisely. In the coupled enzyme assay NADH and GAP are consumed in a one to one ratio. Therefore, by adjusting the concentrations of NADH and GAP so that GAP was limiting it was possible to experimentally determine the GAP concentration by measuring the change in NADH absorbance once TPI had converted all available GAP to DHAP. This was determined for three bottles of GAP to ensure standardisation across experiments. The GAP concentration was determined to be 50 ± 1 mM (mean ± SEM), with all bottles measured falling within this margin of error. This concentration was used when setting up all subsequent assays. For assays which were completed prior to the determination of GAP concentration the concentrations of GAP were retroactively adjusted and the kinetic parameters recalculated. This was possible as all bottles of GAP used during the course of this project were from the same batch.

3.3.3 Steady-state kinetic characterisation of *Spc*TPI, c*Ath*TPI and p*Ath*TPI

Michaelis-Menten steady-state kinetic parameters were determined for each TPI using a coupled enzyme assay described in section 2.3.3. In this assay the initial velocity of the reaction was determined for a range of different GAP concentrations and using direct non-linear fitting, these rate data were then fitted to the Michaelis-Menten equation to determine three kinetic parameters; k_{cat}, K_M and k_{cat}/K_M (Figure 3.3). The k_{cat}, or turnover number, is the maximum number of substrate molecules that can be converted to product per active site per second. The K_M is the concentration of substrate at which an enzyme achieves half its maximum velocity. The ratio of the two, k_{cat}/K_M, is a measure of overall catalytic efficiency. TPIs are known to obey Michaelis-Menten kinetics (Putman *et al.,* 1972).

The values for k_{cat} and K_M were similar for *Spc*TPI, c*Ath*TPI and p*Ath*TPI with the k_{cat} ranging from 12000 to 17000 s^{-1} and the K_M ranging from 0.61 to 0.85 mM (Table 3.1). The narrow range for k_{cat} and K_M is reflected in the near identical k_{cat}/K_M of approximately 2.0×10^7 $s^{-1}M^{-1}$ for all three enzymes (Table 3.1). The observation of near identical kinetic parameters across these three TPI is consistent with previously reported data. However, the previously reported k_{cat}/K_M are

between four and eight times lower than those determined in this project (Table 3.1). The difference in k_{cat}/K_M is primarily due to differences in k_{cat}. The K_M values observed in this project fall within the range of those previously reported (0.40 to 1.3 mM). However, I observed k_{cat} values that are between five and 11 times greater than those previously reported. This trend is most evident for *Spc*TPI where the values of K_M are identical, but an approximately eightfold difference in k_{cat} and consequently k_{cat}/K_M is reported (Table 3.1).

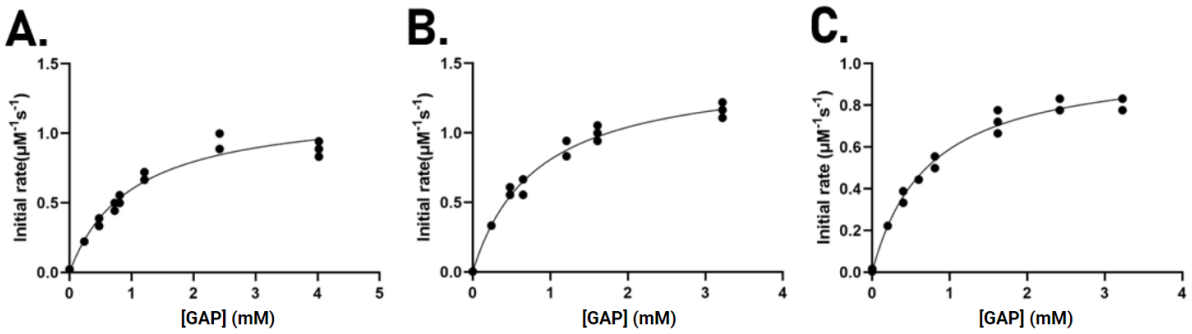


Figure 3.3, Michaelis-Menten kinetics. A. *Spc*TPI. **B.** c*Ath*TPI. **C.** p*Ath*TPI. Measurements for each GAP concentration were performed in quadruplicate. Assays were repeated in at least triplicate to produce the data reported in table 3.1.

Table 3.1, Steady-state kinetic parameters for *Spc*TPI, c*Ath*TPI and p*Ath*TPI.

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}M^{-1}$)	Reference
*Spc*TPI	15000 ± 2000	0.8 ± 0.1	2.0 ± 0.3 ×10^7	This work
	1800 ± 40	0.80 ± 0.02	2.3 ×10^6	Castro-Torres *et al.,* 2018
c*Ath*TPI	17000 ± 2000	0.9 ± 0.1	2.1 ± 0.1 ×10^7	This work
	1100 ± 50	0.47 ± 0.08	2.4 ×10^6	Castro-Torres *et al.,* 2019
	2600 ± 40	0.48 ± 0.05	5.3 ×10^6	López-Castillo *et al.,* 2016
	3400 ± 300	1.3 ± 0.06	2.6 ± 0.3 ×10^6	Dumont *et al.,* 2016
p*Ath*TPI	12000 ± 500	0.61 ± 0.06	2.0 ± 0.2 ×10^7	This work
	2300 ± 40	0.40 ± 0.02	5.8 ×10^6	López-Castillo *et al.,* 2016

Data is reported as mean ± SEM where available. In my experiments n = 3 (p*Ath*TPI) or 4 (c*Ath*TPI and *Spc*TPI) independent data sets, with each independent data set comprised technical quadruplicates for each measured substrate concentration. Data for *Spc*TPI, p*Ath*TPI, and c*Ath*TPI were generated using TPI from a single purification of each protein.

3.3.4 Developing a 2-phosphoglycolate inhibition assay

2PG is a known inhibitor of TPI and is a product of the rubisco catalysed oxygenation reaction. To assess TPI inhibition by 2PG the coupled enzyme assay was modified by the addition of between 20 and 100 µM 2PG. Two 2PG concentrations were measured simultaneously against a negative 2PG control and the resulting initial rates were used to calculate the inhibition constant, K_i. The K_i is a measure of inhibitor potency and is the concentration of inhibitor required to produce half the maximal inhibition. Assuming the competitive model of 2PG inhibition, the K_i for c*Ath*TPI and p*Ath*TPI was 20 ± 10 and 24 ± 3 µM respectively, which is fairly consistent with previously reported data (Table 3.2). However, the competitive model of inhibition poorly fits the data reported in this project, evidenced by high concentrations of GAP not outcompeting 2PG and thus the initial rate is not restored to that of the negative 2PG control (Figure 3.4 A). Although 2PG is not a reported inhibitor of the coupled enzyme GPDH, I ran a trial assay where the GPDH concentration was increased to five times that in the standard assay (Section 3.3.3), to rule out 2PG interference with the coupled enzyme (Appendix D). Increasing the GPDH concentration had no effect on the behaviour of the assay system, therefore 2PG did not appear to be interfering downstream of TPI in this assay. The datasets were then reevaluated using the mixed model of inhibition (Figure 3.4 B). When using a mixed model for 2PG inhibition a greater K_i was observed for both c*Ath*TPI and p*Ath*TPI at 95 ± 50 and 70 ± 7 µM respectively (Table 3.2). The extra sums-of-square F test was used to identify which model was a better fit for the available data. This test reports an F statistic and an associated p-value. If the F statistic is greater than one, then the more complicated model, the mixed model, is correct, if it is less than one then the simpler model, the competitive model, is correct. The p-value identifies whether this difference is significant. For every replicate the F statistic was greater than one and the associated p-values were less than 0.05 therefore, identifying that the mixed model was a significantly better fit of the available data than the competitive model.

Table 3.2, 2PG inhibition of c*Ath*TPI and p*Ath*TPI.

	Competitive K_i (μM)	Mixed K_i (μM)	Reference
c*Ath*TPI	20 ± 10	100 ± 50	This work
p*Ath*TPI	24 ± 3	70 ± 7	This work
	36 ± 4		Flügel *et al.,* 2017
*Ath*TPI*	24		Li *et al.,* 2019
c*Psa*TPI	4 ± 0.2		Anderson, 1971
p*Psa*TPI	15 ± 0.9		Anderson, 1971

Data is reported as mean ± SEM where available. In my experiments n = 2 (c*Ath*TPI) or 3 (p*Ath*TPI) independent data sets, with each independent data set consisting of technical triplicates for each measured combination of 2PG and substrate concentration. For each TPI, all assays were performed using protein from a single purification.
*Pooled c*Ath*TPI and p*Ath*TPI isolated from whole leaf extract.

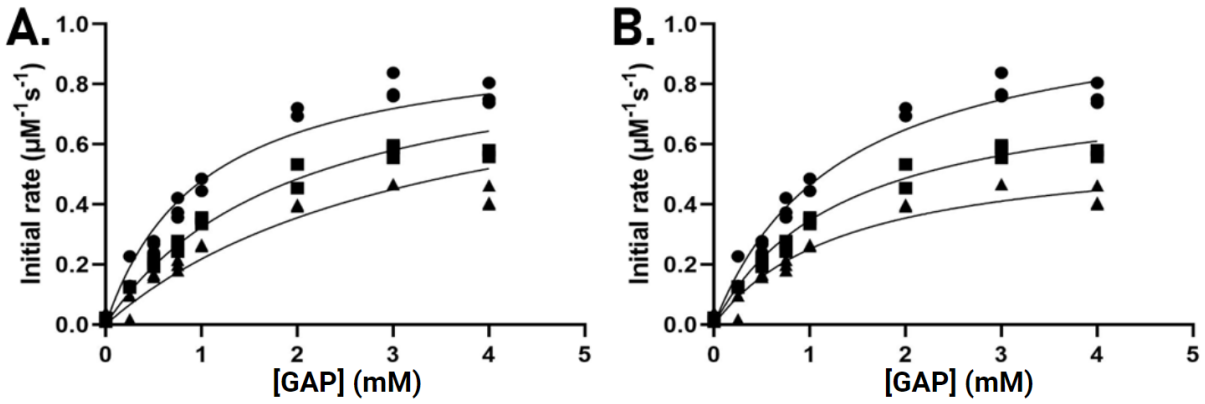


Figure 3.4, Comparison of competitive and mixed models of 2PG inhibition of p*Ath*TPI. A. Competitive inhibition model. **B.** Mixed inhibition model. The rate data are the same in A and B; only the model chosen to fit the data is changed. ●, 0 μM 2PG; ■, 20 μM 2PG; ▲, 50 μM 2PG. Each measurement was performed in triplicate. To identify the better fitting model an extra sum-of-square F test was carried out. The resulting F statistic was 44.6 and the associated p-value was <0.0001, indicating that the mixed model is a significantly better fit for the available data.

3.4 Discussion

In this chapter I set out to characterise *Spc*TPI, c*Ath*TPI and p*Ath*TPI to optimise workflow and place this research in the context of existing literature. The protein expression and purification workflows produced consistently high yields of TPI for use in subsequent kinetic analysis. The steady-state kinetic analysis of all three TPIs reported catalytic efficiencies that are 5 to 11 times greater than has been previously reported primarily due to differences in k_{cat} (Table 3.1). The differences in k_{cat} can be accounted for by differences in methodology, such as the methods of protein purification, protein concentration determination, and the pH of the assay buffer. These differences have a pronounced effect on the final kinetic parameters reported due to the nM-pM concentrations of TPI used in kinetic assays.

3.4.1 On 2-phosphoglycolate inhibition of TPIs from photosynthetic organisms

This project is the first to report a K_i for 2PG inhibition of c*Ath*TPI. Consequently, no previous works have directly compared the K_i for 2PG inhibition of c*Ath*TPI and p*Ath*TPI. The seminal paper investigating differential 2PG inhibition between cytoplasmic and chloroplast TPIs, used TPIs isolated from the pea plant *Pisum sativum* (Anderson *et al.,* 1971, Table 3.2). Anderson and colleagues reported that the cytoplasmic *P. sativum* TPI (c*Psa*TPI) is approximately three times more sensitive to 2PG than the chloroplast isoform (p*Psa*TPI). This project does not observe the same pattern of inhibition using either the competitive or mixed model of inhibition (Table 3.2), albeit for the *A. thaliana* TPIs rather than those from *P. sativum*. This is unsurprising, given that c*Ath*TPI and p*Ath*TPI share ~60 % sequence identity, including complete conservation of their active sites, and their catalytic efficiencies are indistinguishable (Table 3.1).

Previous estimations of the 2PG K_i for c*Psa*TPI, p*Psa*TPI and pooled *Ath*TPI used Lineweaver-Burk plots or a secondary plot as was previously done for p*Ath*TPI (Anderson *et al.,* 1971, Li *et al.,* 2019, Flügel *et al.,* 2017). It is well established that Lineweaver-Burk analysis biases kinetic analysis towards data collected at low substrate concentrations, which have a higher proportion of error, resulting in inaccurate estimates of kinetic parameters (Johnson, 2013). Secondary plots introduce cumulative error due to their derivation from multiple Lineweaver-Burk plots and also biases the resulting K_i towards data collected at higher inhibitor

concentrations (Johnson, 2013). Because direct, nonlinear fitting was used here, I believe that the K_i data generated in this project to be the most accurate estimation of this parameter yet for *Ath*TPIs.

3.4.2 On 2-phosphoglycolate inhibition of TPIs from diverse organisms

At the end of 2021 there were eight values for the 2PG K_i for wild type TPIs in BRENDA, four of which are for the TPI of the human parasite *Trypanosoma brucei*. Notably, the K_i values for c*Psa*TPI, p*Psa*TPI and pooled *Ath*TPI were not reported in BRENDA. Why these data have been excluded from BRENDA is unclear. In BRENDA, all entries assumed a competitive model of 2PG inhibition and the reported range of K_i values was from 19 µM for the TPI of *Saccharomyces cerevisiae* (Go *et al.,* 2010) to 89 mM for the TPI of the psychrophilic bacterium *Morinella marina*, formerly called *Vibrio marinus* (Alvarez *et al.,* 1998). Curiously, the two entries for prokaryotic TPIs were the only entries exceeding 55 µM, the entry for *M. marina* and a K_i of 6 mM for the TPI of *E. coli* (Alvarez *et al.,* 1998). Determining the K_i of *Spc*TPI would be interesting to see whether this trend of prokaryote TPIs being less sensitive to 2PG persists. However, due to time restrictions this could not be carried out during this project.
Assuming the competitive model of 2PG inhibition the estimates of K_i reported in this project fall within the range of those previously reported for other eukaryote TPIs. However, to my knowledge there have been no previously reported 2PG K_i values which assume a mixed model of inhibition.

3.4.3 On the competitive and mixed models of 2PG inhibition

As 2PG is a transition state analogue that occupies the TPI active site, previous authors have used a simple competitive model of inhibition (Lolis and Petsko, 1990). However, in this project I have demonstrated that, under the conditions of my assay system, a mixed model of inhibition is a significantly better fit of the available data (Figure 3.3). Mixed or noncompetitive activity of inhibitors which act at an enzyme active site is not unprecedented (Kovalenko *et al.,* 1997, Walker *et al.,* 2001, Liu *et al.,* 2008) with four different mechanisms of such inhibition being described (Blat, 2010). I believe the mixed inhibition of TPI by 2PG most likely follows a two-step induced fit model (Figure 3.5). Under this model the association-dissociation kinetics (k_1, k_{-1}) are rapid and strictly competitive forming an enzyme inhibitor complex (EI). However,

inhibitor binding also induces a conformational change in the enzyme (k_2) that stabilises the enzyme-inhibitor complex (EI*). The inhibitor cannot dissociate from EI* until the EI* complex reverts back to EI (k_{-2}). The smaller magnitude of k_{-2} relative to substrate binding and dissociation is responsible for the observed mixed activity of inhibitors that occupy enzyme active sites.

Consistent with being a transition state analogue, 2PG is known to induce conformational changes in loop 6 stabilising the closed confirmation of TPI (Castro-Torres *et al.,* 2018). It follows that 2PG cannot disassociate from TPI while the catalytic loop 6 is closed and that as 2PG stabilises the closed conformation, k_{-2} will be less than k_2, which is consistent with the second requirement of a two-step induced fit model of inhibition. What is not consistent with a two-step induced fit model of 2PG inhibition, and why I think this model has been overlooked, is the status of TPI as a 'perfect' and therefore diffusion limited enzyme. By definition, the association-dissociation kinetics limit the catalysis of a 'perfect' enzyme, which is incongruous with a model of 2PG inhibition which is limited by the confirmation changes in loop 6. However, this notion of catalytic 'perfection' is becoming increasingly problematic. The theoretical k_{cat}/K_M of a diffusion limited enzyme is ~ 10^8-10^9 $s^{-1}M^{-1}$, however, most enzymes, including the so-called 'perfect' TPI (Table 3.1), do not possess such exquisite efficiency (Bar-Even *et al.,* 2011). It has previously been reported that TPI is at most partially diffusion limited, with loop motion also being partially rate limiting (Liao *et al.,* 2018). Considering the re-evaluation of TPIs so called 'perfection' it is possible that the association-dissociation kinetics may not, at least in part, be limiting TPI activity. Thus, it appears feasible that 2PG inhibition may act through a mixed model such as the two-step inhibitor binding mechanism where the opening of loop 6 (k_{-2}) is limiting (Figure 3.5).

This figure is unavailable

Figure 3.5, Two-step inhibitor binding mechanism. Inhibitor binding follows two steps with the conversion of the EI complex to the more stable EI*. Stability of the EI* complex is due to the slow rate constant, k_{-2}. Figure taken from Blat, 2010.

3.5 Conclusion

Having established and validated a pipeline for the purification and characterisation of the TPIs from photosynthetic organisms, the following chapters describe implementing this pipeline to investigate the new-to-science TPIs from *C. australis* and *P. umbilicalis*. In particular, subsequent 2PG inhibition assays will be analysed using a mixed model of 2PG inhibition based on the data presented for *Ath*TPIs (Section 3.3.4) and the argument that this conforms to a two-step inhibitor binding model of inhibition (Section 3.4.3).

Chapter 4. Is relaxing purifying selection acting on the chloroplast TPI of a non-photosynthetic parasitic plant?

4.1 Background and hypothesis

Cuscuta australis is an obligate stem parasite belonging to a family commonly known as dodders. There are over 200 known species of dodder, many of which parasitise crop species such as tomato, tobacco and forage legumes (Dawson *et al.,* 1994). Dodders lack roots and leaves and contain only trace amounts of chlorophyll; consequently they cannot sustain themselves by photosynthesis (Yoshida *et al.,* 2016). The *Cuscuta* genus diverged from *A. thaliana* some 117 million years ago, however, adaptations to parasitism have occurred over the past 33 million years since the *Cuscuta* family diverged from their closest nonparasitic relatives, the morning glories (Sun *et al.,* 2018). Since their divergence from the morning glories, *Cuscuta* have undergone extensive relaxed purifying selection (Sun *et al.,* 2018). Purifying selection is natural selection against mutations that are deleterious to overall survival of an organism, while relaxation is the easing of selective pressure which manifests as reduced efficiency or intensity of natural selection (Wertheim *et al.,* 2014). Consequently, relaxed purifying selection is manifested as the accumulation of mutations in regions of the genome that are under reduced selective constraint. In *Cuscuta* relaxed purifying selection is evidenced by large scale gene loss, including genes involved in photosynthesis as well as root and flower development (Sun *et al.,* 2018). Additionally, Sun *et al.,* (2018) identified approximately 3900 genes in *C. australis* undergoing relaxed purifying selection, including the gene of a predicted chloroplast triosephosphate isomerase. This makes sense, as little to no photosynthesis occurs in the *C. australis* chloroplast the chloroplast TPI is presumably rendered redundant. I hypothesised that 33 million years of relaxed purifying selection would be reflected in the structure and function of the chloroplast TPI. In particular, I hypothesised that the mutations would have eroded its catalytic efficiency. To test this hypothesis, I have determined the steady-state kinetic parameters of the cytoplasmic and chloroplast TPIs from *C. australis* to compare with those from *A. thaliana* (Section 3.3.3). Additionally, the structures of the cytoplasmic and chloroplast TPIs from *C. australis* were solved using X-ray crystallography.

4.2 Mining the genome of *Cuscuta australis*

The genome of *C. australis* was released by Sun *et al.* (2018). Using Blastp (Altschul *et al.,* 1990) and the amino acid sequences of c*Ath*TPI and p*Ath*TPI three TPIs were identified encoded in the genome of *C. australis* (Appendix B). I predicted that these genes encoded two cytoplasmic and one chloroplast TPIs based on MSA against the *Ath*TPIs (Figure 4.1 A). These TPIs were dubbed *Cau*TPI, c*Cau*TPI and p*Cau*TPI respectively. p*Cau*TPI was previously identified by Sun *et al.* (2018) as undergoing relaxed purifying selection. The chloroplast localisation of p*Cau*TPI was confirmed using TargetP 2.0 (Armenteros *et al.,* 2019) to identify the chloroplast localisation tag and its cleavage site. The residues upstream of the cleavage site were excluded from the p*Cau*TPI expression construct (Appendix A). c*Cau*TPI and p*Cau*TPI possess ~56 % sequence identity with one another and ~81 % sequence identity with their corresponding *Ath*TPIs. The most divergent TPI was the cytoplasmic isoform *Cau*TPI, with ~70 % identity compared to the other cytoplasmic isoforms and ~53 % identity with both chloroplast isoforms from both *C. australis* and *A. thaliana* (Figure 4.1 B). The sequence divergence of *Cau*TPI is primarily due to a C-terminal extension.

A.

```
                β1                  α1                      β2      α2
pAth MAGSGKFFVGGNWKCNGTKDSIAKLISDLNSATLE--ADVDVVVSPPFVYID  50
pCau MAGSGKCFVGGNWKCNGTKESIVRLISDLNSSKLE--PDVDVVVAPPFLYIE  50
Cau  --MGRKFFVGGNWKCNGTSEEVKKIVSTLNAGEVPPQDVVEVVVSPPFVFLP  50
cAth  --MARKFFVGGNWKCNGTAEEVKKIVNTLNEAQVPSQDVVEVVVSPPYVFLP  50
cCau  --MARKFFVGGNWKCNGTAEEVKKIVTTLNEAEVPSEDDVEVVVSPPFVFLT  50
                         ●  ●

      α2        β3                        α3      β4      α4
pAth QVKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSERRHV 102
pCau QVKSTLTDRIEIAAQNCWIGKGGAFTGEISAEQLKDIGCKWVILGHSERRHV 102
Cau  LAKSSLRPDFLVAAQNCWIKKGGAFTGEISAEMLVNLDIPWVILGHSERRLI 102
cAth LVKSTLRSDFFVAAQNCWVKKGGAFTGEVSAEMLVNLDIPWVILGHSERRAI 102
cCau LVKTLLRPDFSIAAQNCWVRKGGAFTGEVSAEMLINLGIPWVILGHSERRQL 102
                                   ■                   ●

                α4          β5          α5
pAth IGEKDEFIGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADA 154
pCau MGENNEFIGKKAAYASSQGVGIIACIGELLEEREARKTFDVCFQQLKAFADA 154
Cau  LGESNEFVGDKVAYALSQGLKVIACVGETLEQREAGTTMDVVAAQTKAIADK 154
cAth LNESSEFVGDKVAYALAQGLKVIACVGETLEEREAGSTMDVVAAQTKAIADR 154
cCau LNESDDFVGDKVAYALIQGLKVIACVGETLEQRESGATMKVVAAQTEAIFDK 154

                β6                  α6
pAth VPSWDNIVVAYEPVWAIGTGKVASPQQAQEVHVAVRGWLKKNVSEEVASKTR 206
pCau LPSWENVVIAYEPVWAIGTGKVATPEQAQEVHAAIRDWLNKNVSSEVASETR 206
Cau  VSDWSKVVIAYEPVWAIGTGKVATPAQAQEVHAELRKWLESNVSSEVASSTR 206
cAth VTNWSNVVIAYEPVWAIGTGKVASPAQAQEVHDELRKWLAKNVSADVAATTR 206
cCau VSNWSNIVLAYEPVWAIGTGKVATPAQAQEVHSELRKWLHDNVGPEVAASTR 206
                    ●

     β7         α7         β8          α8
pAth IIYGGSVNGGNSAELAKEEDIDGFLVGGASLKGPEFATIVNSVTSKKVAA-- 256
pCau IIYGGSVNGSNCSELAKKEDIDGFLVGGASLKGPDFASIVNSVASKKVTA-- 256
Cau  IIYGGKYFVTILHIT-----SNSYLRSSGLLDTLYMCESIPIFLCFDKIQSD 253
cAth IIYGGSVNGGNCKELGGQADVDGFLVGGASLKPEFI----DIIKAAEVKKSA 254
cCau IIYGGSVSGANCKELATKPDVDGFLVGGASLKPEFI----EIIKSATLKQNA 254

pAth ----------------------
pCau ----------------------
Cau  YVYMGLAHGIGLLAQVAGYM 273
cAth ----------------------
cCau ----------------------
```

B.

	pAth	pCau	Cau	cAth	cCau
pAth	100.0				
pCau	81.3	100.0			
Cau	53.4	53.8	100.0		
cAth	60.4	58.4	73.5	100.0	
cCau	58.8	56.4	70.3	80.7	100.0

Figure 4.1, Multiple sequence alignment of *A. thaliana* and *C. australis* TPIs. A. MSA using Clustal Omega (Sievers *et al.,* 2011). Secondary structural elements are based on the solved structure of p*Ath*TPI (PDB ID: 4OHQ) with dashed lines indicating breaks in secondary structural elements. ● indicates catalytic residues while the ■ indicates key residue of the dimer interface. **B.** Percent identity matrix generated by Clustal Omega. MSA and identity matrix were generated without the chloroplast localisation tags of p*Ath*TPI and p*Cau*TPI.

4.3 Expression screen of *Cau*TPI

The first TPI chosen for analysis was *Cau*TPI as it was the most divergent TPI identified in *C. australis*. However, *Cau*TPI was not soluble under the conditions of the previously established workflow (Section 3.2). Consequently, an expression screen was carried out to try to improve *Cau*TPI solubility. No His_6-*Cau*TPI (31.7 kDa) was soluble at 2, 4 or ~16 hours post induction with IPTG, although a band corresponding to His_6-*Cau*TPI did accumulate in the insoluble fraction (Figure 4.2). As my interests were primarily in the novelty of the *C. australis* chloroplast TPI and no chloroplast localisation tag was identified in the *Cau*TPI sequence this protein was not pursued further.

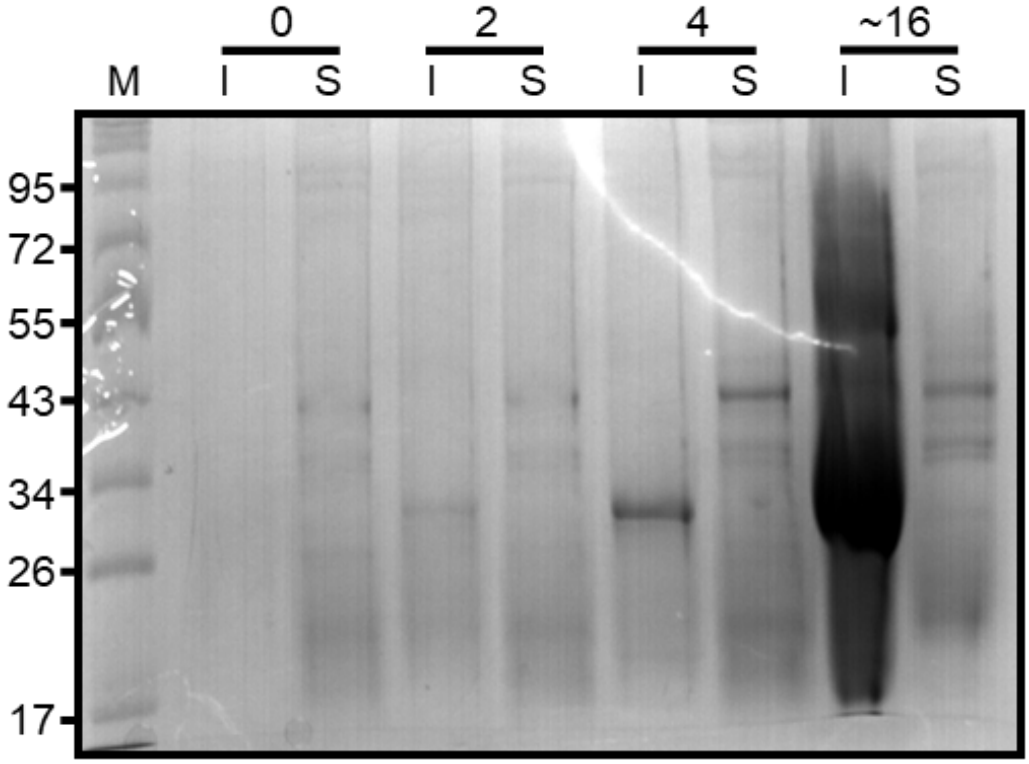


Figure 4.2, Expression screen of *Cau*TPI. Samples of *Cau*TPI were taken at 0, 2, 4 and ~16 hours post induction with IPTG. M, molecular weight standard in kDa. I, insoluble fraction and S, soluble fraction, of whole cell lysis by sonication.

4.4 Expression and purification of c*Cau*TPI and p*Cau*TPI

Both c*Cau*TPI and p*Cau*TPI were solubly expressed and purified with both TPIs behaving as anticipated using this workflow (Figure 4.3). Both c*Cau*TPI and p*Cau*TPI began IMAC elution between 50 and 60 % IMAC buffer B, corresponding to between 200 and 240 mM imidazole. The IMAC purification of c*Cau*TPI saturated the detector which was not observed when purifying p*Cau*TPI. The apparent lower yield of p*Cau*TPI was due to time restrictions that necessitated the flow rate be increased from 1 mL/min to 5 mL/min. Consequently the p*Cau*TPI sample was exposed to the detector for less time and so the yield appears lower. However, SDS-PAGE analysis of representative IMAC fractions showed a comparable yield for both c*Cau*TPI and p*Cau*TPI preparations. The bands present between 26 and 34 kDa, which is consistent with the predicted molecular weight of 29.5 kDa for both His_6-c*Cau*TPI and His_6-p*Cau*TPI, are relatively pure (Figure 4.3 A,C). The TPI containing fractions were then pooled and concentrated prior to further purification by SEC. Both TPIs eluted from the size exclusion column beginning at 10 mL, which is the same elution volume as the corresponding *Ath*TPIs (Section 3.2), indicating that both c*Cau*TPI and p*Cau*TPI are also likely dimers. SDS-PAGE of the resulting fractions demonstrated that both TPIs were sufficiently pure for further characterisation. Only the central fraction of each elution peak was used for further analysis to minimise the presence of any possible degradation products which are observed in the slight tail of the c*Cau*TPI SEC peak (Figure 4.3 B, D).

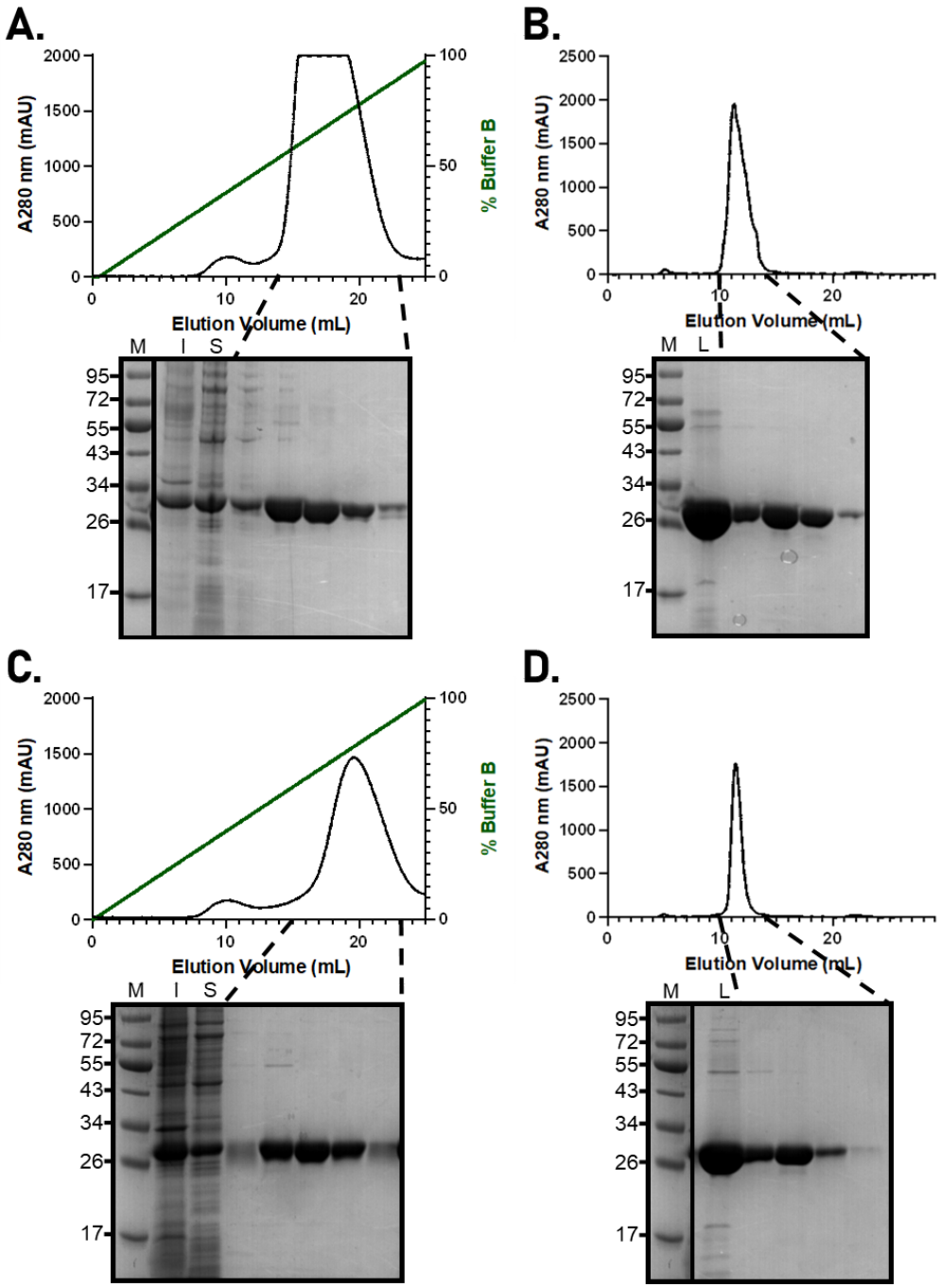


Figure 4.3, Typical purifications of c*Cau*TPI and p*Cau*TPI. A, IMAC trace of c*Cau*TPI. **B,** SEC trace of c*Cau*TPI. **C,** IMAC trace of p*Cau*TPI. **D,** SEC trace of p*Cau*TPI. SDS-PAGE gels beneath traces show representative fractions from the absorbance peaks indicated. M, molecular weight standards (kDa). I, insoluble fraction, and S, soluble fraction, of whole cell lysis by sonication. L, load fraction; the pooled and concentrated IMAC fractions prior to SEC.

4.5 Steady-state kinetic analysis of c*Cau*TPI and p*Cau*TPI

Both c*Cau*TPI and p*Cau*TPI were analysed using the enzyme coupled assay which was established using *Spc*TPI and *Ath*TPIs (Section 3.3.3). The results from steady-state kinetic analysis of the *Cau*TPIs were compared to each other and to the equivalent *Ath*TPIs. No significant difference was observed between c*Cau*TPI and p*Cau*TPI across any kinetic parameter measured (Table 4.1). However, the *Cau*TPIs had approximately half the catalytic efficiency, k_{cat}/K_M, compared to *Ath*TPIs. The values of K_M observed did not differ significantly between the *Cau*TPIs and *Ath*TPIs. However, both c*Cau*TPI and p*Cau*TPI had approximately half the k_{cat}, and consequently k_{cat}/K_M, than the equivalent *Ath*TPIs (Table 4.1).

The loss of catalytic efficiency between p*Cau*TPI and p*Ath*TPI is consistent with my hypothesis that relaxing purifying selection, due to the redundancy of p*Cau*TPI, has eroded the catalytic efficiency of p*Cau*TPI. However, the comparable loss of c*Cau*TPI catalytic efficiency was unanticipated. Additionally, although loss of catalytic efficiency is observed, both c*Cau*TPI and p*Cau*TPI do remain highly efficient enzymes with k_{cat} values that exceed 6000 s^{-1} and an overall efficiency in the order of $\times 10^7$ $s^{-1}M^{-1}$ (Table 4.1).

Table 4.1, Steady-state kinetic parameters c*Cau*TPI and p*Cau*TPI.

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}M^{-1}$)
c*Ath*TPI	17000 ± 2000	0.9 ± 0.1	2.1 ± 0.1 $\times 10^7$
c*Cau*TPI	6900 ± 400	0.58 ± 0.06	1.2 ± 0.1 $\times 10^7$
p*Ath*TPI	12000 ± 500	0.61 ± 0.06	2.0 ± 0.2 $\times 10^7$
p*Cau*TPI	6300 ± 900	0.7 ± 0.2	1.0 ± 0.2 $\times 10^7$

Data is reported as mean ± SEM. *n* = 3 (p*Ath*TPI), 4 (c*Ath*TPI and c*Cau*TPI), or 5 (p*Cau*TPI) independent data sets, with each independent data set comprising technical quadruplicates for each measured substrate concentration. Data for c*Cau*TPI, c*Ath*TPI and p*Ath*TPI were generated using TPI from a single purification. Data for p*Cau*TPI was generated using TPI from two independent purifications.

4.6 X-ray crystallography of c*Cau*TPI and p*Cau*TPI

To attempt to determine a structural basis for the observed difference in catalytic efficiency between *Cau*TPIs and *Ath*TPIs (Table 4.1), I attempted to determine the structures of c*Cau*TPI and p*Cau*TPI by X-ray crystallography (Section 2.5). Crystals were successfully generated for both enzymes. A large coffin shaped crystal of c *Cau*TPI was generated in the PactPremier screen condition B7 (0.2 M sodium chloride 0.1 M MES pH 6.0, 20 % w/v PEG 6000), however, this condition also contained an unknown microbial contaminant. Similarly, a large coffin shaped crystal of p*Cau*TPI was generated in the IndexHT screen condition G8 (0.2 M ammonium acetate, 0.1 M HEPES pH 7.5, 25 % w/v PEG 3350). Both crystals formed within approximately 3 weeks of setting up the crystal trays. Dr. Chelsea Vickers successfully looped these crystals, and many more, which were then shipped to the Australian Synchrotron for data collection. The c*Cau*TPI crystal diffracted to a resolution of 1.81 Å, while the p*Cau*TPI crystal diffracted to a resolution of 1.90 Å.

These datasets were then used to solve the structures of c*Cau*TPI and p*Cau*TPI by molecular replacement. The p*Cau*TPI structure was successfully solved at 1.90 Å (Appendix E). The c*Cau*TPI proved more difficult to solve due to the presence of non-crystallographic symmetry operators in the crystal lattice. However, Emeritus Professor Geoff Jameson (Massey University) was able to overcome this issue and solve the structure of c*Cau*TPI. At the time of writing the c*Cau*TPI structure has not been finalised, consequently structural analysis will focus on p*Cau*TPI.

4.6.1 The conservation of the p*Cau*TPI active site

The p*Cau*TPI structure was solved with four chains in the asymmetric unit (Appendix E) which formed two dimers (Figure 4.4), confirming the oligomeric state inferred from SEC (Section 4.4). Each p*Cau*TPI monomer is composed of the common TIM-barrel fold (Banner *et al.,* 1975) of eightfold alternating of α-helices and β-sheets (Figure 4.4).

Three of the chains in the p*Cau*TPI asymmetric unit were solved with the catalytic loop, loop 6, in the open conformation. However, due to the flexibility of the open loop portions of each loop could not be resolved (Figure 4.4). The final, and highest quality, p*Cau*TPI chain (chain A) was solved with loop 6 in the closed conformation, as evidenced by the superimposition of p*Cau*TPI with the p*Ath*TPI and *Spc*TPI structures (Figure 4.4). The acetate ion in the active site of this

chain mimics the phosphate moiety of the substrate/2PG, forming a hydrogen bond with the amide backbone of G172 and promoting the closed conformation of loop 6 (Figure 4.4). In chain A, the side chain of the catalytic glutamic acid (E166), which is at the N-terminus of loop 6, could not be resolved (Figure 4.4), indicating that E166 sampled multiple conformations in the crystal lattice. The heterogeneity in E166 conformation despite the closure of loop 6 is consistent with previous studies which have identified a degree of independence in the motions of the N- and C- terminal regions of loop 6 (Rozovsky and McDermott, 2001) and with molecular dynamics investigations which have identified multiple subpopulations of open and closed loop 6 conformations (Liao *et al.,* 2018). The orientation of the other catalytic residues, N10, K12 and H96, and the key residue at the dimer interface, T76, is highly conserved between the p*Cau*TPI, *Spc*TPI and p*Ath*TPI structures (Figure 4.4). Other key features such as the sequence identity of loop 6, and the HSERR motif in α-helix four and YGGS motif in loop 7, which are both involved in substrate binding, are fully conserved between the *Cau*TPIs and *Ath*TPIs (Figure 4.1). In light of the maintenance of these known catalytic features it is unsurprising that p*Cau*TPI retains high catalytic efficiency, albeit approximately half that of p*Ath*TPI (Table 4.1). In light of this analysis, it appears that the difference in *Cau*TPI and *Ath*TPI catalytic efficiency is caused by differences that lie outside of the active site, which is unsurprising considering the scale of TPIs catalytic activity. As the TPIs are so exquisitely efficient, with individual reactions occurring at nano- to pico- second time scales, then it follows that even miniscule changes in the positioning of key residues may have a dramatic effect on TPI activity. Thus it is unsurprising that no clear structural basis for the difference in observed *Cau*TPI and *Ath*TPI activity was found and that differences that lie outside of the active site may have a cryptic effect on activity which is not apparent upon manual inspection of the p*Cau*TPI structure.

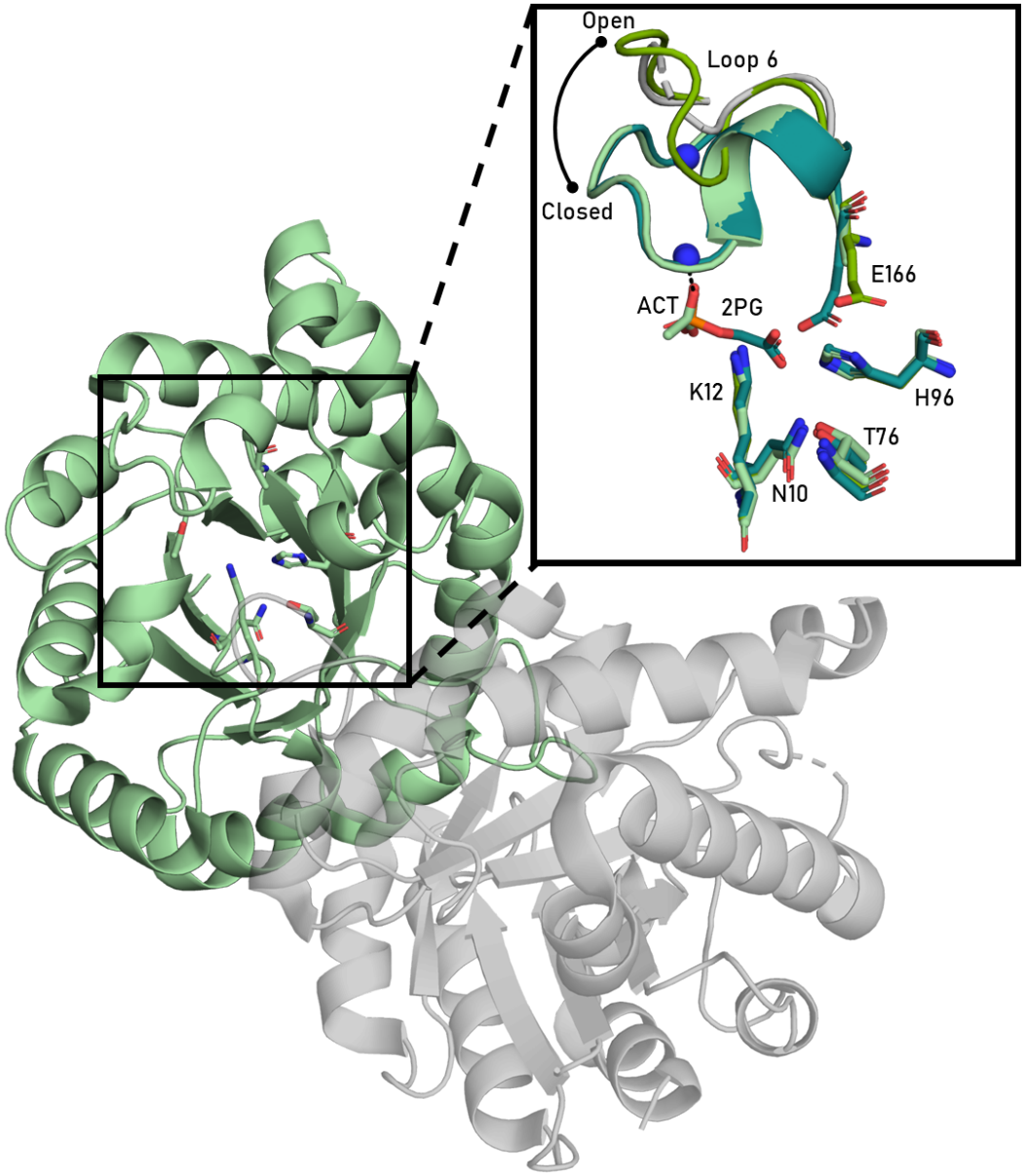


Figure 4.4, Structure of the p*Cau*TPI dimer and active site. Chain A of the p*Cau*TPI dimer is pale green while chain B is grey. The inset depicts the active site of p*Cau*TPI chains A and B compared to *Spc*TPI; teal, and p*Ath*TPI; green, which were solved with loop 6 in the closed and open conformation respectively. Grey dashed lines indicate unresolved residues in loop 6 of p*Cau*TPI chain B. The key catalytic residues are displayed as sticks and are labelled according to their equivalent position in c*Ath*TPI. The blue spheres identify the amide backbone between I171 and G172. Black dashed line identifies the hydrogen bonds between acetate (ACT) and 2-phosphoglycolate (2PG) and the amide backbone in loop 6. Both hydrogen bonds are 2.7 Å.

4.6.2 Where may relaxing purifying selection be acting in the p*Cau*TPI structure?

The conservation of key catalytic features and overall structure of p*Cau*TPI begs the question, where does p*Cau*TPI diverge from p*Ath*TPI? Relaxed purifying selection was identified in the p*Cau*TPI gene by using the ratio of synonymous to nonsynonymous mutations in the p*Cau*TPI gene compared to the chloroplast *TPI* genes of seven reference species (Sun *et al.,* 2018). These seven reference species were the model plant *A. thaliana* and six species of flowering plants in the clade lamiid, to which *C. australis* also belongs. Using Blastp and TargetP 2.0 (Altschul *et al.,* 1990, Armenteros *et al.,* 2019), I identified 15 chloroplast localised TPIs encoded in the genomes of these seven organisms. After excluding two TPIs which were duplicates, one which was an inactive splice variant (López-Castillo *et al.,* 2016), and one which contained an early C-terminal truncation, I performed conservation analysis on the remaining 11 chloroplast TPIs using ConSurf (Landau *et al.,* 2005 Ashkenazy *et al.,* 2010 and 2016) and mapped the results onto the structure of p*Ath*TPI (Figure 4.5 A). p*Cau*TPI was not included in conservation analysis so as not to bias the results towards the p*Cau*TPI structure. This analysis enabled me to identify not only where substitutions between these two TPI have occurred but whether these substitutions occurred at highly mutable or highly conserved sites.

Conservation analysis demonstrated that the amino acid identity of the chloroplast TPIs is high amongst the seven reference organisms, however, due to the small dataset used in this analysis the conservation status of some amino acid positions could not be identified (Figure 4.5 A). There are 47 divergent residues between p*Cau*TPI and p*Ath*TPI, however, 15 of these residues differed between p*Cau*TPI and p*Ath*TPI but not between p*Cau*TPI and c*Ath*TPI and consequently, were excluded from analysis under the assumption that these substitutions occurred in p*Ath*TPI after the divergence of *C. australis* and *A. thaliana* and therefore could not be a result of relaxing purifying selection acting upon p*Cau*TPI. Of the remaining 32 residues that differed between p*Cau*TPI and p*Ath*TPI, nine were in positions that were otherwise highly conserved across the set of chloroplast TPIs (F7C, K24R, D50E, D107N, V124I, D159E, K204E, E241D and T250A, residue in p*Ath*TPI is identified first), five of these nine differing residues maintain amino acid properties (Figure 4.5 A). Of the remaining 23 residues 9 were identified at positions that varied amongst the set of chloroplast TPIs, while 14 residues were present at positions where conservation status could not be determined (Figure 4.5 A). The residues which

differed between p*Cau*TPI and p*Ath*TPI were primarily located in the external α-helices of the TIM-barrel and had solvent exposed sidechains (Figure 4.5 A). The location of differing residues appeared independent of whether the residue was at a conserved or variable position. Only two sequence differences between p*Cau*TPI and p*Ath*TPI were observed in the β-sheets that compose the centre of the TIM-barrel, I124 and I163 in p*Cau*TPI, although the side chains of these residues are not positioned into the cavity of the barrel and additionally they conserve amino acid properties, being valines in the p*Ath*TPI structure. For this analysis it can be inferred where in the p*Cau*TPI structure relaxed purifying selection is acting. The residues which are conserved across the seven chloroplast TPIs but not in p*Cau*TPI provide the strongest evidence for relaxed purifying selection. Thus it can be inferred that relaxed purifying selection is primarily acting upon the residues which compose the α-helices.

In Figure 4.5 A, my analysis was designed to specifically identify relaxed purifying selection in p*Cau*TPI compared to other chloroplast TPIs. However, I also broadened the search for patterns of conservation using a more expansive dataset in order to assess whether the patterns of amino acid conservation are maintained when sampling TPIs across broader phylogenetic space (Figure 4.5 B). This analysis used 59 TPIs for which a representative X-ray crystallographic structure was available (Appendix F). 52 of these structures were obtained from the protein data bank (PDB) and the remaining seven were solved by Dr. Chelsea Vickers. All the TPI structures were wild type, except for the TPI of *Giardia intestinalis* for which no wild type structure is available, consequently the C202A mutant (PDB ID: 4B17) was used. This list of TPI structures was curated by Dr. Vickers. Unsurprisingly, when sampling TPIs from broader phylogenetic origins less conservation is observed overall, particularly in the α-helices which compose the outside of the TIM-barrel (Figure 4.5 B). The regions which are conserved include many residues at the centre of the barrel, including the catalytic residues, and the residues in loop 6 (Figure 4.5 B). When inspecting the aforementioned divergent residues, 23 were present in highly variable positions, three were in neutral positions and six were present in conserved positions although notably none of these positions were strongly conserved (Figure 4.5 B). Of those in mutable positions, five, N107, I124, I163, T178 and I189 of p*Cau*TPI, were present in either mutable or indeterminate positions in the previous Consurf analysis, while one, N106, was previously identified as a variable position (Figure 4.5).

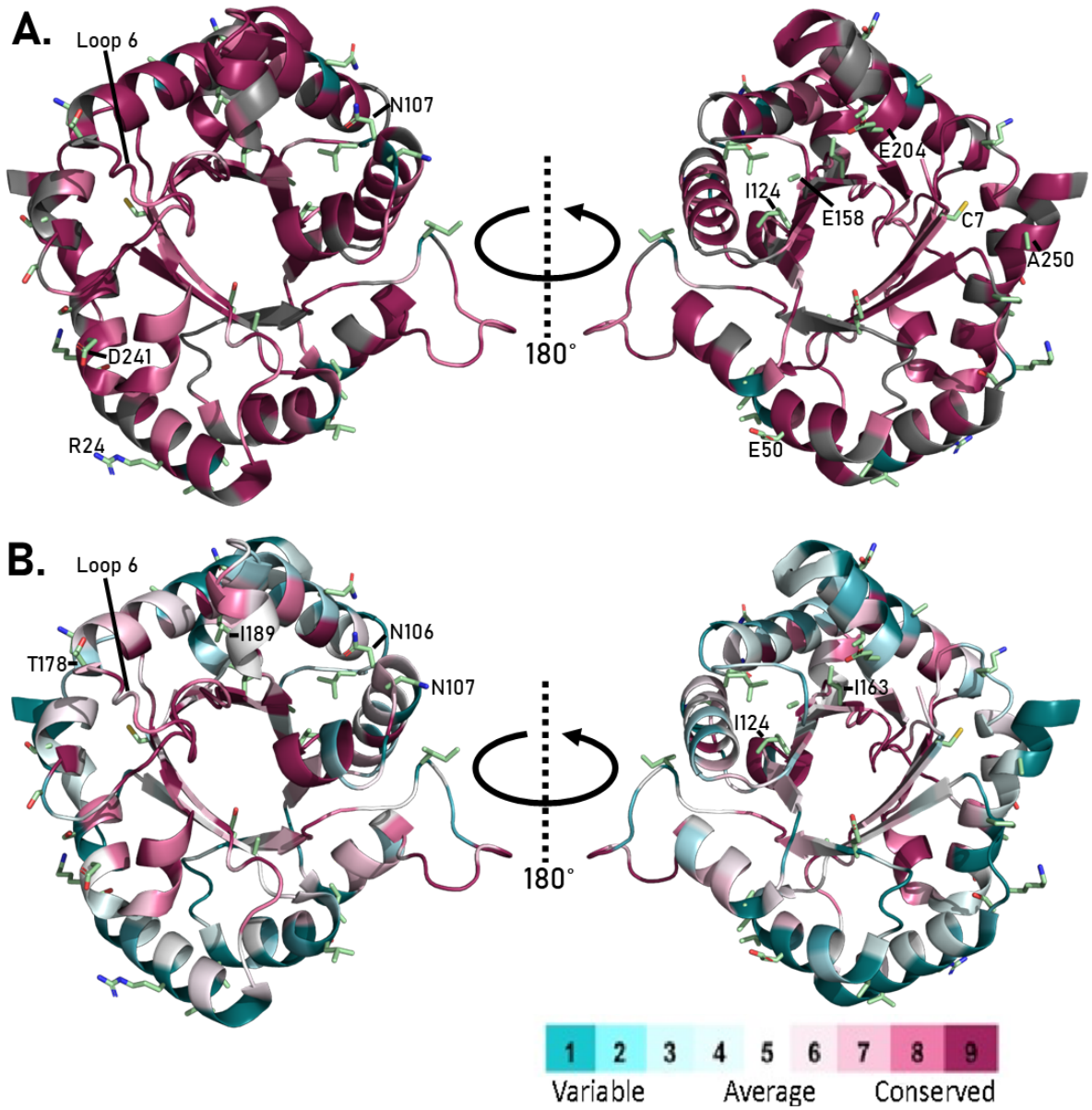


Figure 4.5, Amino acid conservation analysis of p*Cau*TPI and p*Ath*TPI. A. Consurf conservation analysis (Landau *et al.,* 2005 Ashkenazy *et al.,* 2010 and 2016) using the chloroplast TPIs of the seven reference species used to identify relaxing purifying selection in p*Cau*TPI (Sun *et al.,* 2018). **B.** Consurf conservation analysis using TPIs with representative wild type structures available. In both **A** and **B**, the conservation analysis is mapped onto the structure of p*Ath*TPI (PDB ID: 4OHQ). Positions of indeterminate conservation are in dark grey. The side chains of p*Cau*TPI residues that are divergent from p*Ath*TPI are shown as pale green sticks and are overlaid onto the p*Ath*TPI structure. The divergent residues in conserved positions are numbered according to the equivalent position in c*Ath*TPI. The side chain of E158 was not fully resolved in the p*Cau*TPI structure. A full list of the TPIs used in **B.** can be found in appendix F.

4.6.3 How different are c*Cau*TPI and p*Cau*TPI from other known TPI structures?

To generate a more holistic assessment of how c*Cau*TPI and p*Cau*TPI compare to other known TPI structures a TPI 'structure space' was generated by correspondence analysis. The structure space was generated using the Dali server (Holm and Rosenström, 2010) and the 59 representative TPI structures (Section 4.6.2, Appendix F), plus the structures of c*Cau*TPI and p*Cau*TPI. As the c*Cau*TPI is not yet finalised this correspondence analysis is preliminary.

In general the TPIs cluster according to their taxonomic domain (Figure 4.6 A). Visualisation of correspondence analysis dimensions one and two, accounting for 58.8 % and 14.4 % of the explained variation respectively, generated distinct clusters for archaeal, bacterial and eukaryotic TPIs. Closer inspection of the eukaryotic TPI structure space reveals no subcluster of TPIs from photosynthetic eukaryotes (Figure 4.6 B). Additionally, there are too few TPIs from photosynthetic eukaryotes to identify distinct subclusters for cytoplasmic and chloroplast TPIs. However, c*Cau*TPI and p*Cau*TPI are in a distinct subcluster from the TPIs of other photosynthetic eukaryotes. These two subclusters represent TPIs solved with loop 6 in the open and closed conformations. There is one exception, the *Gallus gallus* TPI (*Gga*TPI) was solved in the open conformation but clusters more closely with closed eukaryote TPIs (Figure 4.6 B, D). This 2.5 Å structure was the first TIM barrel structure solved in 1975 (PDB ID: 1TIM, Banner *et al.*). However, *Gga*TPI is also of comparatively lower quality than subsequent TPI structures, which may confound the correspondence analysis. No higher quality structures of wildtype *Gga*TPI have been solved, consequently I cannot confirm this.

c*Cau*TPI and chain A of p*Cau*TPI are the first TPIs from photosynthetic eukaryotes solved with loop 6 in the closed conformation (Figure 4.6 D). Accordingly, c*Cau*TPI and p*Cau*TPI cluster more closely with other eukaryotic TPIs solved in the same conformation than with the TPIs from other photosynthetic eukaryotes. The separation of TPIs based on loop 6 conformation is also present when inspecting dimension three, which accounts for 9.9 % of the explained variation (Figure 4.6 C). Consequently, c*Cau*TPI and p*Cau*TPI cluster more closely with *Spc*TPI and other bacteria TPIs solved with a closed loop 6 than the open TPIs of other photosynthetic eukaryotes. Curiously, *Gga*TPI clusters with the other open TPI when inspecting dimension three. The other six dimensions, which range from 4.2 and 1.7 % of the explained variation in this analysis, show no clear pattern of clustering (Appendix G).

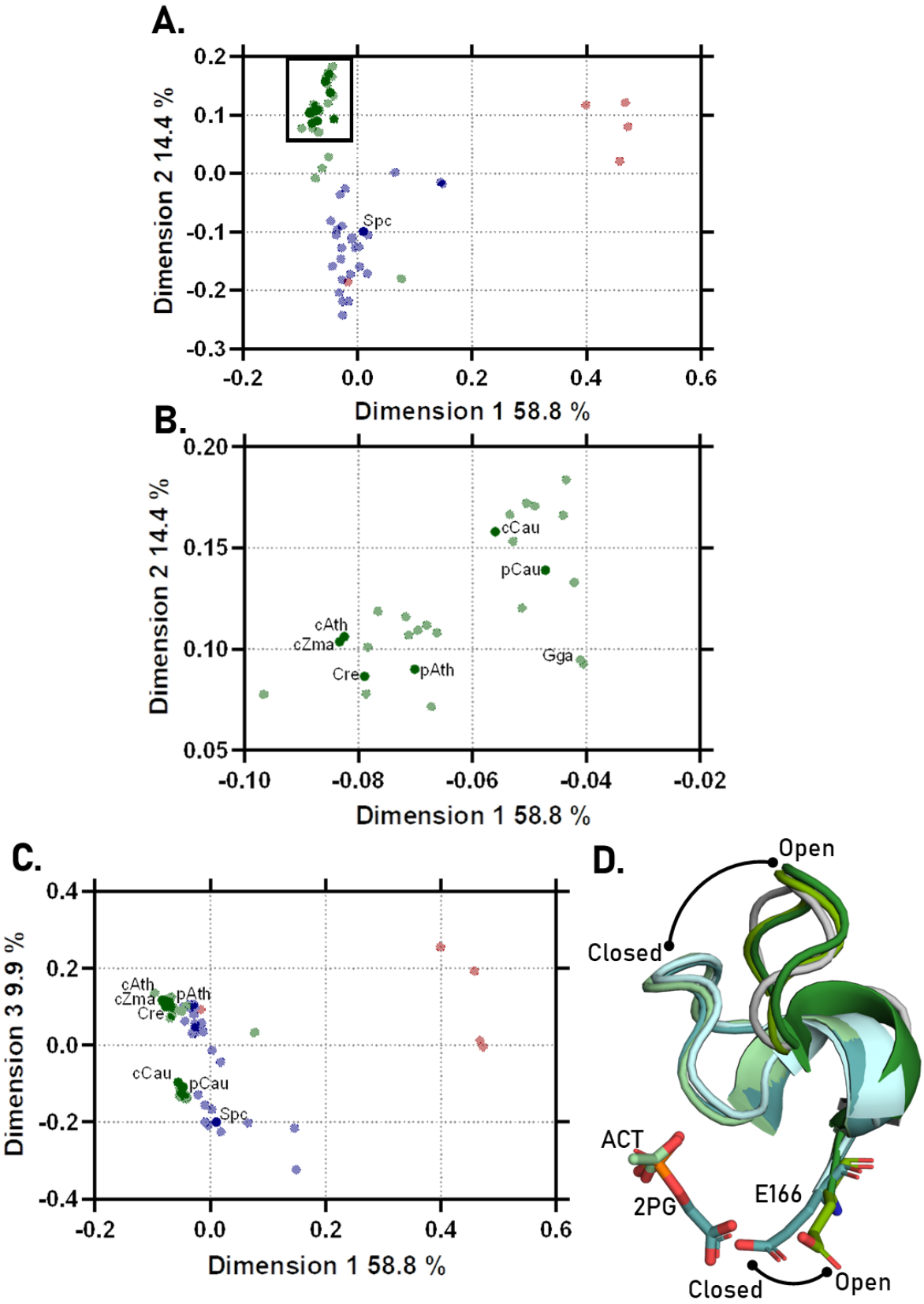

A.
Dimension 2 14.4 %
Dimension 1 58.8 %
Spc
B.
Dimension 2 14.4 %
Dimension 1 58.8 %
cCau
pCau
cAth
cZma
Cre
pAth
Gga
C.
Dimension 3 9.9 %
Dimension 1 58.8 %
cAth
cZma
pAth
Cre
cCau
pCau
Spc
D.
Open
Closed
ACT
2PG
E166
Closed
Open

Figure 4.6, Dali correspondence analysis of c*Cau*TPI and p*Cau*TPI. A. First and second dimension of correspondence analysis using chain A of p*Cau*TPI and the preliminary structure of c*Cau*TPI. The square identifies the eukaryotic TPIs which are shown in more detail in panel **B**. **C.** First and third dimension of correspondence analysis using chain A of p*Cau*TPI and preliminary structure of c*Cau*TPI. Eukaryote TPIs are in green, prokaryote TPIs are in blue and archaea TPIs are in red. The TPIs from photosynthetic organisms are shown as darker data points and labelled appropriately; c*Ath*, cytoplasmic TPI of *Arabidopsis thaliana* (PDB ID: 4OBT); p*Ath*, chloroplast TPI of *A. thaliana* (PDB ID: 4OHQ); c*Cau*, cytoplasmic TPI of *Cuscuta australis* (preliminary structure); p*Cau*, chloroplast TPI of *C. australis* (PDB ID: 7SKJ); *Cre*, *Chlamydomonas reinhardtii* (PDB ID: 4MKN); *Spc*, *Synechocystis* sp. PCC6803 (PDB ID: 6BVE); c*Zma*, cytoplasmic TPI of *Zea mays* (PDB ID: 6CG9). The TPI of *Gallus gallus* is labelled *Gga* (PDB ID: 1TIM). A full list of the TPIs used in this analysis can be found in appendix F. **D.** Comparison of TPI loop 6 and E166 in open and closed conformations. Teal, *Spc*TPI. Pale green, p*Cau*TPI. Pale blue, c*Cau*TPI (preliminary structure). Gray, *Gga*TPI. Green, p*Ath*TPI. Dark green, c*Ath*TPI. ACT, acetate ion. 2PG, 2phosphoglycolate. E166 of p*Ath*TPI and *Spc*TPI are depicted as sticks in open and closed conformation, respectively. Residues are numbered according to the equivalent position in c*Ath*TPI.

4.7 Discussion

The goal of this chapter was to test how 33 million years of relaxing purifying selection has manifested in the structure and function of the redundant chloroplast TPI from the non-photosynthetic parasitic dodder, *C. australis*. Surprisingly, *C. australis* maintains a highly active chloroplast TPI, although it does possess a k_{cat} approximately two fold lower than the *A. thaliana* orthologue (Table 5.1). In an attempt to explain this difference, I solved the structures of both the cytoplasmic and chloroplast TPIs of *C. australis*. Although the structure of c*Cau*TPI remains preliminary, analysis of the p*Cau*TPI structure demonstrates the conservation of key catalytic residues which is consistent with the reported efficiency of this enzyme. Divergent residues between p*Cau*TPI and p*Ath*TPI, and thus the traces of relaxed purifying selection, are primarily located in the α-helices that compose the exterior of the TIM-barrel fold.

4.7.1 On why *C. australis* maintains an efficient chloroplast TPI

Consistent with my initial hypothesis, I report a reduction in the catalytic efficiency of p*Cau*TPI compared to p*Ath*TPI, however, p*Cau*TPI remains an exquisitely efficient enzyme. This begs the question, if *C. australis* can no longer effectively photosynthesise, why does it retain an efficient chloroplast localised TPI?

One possible reason that the chloroplast TPI of *C. australis* is retained is that it is not redundant, as I initially assumed. Instead this TPI may be carrying out a yet unknown function in the non-photosynthetic chloroplast of *C. australis*. The retention of chloroplast localised TPIs in non-photosynthetic degenerate chloroplasts is not unprecedented. Many apicomplexan parasites, such as the malaria causing parasite *Plasmodium falciparum*, possess a degenerate non-photosynthetic chloroplast, called an apicoplast, and an apicoplast localised TPI that is believed to be required for isoprenoid synthesis (Ralph *et al.,* 2004, Lim and McFadden, 2010). Additionally, the non-photosynthetic Euglena, *Euglena longa*, possesses a degenerate non-photosynthetic chloroplast and a chloroplast localised TPI (Füssy *et al.,* 2020). Unlike in the apicoplast, the degenerate chloroplasts of *E. longa* do not carry out isoprenoid biosynthesis, instead the chloroplast TPI is part of a linearised CBB cycle which is speculated to be required for redox balance in *E. longa* (Füssy *et al.,* 2020). At first glance such adaptations of TPIs to new metabolic contexts appear antithetical to the identification of relaxed purifying selection acting upon the p*Cau*TPI by Sun *et al.* (2018). However, it is feasible that in a new hypothetical biological context p*Cau*TPI may not require the same exquisite efficiency necessary for photosynthesis, consequently a degree of relaxed purifying selection and subsequent loss of activity is tolerated. This would be consistent with p*Cau*TPI and p*Ath*TPI divergent residues being primarily located in the α-helices, which are fairly well conserved amongst chloroplast localised TPIs but are not well conserved when considering TPI from more diverse phylogenetic origins. Mutations in the α-helices appear to be tolerated, likely as they have negligible effect on catalytic activity, as observed by the maintenance of p*Cau*TPIs high catalytic efficiency, although these mutations may alter protein interactions of TPI. In contrast, more disruptive mutations in the key catalytic features may not be tolerated, hence why they are not observed in p*Cau*TPI.

Alternatively, it is possible that p*Cau*TPI is an artefact and that the 33 million years since dodders diverged from their last free living ancestors has not been sufficient for relaxed purifying selection to evolve TPI novelty or cleanse p*Cau*TPI from the *C. australis* genome. This is plausible given the paradigm of parasite evolution. The transition from free living to parasitic is incremental with facultative parasites acting as a 'stepping stone' to obligate parasitism (Luong and Mathot, 2019). Facultative parasitic plants are capable of fulfilling their lifecycle as free living autotrophs, therefore, by definition, they must maintain photosynthesis (Lyko and Wicke, 2021). In contrast obligate parasites are host dependent and have varying photosynthetic ability. Obligate parasitic plants can be further divided into hemiparasites, which retain their photosynthetic ability, and holoparasites, which do not. Dodders have varying photosynthetic ability and there is debate whether some species, including *C. australis*, are hemi- or holo-parasitic. This would suggest dodders have only recently, or are currently transitioning to, holoparasitism (Nickrent, 2020). Consequently, relaxed purifying selection may yet result in the evolution of p*Cau*TPI novelty, or the cleaning of p*Cau*TPI from the *C. australis* genome, as dodders further adapt to holoparasitism.

4.7.2 Comparing the TPI of *C. australis* to those of other parasitic plants

At the end of 2021, there were four additional parasitic plant genomes available, for: the facultative parasite *Phtheirospermum japonicum* (Cui *et al.,* 2020); the hemiparasite *Striga asiatica* (Yoshida *et al.,* 2019); the recently holoparasitic dodder *Cuscuta campestris* (Vogel *et al.,* 2018); and the holoparasite *Sapria himalayana* (Cai *et al.,* 2021). *P. japonicum*, *S. asiatica* and *C. campestris* each have at least one chloroplast TPI, whereas no chloroplast TPI was identified in *S. himalayana*. *S. himalayana* is part of the enigmatic rafflesiaceae family, a family of root endoparasites which diverged from their last photosynthetic ancestor some 100 million years ago (Pelser *et al.,* 2019). Since divergence, *S. himalayana* has lost 44 % of conserved plant genes, in contrast to the 11.7 % lost in *C. australis* (Cai *et al.,* 2021, Sun *et al.,* 2018). In part this difference in gene loss is due to the reduced selection pressures associated with endoparasitism, as by living inside the host the endoparasite is exposed to a consistent external environment

(Thorogood *et al.,* 2021).The lack of a chloroplast TPI in *S. himalayana* is consistent with relaxing purifying selection cleansing photosynthesis related genes and supports the argument that given more time evolution will likely cleanse the *C. australis* chloroplast TPI. Curiously, no cytoplasmic TPI was identified in the facultative parasite *P. japonicum* which may indicate that this organism has a glycolytic pathway split across the chloroplast and cytoplasm, as in the green alga *C. reinhardtii* (Johnson and Alric, 2013), and indicating flexibility in both glycolysis and the CBB cycle of plants during adaptation to parasitism.

4.8 Conclusion

In this chapter I report that p*Cau*TPI has a catalytic efficiency half that of its *A. thaliana* orthologue which is consistent with my initial hypothesis of relaxed purifying selection manifesting in TPI structure and function as reduced catalytic efficiency. However, the catalytic efficiency of p*Cau*TPI remains high and is approximately equal to that of c*Cau*TPI. Thus it appears that evolution of parasitism and the loss of photosynthesis in *C. australis* has subtler consequences on the adaptations of TPIs than I initially considered. Considering that *C. australis* has only recently transitioned to holoparasitism, revisiting this question in another 33 million years may provide further insights into how relaxed purifying selection has shaped the evolution of TPIs in *C. australis*.

Chapter 5. Characterising TPIs from the underexplored phylum Rhodophyta

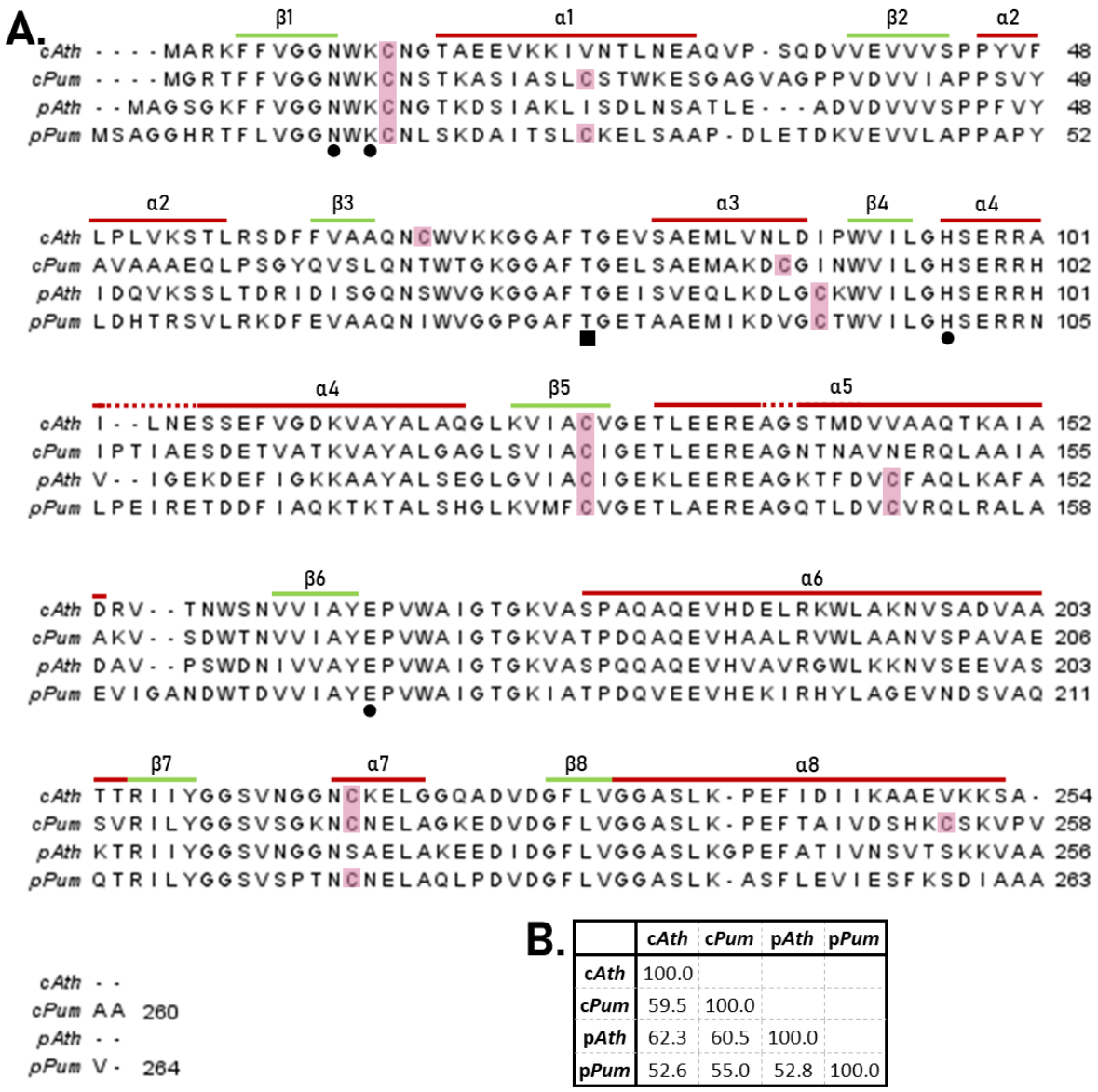
5.1 Background and hypothesis

The division between red algae (Rhodophyta) and green algae is one of the deepest divisions amongst photosynthetic eukaryotes, diverging approximately 900 million years ago (Berney and Pawlowski, 2006). There are over 7000 species in the phylum Rhodophyta (Guiry and Guiry, 2022). Rhodophyta are primarily aquatic and include economically important species, such as intertidal seaweeds in the genera *Porphyra* and *Chondrus* which are widely consumed as a food stuff. Red algae are of particular interest to molecular biologists due to the efficiency of their rubiscos, which are twice as selective for CO_2 over O_2 than those of C3 crop plants while maintaining a similar catalytic rate (Flamholz *et al.,* 2019). Consequently, considerable effort has been made to 'improve' photosynthesis through the recombinant expression of red algal rubiscos in crop plants (Lin and Hanson, 2017). It is predicted that successful recombinant expression of a red algal rubisco would result in reduced rates of 2PG production and therefore reduce photosynthetic in-efficiency due to photorespiration. Despite keen interest in the rubiscos of Rhodophyta, many of the other enzymes in the Rhodophyta CBB cycle have been neglected.
In this chapter I characterise the TPIs of the red alga *Porphyra umbilicalis*. Of particular interest was the sensitivity of *P. umbilicalis* TPIs to 2PG, due to the renowned low oxygenation activity of red algal rubiscos (Read and Tabita, 1994, Uemera *et al.,* 1996, Whitney *et al.,* 2001). I hypothesised that the reduced production of 2PG in the red algal chloroplast would be reflected in the sensitivity of red algal TPIs to 2PG. In particular, I hypothesised that red algal TPIs would have increased sensitivity to 2PG due to reduced selection to maintain catalytic activity in the presence of 2PG. To test this hypothesis, I have determined the steady-state kinetic parameters and 2PG inhibition constants of the cytoplasmic and chloroplast TPIs from *P. umbilicalis* for comparison with the TPIs from *A. thaliana* (Section 3.3.3). This is the first reported kinetic characterisation of TPIs from the phylum Rhodophyta.

5.2 Mining the Genome of *Porphyra umbilicalis*

The genome of *P. umbilicalis* was released five years ago (Brawley *et al.,* 2017). Using Blastp (Altschul *et al.,* 1990) and the amino acid sequences of c*Ath*TPI and p*Ath*TPI I identified two TPIs encoded in the genome of *P. umbilicalis*. I predicted that these TPIs were one cytoplasmic TPI isoform and one chloroplast TPI isoform based on MSA against the *Ath*TPIs (Figure 5.1 A). These TPIs were dubbed c*Pum*TPI and p*Pum*TPI. TargetP 2.0 (Armenteros *et al.,* 2019) was used to confirm the chloroplast localisation of p*Pum*TPI. Although TargetP 2.0 did confirm chloroplast localisation of p*Pum*TPI it did not precisely identify the position of the localisation tag cleavage site, instead identifying a 35 amino acid region in which the site was located (Appendix B). This region included catalytic residues N10 and K12. The cleavage site was instead inferred based on the MSA, observing that the cleavage sites in green plants occurred at a methionine residue, and that an equivalent methionine was present in the 35 amino acid region identified by TargetP 2.0. Consequently, I deduced that this methionine was the probable cleavage site in p*Pum*TPI. Residues upstream of this methionine were excluded from the p*Pum*TPI expression construct (Appendix A).

c*Pum*TPI and p*Pum*TPI share 55 % sequence identity (Figure 5.1 B). c*Pum*TPI has approximately 60 % sequence identity with both *Ath*TPIs, which is only slightly less than the 62 % between c*Ath*TPI and p*Ath*TPI. In contrast, p*Pum*TPI has approximately 53 % sequence identity with both *Ath*TPIs. Based on sequence alone this indicates that p*Pum*TPI is the most divergent TPI investigated in this project.



	cAth	cPum	pAth	pPum
cAth	100.0			
cPum	59.5	100.0		
pAth	62.3	60.5	100.0	
pPum	52.6	55.0	52.8	100.0

Figure 5.1, Multiple sequence alignment of *A. thaliana* and *P. umbilicalis* TPIs. A. MSA using Clustal Omega (Seivers *et al.,* 2011). Secondary structural elements are based on the solved structure of p*Ath*TPI (PDB ID: 4OHQ) with dashed lines indicating breaks in secondary structural elements. Positions of cysteine residues are highlighted in pink. ● indicates catalytic residues while the ■ indicates key residue of the dimer interface. **B.** Percent identity matrix generated by Clustal Omega. MSA and identity matrix were generated without the chloroplast localisation tags of p*Ath*TPI and p*Pum*TPI.

5.3 Expression and Purification of c*Pum*TPI and p*Pum*TPI

Both c*Pum*TPI and p*Pum*TPI were solubly expressed and purified as previously described and yielded protein comparable to previous preparations of *Ath*TPIs (section 3.2). Both *Pum*TPIs behaved as anticipated during IMAC purification beginning elution at approximately 50% buffer B, corresponding to 200 mM imidazole, and saturating the detector (Figure 5.2 A, C). Subsequent SDS-PAGE analysis of the resulting IMAC fractions revealed relatively pure protein bands between 26 and 34 kDa, which is consistent with the predicted molecular weight of 30.5 and 29.1 kDa for His_6-c*Pum*TPI and His_6-p*Pum*TPI respectively (Figure 5.2 A, B). When *Pum*TPI fractions were pooled and concentrated prior to SEC, white crystalline aggregates rapidly formed in the concentrator. The addition of the reducing agent DTT did not stop aggregate formation. These aggregates were not observed for any other TPI purified during this project. These aggregates were removed by centrifugation and filtering prior to SEC. As was observed for *Ath*TPIs, both c*Pum*TPI and p*Pum*TPI eluted from the SEC at approximately 10 mL indicating that these TPIs are likely also dimers (Figure 5.2 B, D). However, unlike previous TPI preparations a pronounced shoulder is observed trailing the elution peaks of both *Pum*TPIs (Figure 5.2 B, D). This shoulder is perhaps indicative that a heterogeneous population of TPI monomers and dimers present in solution. Alternatively, these shoulders may be indicative of TPI degradation. However, no bands indicative of degradation were observed during subsequent SDS-PAGE analysis (Figure 5.2 B, D). Both c*Pum*TPI and p*Pum*TPI were sufficiently pure for further analysis, although only the central fraction of each peak was used to minimise any heterogeneity or degradation products present. Attempts to crystallize these proteins to solve their structures by X-ray crystallography were unsuccessful, likely due to the instability which was observed during purification.

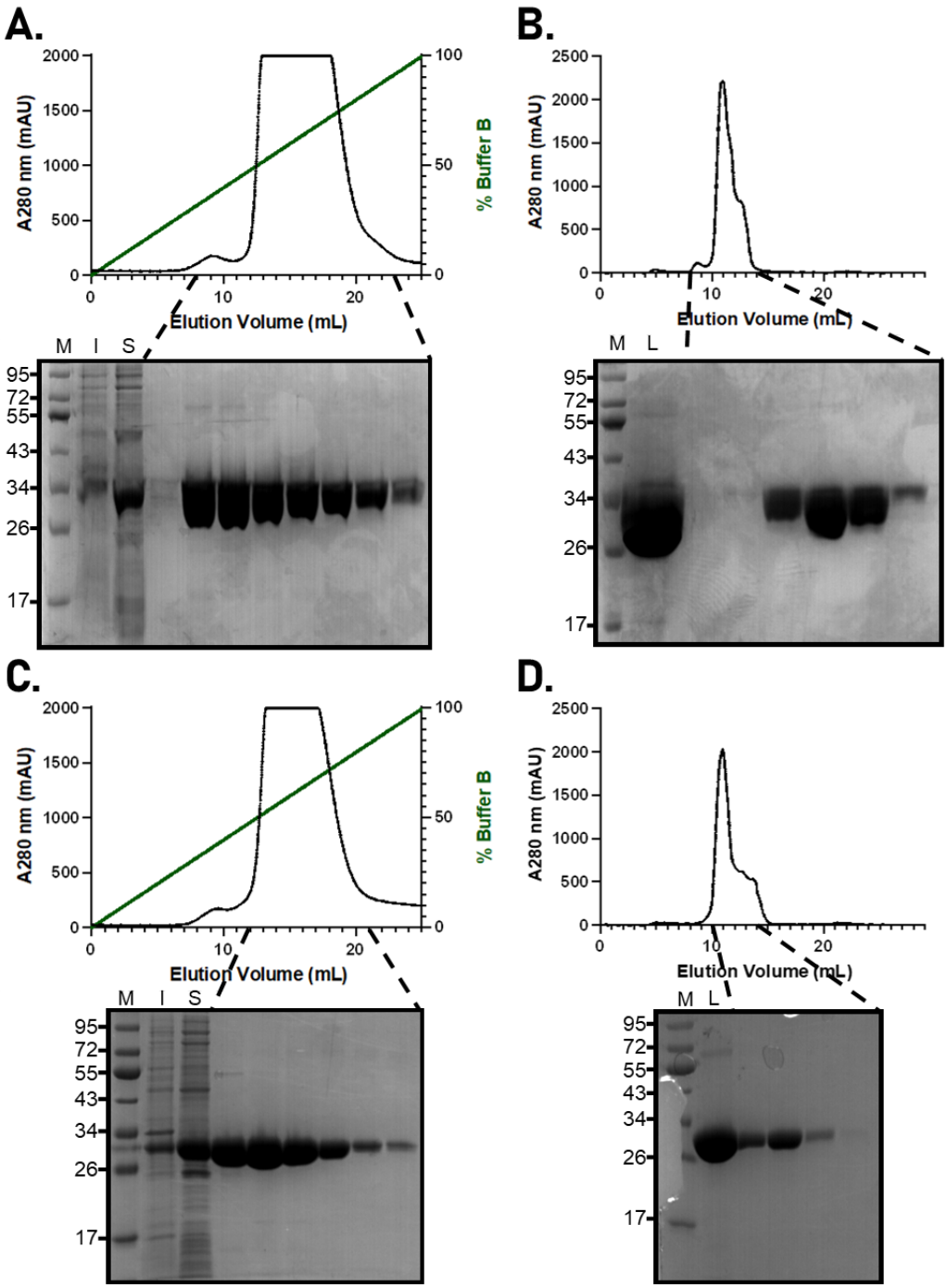


Figure 5.2, Purification of c*Pum*TPI and p*Pum*TPI. A, IMAC trace of c*Pum*TPI. **B,** SEC trace of c*Pum*TPI. **C,** SEC trace of p*Pum*TPI. **D,** SEC trace of p*Pum*TPI. SDS-PAGE gels beneath traces show representative fractions from the absorbance peaks indicated. M, molecular weight standards (kDa). I, insoluble fraction, and S, soluble fraction, of whole cell lysis by sonication. L, load fraction; the pooled and concentrated IMAC fractions prior to SEC.

5.4 Steady-state kinetic characterisation of c*Pum*TPI and p*Pum*TPI

5.4.1 The instability of p*Pum*TPI

Initial attempts to gather steady-state kinetic data on p*Pum*TPI were unsuccessful as this enzyme rapidly lost activity during trial assays, which was not observed for c*Pum*TPI. Fluorescence thermal shift (FTS) assays were performed to estimate the melting temperature (T_m) of p*Pum*TPI and to ensure it was not denaturing during the assay. c*Pum*TPI was used as a control and the assays were carried out in duplicate. The T_m values of c*Pum*TPI and p*Pum*TPI were estimated to be 46.6 ± 0.2 and 42.5 ± 0.3 ℃, respectively (mean ± SEM). The T_m of p*Pum*TPI is greater than the assay temperature (30 ℃) thus it is unlikely that p*Pum*TPI was undergoing temperature dependent denaturation during kinetic analysis.

Notably during FTS assays p*Pum*TPI had a significantly higher initial fluorescence compared to c*Pum*TPI, indicating the presence of exposed hydrophobic regions (Appendix H). The addition of 1 mM DTT to p*Pum*TPI decreased the initial absorbance during FTS assays, although did not decrease initial absorbance to what was observed for c*Pum*TPI. The addition of DTT also resulted in a slight increase in p*Pum*TPI T_m. The same effects of DTT were not observed for c*Pum*TPI. Based on this data I hypothesised that p*Pum*TPI was undergoing oxidation dependent aggregation during kinetics assays. This was investigated further using DTT dependent stability assays of p*Pum*TPI (Figure 5.3). In the absence of DTT, thawed aliquots of p*Pum*TPI that were incubated in SEC buffer (Table 2.1) lost almost all activity within 10 minutes. However, the addition of 10 mM DTT immediately upon thawing mitigated activity loss. Although an initial decrease in activity was still observed, 1 mM DTT in assays stabilised p*Pum*TPI and this enzyme remained active after 30 minutes. No appreciable loss of p*Pum*TPI activity was observed between 5 and 30 minutes in the presence of DTT. In contrast, no difference in c*Pum*TPI activity was observed in the presence or absence of DTT. Based on this analysis subsequent p*Pum*TPI assays were carried out in the presence of 1 mM DTT to minimise oxidation dependent loss of catalytic activity.

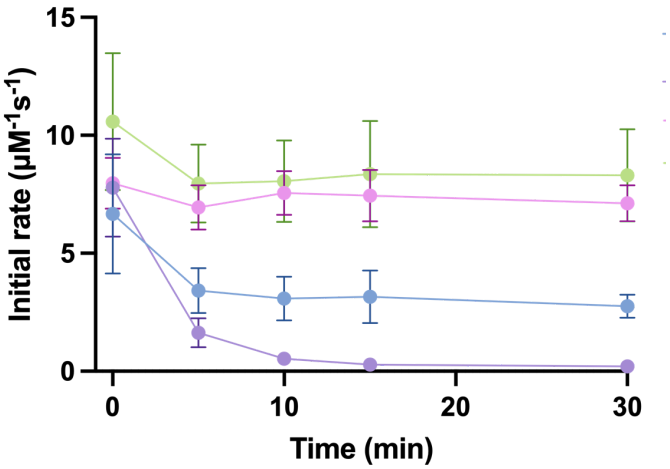


Figure 5.3, DTT time course analysis of c*Pum*TPI and p*Pum*TPI. Upon thawing the TPIs were immediately diluted in SEC buffer in the presence or absence of 10 mM DTT. These dilutions were incubated at room temperature until aliquots were taken and diluted 1 in 10 upon initiating the assay. The final concentration of DTT in the assay was 1 mM. Green, c*Pum*TPI. Pink, c*Pum*TPI + DTT. Purple, p*Pum*TPI. Blue, p*Pum*TPI + DTT. Data points are the mean ± SEM of three technical replicates. Where error bars are not visible the, the error is smaller than the width of the data point.

5.4.2 Kinetics and 2-phosphoglycolate inhibition of c*Pum*TPI and p*Pum*TPI

Steady-state kinetic analysis of both c*Pum*TPI and p*Pum*TPI was undertaken and the results were compared to the equivalent *Ath*TPIs (Table 5.1). Both the k_{cat} and K_M of c*Pum*TPI are approximately half that of c*Ath*TPI, consequently the k_{cat}/K_M of these two enzymes are approximately equal. In contrast, the k_{cat}/K_M of p*Pum*TPI is approximately one fifth that of p*Ath*TPI and c*Pum*TPI. This difference in catalytic efficiency is due to p*Pum*TPI having a reduced k_{cat} and twice the K_M compared to p*Ath*TPI and c*Pum*TPI (Table 5.1). The comparisons between p*Pum*TPI and the other TPI are confounded by this enzyme's rapid loss of activity. To mitigate this p*Pum*TPI assays were carried out immediately after thawing p*Pum*TPI, and in the presence of 1 mM DTT. In a control experiment, p*Pum*TPI was preincubated with DTT at room temperature for one hour and then assayed. As expected (Figure 5.3), preincubation reduced the apparent k_{cat} ten-fold to 500 s^{-1}, while the apparent K_M was 1.7 mM resulting in a reduced k_{cat}/K_M of 3.2×10^5 $s^{-1}M^{-1}$.

In the 2PG inhibition assays a mixed model of inhibition was used based on the results obtained for *Ath*TPIs (Section 3.3.4). As was demonstrated for *Ath*TPIs a mixed model of inhibition was a

significantly better fit for the available data than a competitive model for each *Pum*TPI replicate (Figure 5.4). No difference in K_i was observed between p*Pum*TPI and any of c*Pum*TPI, c*Ath*TPI or p*Ath*TPI (Table 5.1). This result contradicts my initial hypothesis that the sensitivity of p*Pum*TPI to 2PG would increase, and hence K_i would decrease, compared to c*Pum*TPI and *Ath*TPIs.

Table 5.1, Comparison of steady-state kinetic parameters for *Pum*TPIs and *Ath*TPIs

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}M^{-1}$)	K_i (μM) 2PG Mixed
c*Ath*TPI	17000 ± 2000	0.9 ± 0.1	2.1 ± 0.1 ×10^7	100 ± 50
c*Pum*TPI	10000 ± 2000	0.51 ± 0.09	2.2 ± 0.5 ×10^7	65 ± 7
p*Ath*TPI	12000 ± 500	0.61 ± 0.06	2.0 ± 0.2 ×10^7	70 ± 7
p*Pum*TPI*	5000 ± 1000	1.2 ± 0.2	4.2 ± 0.3 ×10^6	80 ± 30

*Assays were run with 1 mM DTT
Data is reported as mean ± SEM. For k_{cat}, K_M and k_{cat}/K_M; n = 3 (p*Ath*TPI, p*Pum*TPI), 4 (c*Ath*TPI), or 5 (c*Pum*TPI) independent data sets. Each dataset contained technical quadruplicates for each measured substrate concentration. For K_i; n = 2 (c*Ath*TPI, c*Pum*TPI and p*Pum*TPI) and 3 (p*Ath*TPI). Each dataset contained technical triplicates for each measured combination of substrate and 2PG concentration. Data for p*Ath*TPI, and c*Ath*TPI were generated using TPI from a single purification of each protein. Data for c*Pum*TPI and p*Pum*TPI was generated using TPI from two independent purifications.

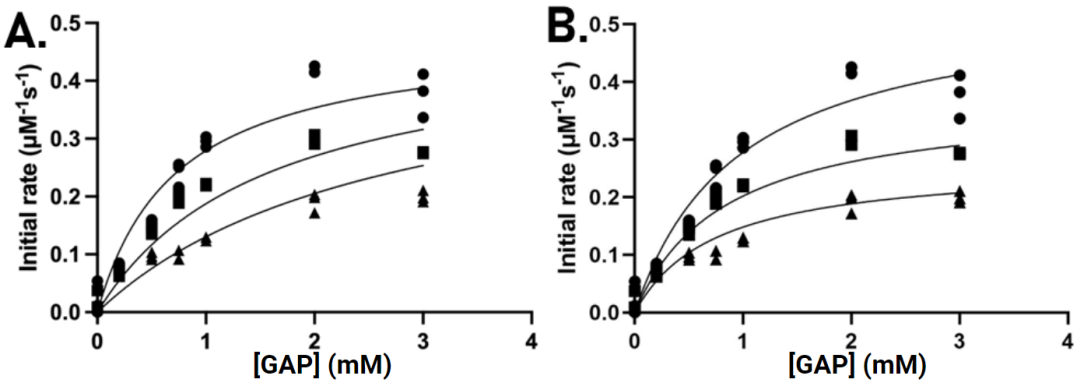


Figure 5.4, Comparison of competitive and mixed models of 2PG inhibition for p*Pum*TPI. A. Competitive inhibition model. **B.** Mixed inhibition model. The rate data are the same in A and B; only the model chosen to fit the data is changed. ●, 0 μM 2PG; ■, 30 μM 2PG; ▲, 70 μM 2PG. Each measurement was performed in triplicate. To identify the better fitting model an extra sum-of-square F test was carried out. The resulting F statistic was 31.5 with an associated p-value <0.0001 indicating that the mixed model is a significantly better fit for the available data.

5.5 Structural modelling of c*Pum*TPI and p*Pum*TPI

As c*Pum*TPI and p*Pum*TPI were not amenable to crystallisation, their structures could not be solved by X-ray crystallography. Instead, structural models of these TPIs were generated using a modified version of AlphaFold v2.1.0 (Jumper *et al.,* 2021) which was available through GoogleColab (Figure 5.5). Due to the comparable elution volume of *Pum*TPIs and *Ath*TPIs during SEC (Figures 3.1, 5.2), I modelled both *Pum*TPIs as dimers (Figure 5.5). AlphaFold v2.1.0 produced high accuracy models for both proteins based on the predicted local distance difference test (pLDDT) provided by AlphaFold v2.1.0; with the pLDDT score for the majority of residues exceeding 90 % indicating high accuracy (Appendix I, Tunyasuvunakool *et al.,* 2021). The regions of each model predicted with lower accuracy are on the exterior of the TIM-barrel away from the dimer interface. These models were used to formulate a hypothesis on the structural basis which underlies the dramatic difference in sensitivity to oxidation.

As I was interested in oxidation sensitivity, I began by identifying p*Pum*TPI specific cysteine residues. Both c*Pum*TPI and p*Pum*TPI have six cysteine residues, while both *Ath*TPIs contain four cysteines (Figure 5.1 A). Four of the cysteines are conserved between c*Pum*TPI and p*Pum*TPI; C13, C24, C130 and C221 (Figure 5.1 A). The remaining cysteines of p*Pum*TPI, C92 and C149, are between α-helix three and β-sheet four and in α-helix five respectively. Both C92 and C149 do not have a solvent accessible sidechain, consequently, it appears unlikely that either of these residues are responsible for the differential oxidation sensitivity of c*Pum*TPI and p*Pum*TPI. This led me to inspect the remaining cysteine residues more closely.

C13 is conserved in the TPIs of photosynthetic eukaryotes; it is present at the dimer interface and is adjacent to the catalytic residues N10 and K12 (Zaffagini *et al.,* 2014). In c*Ath*TPI, p*Ath*TPI, and the TPI of the green alga *Chlamydomonas reinhardtii*, the oxidation of C13 results in reduced catalytic activity (Zaffagini *et al.,* 2014, Lopez-Castillo *et al.,* 2016). In light of this data the dimer interfaces of the *Pum*TPIs were inspected. In the models of c*Pum*TPI and p*Pum*TPI a two amino acid insertion, which disrupts the fourth α helix (Figure 5.1 A), forms additional contacts between the TPI monomers (Figure 5.5). In c*Pum*TPI this insertion is a PT, while in p*Pum*TPI this insertion is a PE (Figure 5.1 A, 5.5). In both models the T/E105 are buried in a pocket formed by residues from loop three and the fourth α helix of the other subunit. In p*Pum*TPI this pocket is formed by three nonpolar residues, W69, V70 and F112, whereas in c*Pum*TPI there are two polar threonine residues present in the pocket at positions 70 and 112

(Figure 5.5). I predict that the position of the polar, charged E105 into this hydrophobic pocket is destabilising the p*Pum*TPI dimer interface, therefore increasing solvent exposure of C13 making it prone to oxidation with subsequent loss of activity. The equivalent insertion in c*Pum*TPI does not destabilise the dimer interface to the same extent due to T105 being accommodated by the other polar residues in this pocket (Figure 5.5). This hypothesis is consistent with the ability of DTT to partially stop the loss of p*Pum*TPI activity and with the T_m of p*Pum*TPI being 4 ℃ lower than that of c*Pum*TPI (Section 5.4.2).

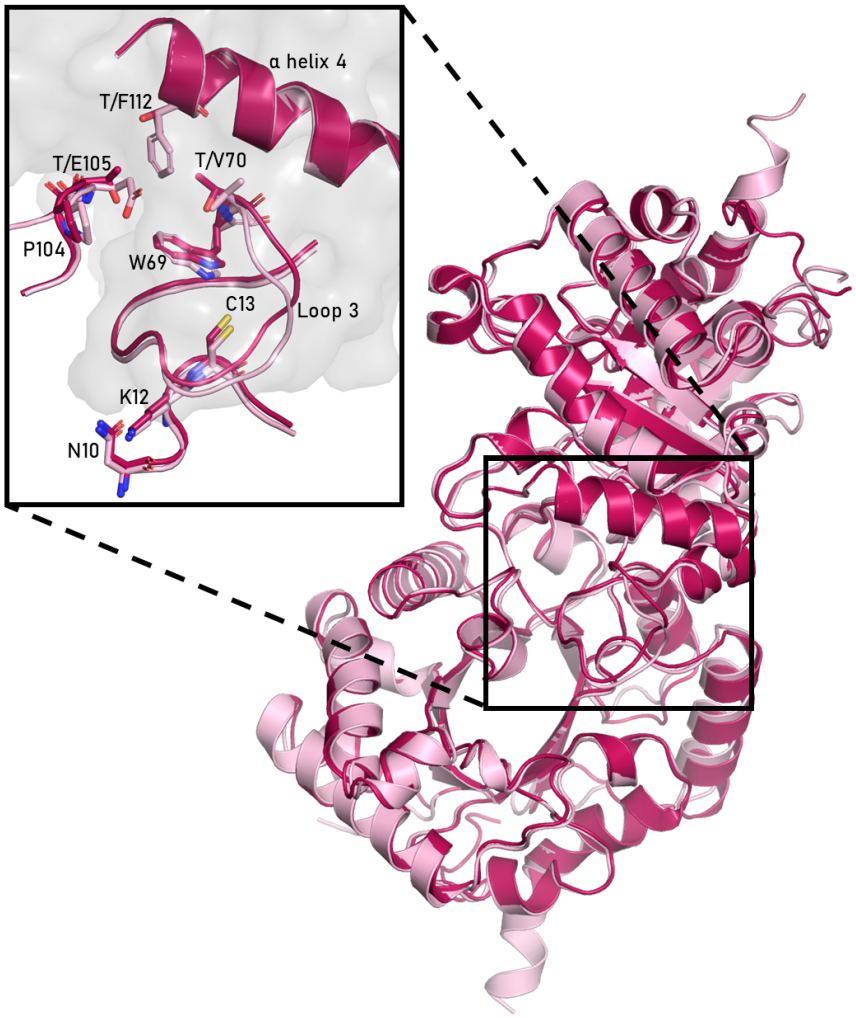


Figure 5.5, Comparison of the c*Pum*TPI and p*Pum*TPI dimer interfaces. Structural models were generated using AlphaFold v2.1.0 (Jumper *et al.,* 2021). Red, c*Pum*TPI. Pink, p*Pum*TPI. The inset highlights the predicted interactions of the PT/PE insert at the dimer interface, with the pale grey surface identifying regions on the opposite subunit. Residues are labelled according to the equivalent position in c*Pum*TPI.

5.6 Discussion

The goal of this chapter was to test if the low rate of rubisco catalysed oxygenation, and therefore the reduced production of the TPI inhibitor 2PG, in red algae was manifested in the structure and function of the chloroplast TPI of *P. umbilicalis*. Contrary to my initial hypothesis that the p*Pum*TPI would be more sensitive to inhibition by 2PG, I report no difference in 2PG sensitivity between any of c*Pum*TPI, p*Pum*TPI and *Ath*TPIs. However, unexpectedly I have demonstrated that the p*Pum*TPI is exquisitely sensitive to oxidation and possesses reduced catalytic efficiency compared to c*Pum*TPI and *Ath*TPIs.

5.6.1 On the oxidation sensitivity of p*Pum*TPI

I have hypothesised that the dramatic difference between c*Pum*TPI and p*Pum*TPI oxidation sensitivity and catalytic efficiency is due to the identity of a two amino acid insertion at the dimer interface of these two TPI (Section 5.5). However, this begs the question, how well conserved are these insertions, and the predicted interacting residues, amongst Rhodophyta TPIs? Previous phylogenetic studies have identified that this insertion is unique to Rhodophyta and Euglena TPIs (Sun *et al.,* 2008). Euglena are the only photosynthetic genera in the eukaryotic clade Discoba (Schoch *et al.,* 2020), and are more closely related to the parasitic trypanosomes and giardia than to other photosynthetic eukaryotes (Zakryś *et al.,* 2017). The common insertion in Rhodophyta and Euglena TPIs is due to an ancestral horizontal gene transfer event where Euglena acquired a TPI from a red alga (Sun *et al.,* 2008). However, I am the first to report biochemical data on a TPI from any of these organisms.

At the end of 2021 there were nine additional Rhodophyta genomes available. Using Blastp (Altschul *et al.,* 1990) I identified 13 additional Rhodophyta TPIs encoded in these nine genomes. The same analysis was not performed for Euglena TPIs due to the prevalence of genetic mosaicism amongst this genus (Füssy *et al.,* 2020) meaning I could not be certain that the identified TPIs were of Rhodophyta origin. Of the 15 Rhodophyta TPIs identified C13 was conserved in all 15, while the two amino acid insert in the region disrupting α4 helix was present in 13, the exceptions being the cytoplasmic TPIs of *Chondrus crispus* and *Gracilariopsis chorda*. The consensus sequence of the insert was P104 E/Q105 while the consensus of the pocket residues was W69, V/I70, however, there was no clear consensus between threonine and phenylalanine residues at position 112 (Appendix J). The substitutions in the consensus sequence

at positions 70 and 105 conserve amino acid properties. This consensus sequence is consistent with p*Pum*TPI, while c*Pum*TPI is the only known Rhodophyta TPI with threonines at positions 70 and 105. This indicates that the dimer interface of p*Pum*TPI is more typical amongst Rhodophyta than that of c*Pum*TPI. If my hypothesis (Section 5.5) is correct, this also indicates that the oxidation sensitivity of p*Pum*TPI is the rule for Rhodophyta TPIs, both cytoplasmic and chloroplast, while the oxidation insensitivity of c*Pum*TPI is the outlier.

5.6.2 On the 2-phosphoglycolate sensitivity of *Pum*TPIs

In *A. thaliana* the partial inhibition of p*Ath*TPI by 2PG has been proposed to regulate flux through the CBB cycle (Section 1.4.1, Sharkey, 2018). The partial inhibition of p*Ath*TPI maintains the TPI reaction away from its equilibrium position, thus the relative concentrations of GAP and DHAP are closer to the optimal 1:1 ratio for the subsequent fructose 1,6-bisphosphate aldolase and transketolase catalysed reactions. However, in contrast to my initial hypothesis, I have discovered no difference in 2PG sensitivity between p*Pum*TPI and p*Ath*TPI. As Rhodophyta rubiscos catalyse less 2PG formation via the oxygenation reaction, it follows that 2PG will be present at a reduced concentration in the chloroplasts of *P. umbilicalis*. In turn, this may mean that p*Pum*TPI is relatively uninhibited, compared to p*Ath*TPI, consequently the GAP:DHAP ratio would lie closer to its equilibrium of 1:20 which would be deleterious for the subsequent reactions in the CBB cycle. The consequences of low 2PG production in Rhodophyta may be further exacerbated by the presence of carbon concentrating mechanisms, which are prevalent in marine macroalgae (Johnston and Raven 1987). Carbon concentration mechanisms reduce rubisco catalysed oxygenation by increasing the local concentration of CO_2 at the active site of rubisco. Although these mechanisms have not been investigated in *P. umbilicalis* they have been identified in various species from the related genus *Pyropia* (Mercado *et al.,* 1997, Wang *et al.,* 2020, Zhang *et al.,* 2020). Although I speculate that p*Pum*TPI being relatively uninhibited would be deleterious for overall flux through the CBB cycle, I also report that p*Pum*TPI has relatively lower activity than p*Ath*TPI (Table 5.1). Additionally, there is no available data for the relative expression levels of p*Pum*TPI and p*Ath*TPI. Therefore, it remains possible that the GAP:DHAP ratio in *P. umbilicalis* chloroplasts, and Rhodophyta chloroplasts in general, is tuned by a yet unknown 2PG independent mechanism.

5.7 Conclusion

Rhodophyta have become of interest to plant biotechnologists due to the high selectivity of their rubiscos for CO_2 over O_2. However, other enzymes in the Rhodophyta CBB cycle have long been neglected. In this chapter I report the first kinetic characterisation of Rhodophyta TPIs and have identified the relatively low activity and unique oxidation sensitivity of p*Pum*TPI. This result highlights the underexplored diversity of CBB cycle enzymes amongst Rhodophyta. Further investigation of these enzymes may provide serendipitous opportunities to further engineer photosynthesis in crop plants by exploiting the evolved adaptations that already exist in the chloroplasts and chloroplast localised enzymes of red algae.

6. Future directions and final remarks

6.1 On chapter 3; the TPIs of *A. thaliana* and *Synechocystis*

Previous investigations into three of these TPIs, *Spc*TPI, c*Ath*TPI and p*Ath*TPI, have formed much of the basis for understanding of the role of TPI in the CBB cycle and how adaptation to the chloroplast has shaped TPI evolution (Dumont *et al.,* 2016, López-Castillo *et al.,* 2016, Castro-Torres *et al.,* 2018). At the outset of this investigation I intended to use these three TPIs as a benchmark for the subsequent characterisation of the TPIs of *C. australis* and *P. umbilicalis*, however, the results of my investigation into 2PG inhibition of the *A. thaliana* TPIs were unanticipated. Although widely considered a competitive inhibitor (Anderson *et al.,* 1971, Lolis and Petsko, 1990, Flügel *et al.,* 2017, Li *et al.,* 2019), my data demonstrated that the 2PG inhibition of *Ath*TPIs is better explained by a mixed model (Table 3.2, Figure 3.4). I hypothesised that the mixed inhibition of 2PG is due to a two-step inhibitor binding mechanism where the opening of the catalytic loop, loop 6, is at least partially rate limiting (Section 3.4.3, Figure 3.5). One of the hallmarks of a two-step inhibitor binding mechanism is time dependent inhibition (Blat, 2010). If 2PG is a two-step inhibitor of TPI, then preincubation of TPI and 2PG would favour the accumulation of the enzyme-inhibitor complex (EI*, Figure 3.5), which is 'resistant' to competition with the substrate and consequently, the apparent K_i of 2PG would decrease. This could be tested by adjusting the design of the 2PG inhibition assays where instead of adding 2PG to the reagent mixture it is added to the dilution of TPI prior to initiating the assay.

6.2 On chapter 4; relaxed purifying selection and the TPIs of *C. australis*

In chapter four, I characterised the TPIs of the parasitic plant *C. australis*. I hypothesised that relaxed purifying selection acting upon p*Cau*TPI, due to the loss of photosynthesis, would erode the catalytic efficiency of this TPI compared to its *A. thaliana* orthologue. Consistent with this hypothesis I have demonstrated that although p*Cau*TPI remains a highly efficient enzyme, it is half as efficient as p*Ath*TPI (Table 4.1). Unexpectedly, I also demonstrate that c*Cau*TPI has

experienced an equivalent loss in catalytic efficiency (Table 4.1). In an attempt to uncover a structural basis for the observed differences in catalytic efficiency I solved the structure of both c*Cau*TPI and p*Cau*TPI by X-ray crystallography. Although the structure of c*Cau*TPI remains preliminary, structural analysis of p*Cau*TPI identified that key catalytic features of TPI are conserved which is consistent with the maintenance of high p*Cau*TPI catalytic efficiency (Section 4.6.1). Additionally, conservation analysis of p*Cau*TPI (Section 4.6.2) demonstrated that the effects of relaxed purifying selection are primarily observed in the α-helices of the TIM-barrel and although conserved amongst chloroplast TPIs are far less conserved when sampling TPIs across broader phylogenetic space. I have postulated on why p*Cau*TPI may maintain such high catalytic efficiency (Section 4.7). Characterising the TPIs from other organisms identified in this discussion, such as the chloroplast TPI of *E. longa* postulated to be involved in a linearised CBB cycle involved in redox regulation (Füssy *et al.,* 2020), or the cytoplasmic TPI of *P. japonicum,* a facultative parasite which maintains photosynthesis despite lacking a chloroplast localised TPI, may provide further insight into how the evolution of parasitism in plants and any associated relaxed purifying selection, has shaped the evolution of TPI and further, the CBB cycle as a whole.

6.3 On chapter 5; the TPIs of the red alga *P. umbilicalis*

In this project I present the first reported kinetic characterisation of a Rhodophyta TPI. I began my investigation of the TPIs from the understudied red alga *P. umbilicalis* with a keen interest in the sensitivity of p*Pum*TPI to inhibition by 2PG, as red algae are renowned for the increased selectivity of their rubiscos for CO_2 over O_2 and consequently the reduced production of 2PG. Contrary to my initial hypothesis, I reported no difference in 2PG sensitivity between any of c*Pum*TPI, p*Pum*TPI, c*Ath*TPI or p*Ath*TPI (Table 5.1). While it will undoubtedly be informative to study other enzymes from the CBB cycle of *P. umbilicalis* and other red algae, I speculate that p*Pum*TPI being relatively uninhibited compared to p*Ath*TPI would have a deleterious effect on overall flux through the CBB cycle of *P. umbilicalis* (Section 5.6.2).

Although not part of my initial aim for this chapter, I discovered differential sensitivity of p*Pum*TPI and c*Pum*TPI to oxidation (Section 5.4.1). I hypothesised that this difference is due to the differential identity of a two amino acid insertion at the dimer interface of these two enzymes (Section 5.5). By mining the available genomes of other red algae, I speculated that the exquisite

sensitivity of p*Pum*TPI is more typical of Rhodophyta TPIs (Section 5.6.1). To test my hypothesis mutant *Pum*TPIs have been ordered where the identity of the residues of interest at the dimer interface, T/V70, T/E105 and T/F112 (Figure 5.5), have been swapped between c*Pum*TPI and p*Pum*TPI. For c*Pum*TPI, I predict that these mutations will destabilise the dimer interface resulting in a decreased T_m and an increase in redox sensitivity and consequent loss of TPI activity. Whereas, for p*Pum*TPI I predict that the mutations will have the opposite effect.

6.4 Final remarks

In this project I have purified and characterised seven TPIs, including solving the structure of two of these TPIs by X-ray crystallography, from four photosynthetic organisms. Using structure/function and kinetic paradigms I have investigated how adaptation to the chloroplast may have shaped TPI evolution and based on this data speculated on the implications on the activity of the Rhodophyta CBB cycle and the impacts of relaxed purifying selection on the TPIs of parasitic plants.

This project highlights the diversity of biochemistries in yet underexplored phyla. However, there remains groups of photosynthetic eukaryotes, such as glaucophyte algae and Euglenids, which are yet to be explored in the context of TPI evolution and oxygenic photosynthesis.

In addition to highlighting the wealth of existing biochemical diversity present in nature, my investigation into the TPI inhibitor, 2PG, begins to question the long held assumption of a strictly competitive model of 2PG inhibition.

In light of this project it is clear how much there is still to learn in the field of enzymology even in regards to TPI, one of the most biochemically well characterised enzymes currently present in literature.

7. Appendices

Appendix A; Table of protein expression constructs used in this study.

Name	Construct Sequence	MW (kDa)	E ($M^{-1}cm^{-1}$)	E %	pI
His_6p*Pu* *m*TPI	MGSSHHHHHHSSGLVPRGSHMSAGGHRTFLVGGNWKCNLSKDAITSLCKELSAAPDLETDKV EVVLAPPAPYLDHTRSVLRKDFEVAAQNIWVGGPGAFTGETAAEMIKDVGCTWVILGHSERRNL PEIRETDDFIAQKTKTALSHGLKVMFCVGETLAEREAGQTLDVCVRQLRALAEVIGANDWTDVVI AYEPVWAIGTGKIATPDQVEEVHEKIRHYLAGEVNDSVAQQTRILYGGSVSPTNCNELAQLPDVD GFLVGGASLKASFLEVIESFKSDIAAAV	30.5	33835	1.109	5.5
His_6p*Pu* *m*TPI V116T E151T F158T	MGSSHHHHHHSSGLVPRGSHMSAGGHRTFLVGGNWKCNLSKDAITSLCKELSAAPDLETDKV EVVLAPPAPYLDHTRSVLRKDFEVAAQNIWTGGPGAFTGETAAEMIKDVGCTWVILGHSERRNL PTIRETDDTIAQKTKTALSHGLKVMFCVGETLAEREAGQTLDVCVRQLRALAEVIGANDWTDVVI AYEPVWAIGTGKIATPDQVEEVHEKIRHYLAGEVNDSVAQQTRILYGGSVSPTNCNELAQLPDVD GFLVGGASLKASFLEVIESFKSDIAAAV	30.5	33835	1.107	5.59
His_6c*Pu* *m*TPI	MGSSHHHHHHSSGLVPRGSHMGRTFFVGGNWKCNSTKASIASLCSTWKESGAGVAGPPVDV VIAPPSVYAVAAAEQLPSGYQVSLQNTWTGKGGAFTGELSAEMAKDCGINWVILGHSERRHIPTI AESDETVATKVAYALGAGLSVIACIGETLEEREAGNTNAVNERQLAAIAAKVSDWTNVVIAYEPVW AIGTGKVATPDQAQEVHAALRVWLAANVSPAVAESVRILYGGSVSGKNCNELAGKEDVDGFLVG GASLKPEFTAIVDSHKCSKVPVAA	29.1	46325	1.598	6.11
His_6c*Pu* *m*TPI T70V T105E T112F	MGSSHHHHHHSSGLVPRGSHMGRTFFVGGNWKCNSTKASIASLCSTWKESGAGVAGPPVDV VIAPPSVYAVAAAEQLPSGYQVSLQNTWVGKGGAFTGELSAEMAKDCGINWVILGHSERRHIPEI AESDEFVATKVAYALGAGLSVIACIGETLEEREAGNTNAVNERQLAAIAAKVSDWTNVVIAYEPVW AIGTGKVATPDQAQEVHAALRVWLAANVSPAVAESVRILYGGSVSGKNCNELAGKEDVDGFLVG GASLKPEFTAIVDSHKCSKVPVAA	29.2	46325	1.587	5.99
His_6p*Ath* TPI	MGSSHHHHHHSSGLVPRGSHMAGSGKFFVGGNWKCNGTKDSIAKLISDLNSATLEADVDVVV SPPFVYIDQVKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSERRHVIGEKD EFIGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADAVPSWDNIVVAYEPVWAIGT GKVASPQQAQEVHVAVRGWLKKNVSEEVASKTRIIYGGSVNGGNSAELAKEEDIDGFLVGGASL KGPEFATIVNSVTSKKVAA	29.2	39210	1.342	6.41

His_6*cAth* TPI	MGSSHHHHHHSSGLVPRGSHMARKFFVGGNWKCNGTAEEVKKIVNTLNEAQVPSQDVVEVV VSPPYVFLPLVKSTLRSDFFVAAQNCWVKKGGAFTGEVSAEMLVNLDIPWVILGHSERRAILNES SEFVGDKVAYALAQGLKVIACVGETLEEREAGSTMDVVAAQTKAIADRVTNWSNVVIAYEPVWAI GTGKVASPAQAQEVHDELRKWLAKNVSADVAATTRIIYGGSVNGGNCKELGGQADVDGFLVGG ASLKPEFIDIIKAAEVKKSA	29.2	39210	1.343	6.21
His_6*Cau*T PI	MGSSHHHHHHSSGLVPRGSHMGRKFFVGGNWKCNGTSEEVKKIVSTLNAGEVPPQDVVEVV VSPPFVFLPLAKSSLRPDFLVAAQNCWIKKGGAFTGEISAEMLVNLDIPWVILGHSERRLILGESN EFVGDKVAYALSQGLKVIACVGETLEQREAGTTMDVVAAQTKAIADKVSDWSKVVIAYEPVWAIG TGKVATPAQAQEVHAELRKWLESNVSSEVASSTRIIYGGKYFVTILHITSNSYLRSSGLLDTLYMC ESIPIFLCFDKIQSDYVYMGLAHGIGLLAQVAGYM	31.7	46660	1.47	6.06
His_6c*Cau* TPI	MGSSHHHHHHSSGLVPRGSHMARKFFVGGNWKCNGTAEEVKKIVTTLNEAEVPSEDDVEVVV SPPFVFLTLVKTLLRPDFSIAAQNCWVRKGGAFTGEVSAEMLINLGIPWVILGHSERRQLLNESD DFVGDKVAYALIQGLKVIACVGETLEQRESGATMKVVAAQTEAIFDKVSNWSNIVLAYEPVWAIGT GKVATPAQAQEVHSELRKWLHDNVGPEVAASTRIIYGGSVSGANCKELATKPDVDGFLVGGASL KPEFIEIIKSATLKQNA	29.5	37720	1.332	6.21
His_6p*Cau* TPI	MGSSHHHHHHSSGLVPRGSHMAGSGKCFVGGNWKCNGTKESIVRLISDLNSSKLEPDVDVVV APPFLYIEQVKSTLTDRIEIAAQNCWIGKGGAFTGEISAEQLKDIGCKWVILGHSERRHVMGENN EFIGKKAAYASSQGVGIIACIGELLEEREARKTFDVCFQQLKAFADALPSWENVVIAYEPVWAIGT GKVATPEQAQEVHAAIRDWLNKNVSSEVASETRIIYGGSVNGSNCSELAKKEDIDGFLVGGASL KGPDFASIVNSVASKKVTA	29.5	39335	1.279	6.31
His_6*Spc*T PI	MGSSHHHHHHSSGLVPRGSHMRKIIIAGNWKMHKTQAEAQAFLQGFKPLIEDAAESREVVLCV PFTDLSGMSQQLHGGRVRLGAQNVHWEASGAYTGEISAAMLTEIGIHYVVIGHSERRQYFGET DETANLRVLAAQKAGLIPILCVGESKAQRDAGETEQVIVDQVKKGLVNVDQSNLVIAYEPIWAIGT GDTCAATEANRVIGLIREQLTNSQVTIQYGGSVNANNVDEIMAQPEIDGALVGGASLEPQSFARIV NFQP	31.7	24075	0.854	5.82

Blue text identifies the His_6 expression tag used for purification
Red text identifies residues mutated from the wildtype construct.

Appendix B; The TPI sequences of organisms investigated in this study

Name	Accession Code	Amino Acid Sequence
*Spc*TPI	WP_010873374.1	MRKIIIAGNWKMHKTQAEAQAFLQGFKPLIEDAAESREVVLCVPFTDLSGMSQQLHGGRVRLGAQNVHWEASGAYTG EISAAMLTEIGIHYVVIGHSERRQYFGETDETANLRVLAAQKAGLIPILCVGESKAQRDAGETEQVIVDQVKKGLVNVDQS NLVIAYEPIWAIGTGDTCAATEANRVIGLIREQLTNSQVTIQYGGSVNANNVDEIMAQPEIDGALVGGASLEPQSFARIVNF QP
c*Ath*TPI	NP_191104.1	MARKFFVGGNWKCNGTAEEVKKIVNTLNEAQVPSQDVVEVVVSPPYVFLPLVKSTLRSDFFVAAQNCWVKKGGAFTG EVSAEMLVNLDIPWVILGHSERRAILNESSEFVGDKVAYALAQGLKVIACVGETLEEREAGSTMDVVAAQTKAIADRVTN WSNVVIAYEPVWAIGTGKVASPAQAQEVHDELRKWLAKNVSADVAATTRIIYGGSVNGGNCKELGGQADVDGFLVGGA SLKPEFIDIIKAAEVKKSA
p*Ath*TPI	NP_001323967.1	MAATSLTAPPSFSGLRRISPKLDAAAVSSHQSFFHRVNSSTRLVSSSSSSHRSPRGVVAMAGSGKFFVGGNWKCNGT KDSIAKLISDLNSATLEADVDVVVSPPFVYIDQVKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSE RRHVIGEKDEFIGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADAVPSWDNIVVAYEPVWAIGTGKVASP QQAQEVHVAVRGWLKKNVSEEVASKTRIIYGGSVNGGNSAELAKEEDIDGFLVGGASLKGPEFATIVNSVTSKKVAA
p*Ath*TPI (splice variant)	NP_001077931.1	MAATSLTAPPSFSGLRRISPKLDAAAVSSHQSFFHRVNSSTRLVSSSSSSHRSPRGVVAMAGSGKNGTKDSIAKLISDL NSATLEADVDVVVSPPFVYIDQVKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSERRHVIGEKDEF IGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADAVPSWDNIVVAYEPVWAIGTGKVASPQQAQEVHVAV RGWLKKNVSEEVASKTRIIYGGSVNGGNSAELAKEEDIDGFLVGGASLKGPEFATIVNSVTSKKVAA
*Cau*TPI	RAL36983.1	MGRKFFVGGNWKCNGTSEEVKKIVSTLNAGEVPPQDVVEVVVSPPFVFLPLAKSSLRPDFLVAAQNCWIKKGGAFTGE ISAEMLVNLDIPWVILGHSERRLILGESNEFVGDKVAYALSQGLKVIACVGETLEQREAGTTMDVVAAQTKAIADKVSDW SKVVIAYEPVWAIGTGKVATPAQAQEVHAELRKWLESNVSSEVASSTRIIYGGKYFVTILHITSNSYLRSSGLLDTLYMCE SIPIFLCFDKIQSDYVYMGLAHGIGLLAQVAGYM
c*Cau*TPI	RAL37417.1	MARKFFVGGNWKCNGTAEEVKKIVTTLNEAEVPSEDDVEVVVSPPFVFLTLVKTLLRPDFSIAAQNCWVRKGGAFTGEV SAEMLINLGIPWVILGHSERRQLLNESDDFVGDKVAYALIQGLKVIACVGETLEQRESGATMKVVAAQTEAIFDKVSNWS NIVLAYEPVWAIGTGKVATPAQAQEVHSELRKWLHDNVGPEVAASTRIIYGGSVSGANCKELATKPDVDGFLVGGASLK PEFIEIIKSATLKQNA
p*Cau*TPI	RAL50862.1	MASAALSCQLSCPIARPNFSLRRSIPGSSTRSATQSFHTIDSRLRLVCSGKGCRSVVAMAGSGKCFVGGNWKCNGTK ESIVRLISDLNSSKLEPDVDVVVAPPFLYIEQVKSTLTDRIEIAAQNCWIGKGGAFTGEISAEQLKDIGCKWVILGHSERRH VMGENNEFIGKKAAYASSQGVGIIACIGELLEEREARKTFDVCFQQLKAFADALPSWENVVIAYEPVWAIGTGKVATPEQ

		AQEVHAAIRDWLNKNVSSEVASETRIIYGGSVNGSNCSELAKKEDIDGFLVGGASLKGPDFASIVNSVASKKVTA
c*Pum*TPI	OSX73519.1	MGRTFFVGGNWKCNSTKASIASLCSTWKESGAGVAGPPVDVVIAPPSVYAVAAAEQLPSGYQVSLQNTWTGKGGAFT GELSAEMAKDCGINWVILGHSERRHIPTIAESDETVATKVAYALGAGLSVIACIGETLEEREAGNTNAVNERQLAAIAAKV SDWTNVVIAYEPVWAIGTGKVATPDQAQEVHAALRVWLAANVSPAVAESVRILYGGSVSGKNCNELAGKEDVDGFLVG GASLKPEFTAIVDSHKCSKVPVAA
p*Pum*TPI	OSX70626.1	MAFITSGAAAVSGASLSARRAFTGAAVCPTAAATAAPVAAAWRMSAGGHRTFLVGGNWKCNLSKDAITSLCKELSAAP DLETDKVEVVLAPPAPYLDHTRSVLRKDFEVAAQNIWVGGPGAFTGETAAEMIKDVGCTWVILGHSERRNLPEIRETDD FIAQKTKTALSHGLKVMFCVGETLAEREAGQTLDVCVRQLRALAEVIGANDWTDVVIAYEPVWAIGTGKIATPDQVEEVH EKIRHYLAGEVNDSVAQQTRILYGGSVSPTNCNELAQLPDVDGFLVGGASLKASFLEVIESFKSDIAAAV

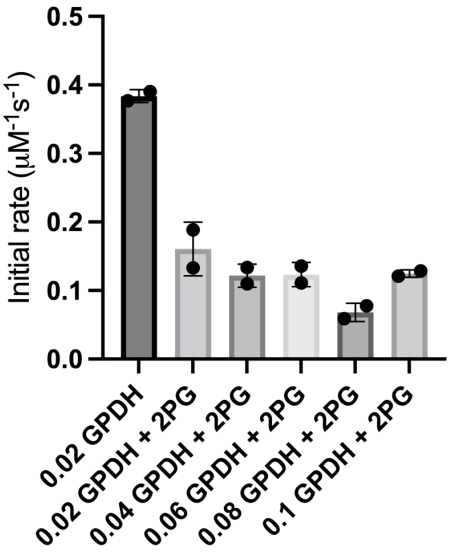
Green text indicates chloroplast localisation tags which were identified using TargetP 2.0 and MSA
Underline; Region identified by TargetP 2.0 as probable location of chloroplast localisation tag cleavage site in pPum

Appendix C, Summary of refined crystal screens

Protein	Original Condition	Adjusted conditions
c*Pum*TPI	Nuc-Pro H8; 50 mM MES pH 6.0, 20 mM $MgCl_2$, 15 % (w/v) 2-propanol	1-100 mM MgCl2, 5-20 % (w/v) 2-propanol
c*Cau*TPI	PACT Premier A2; 100 mM SPG* pH 5.0, 25 % (w/v) PEG 1500	100 mM SPG pH 4-6.7, 17.5-30 % (w/v) PEG 1500
c*Cau*TPI	PACT Premier B7; 100 mM MES pH 6.0, 200 mM NaCl, 20 % (w/v) PEG 6000	50-300 mM NaCl, 15-22.5 % (w/v) PEG 6000
c*Cau*TPI	Index HT G3; 200 mM $LiSO_4$ monohydrate, 100 mM Bis-Tris pH 6.5, 25 % (w/v) PEG 3350	50-300 mM $LiSO_4$ monohydrate, 20-27.5 % (w/v) PEG 3350
c*Cau*TPI	Index HT G4; 200 mM $LiSO_4$ monohydrate, 100 mM HEPES pH 7.5, 25 % (w/v) PEG 3350	50-300 mM $LiSO_4$ monohydrate, 20-27.5 % (w/v) PEG 3350
p*Cau*TPI	PACT Premier A3; 100 mM SPG* pH 6.0, 25 % (w/v) PEG 1500	100 mM SPG pH 4-6.7, 17.5-30 % (w/v) PEG 1500
p*Cau*TPI	Index HT G7; 200 mM ammonium acetate, 100 mM Bis-Tris pH 6.5, 25 % (w/v) PEG 3350	50-300 mM ammonium acetate, 20-27.5 % (w/v) PEG 3350
p*Cau*TPI	JCSG-plus A8; 200 mM ammonium formate, 20 % (w/v) PEG 3350	50-300 mM ammonium formate, 20-25 % (w/v) PEG 3350

* SPG buffer; succinic acid, sodium phosphate monobasic monohydrate and glycine in molar ratio 2:7:7

Appendix D, Troubleshooting 2PG inhibition assays; testing for inhibition of GPDH by 2PG



c*Ath*TPI activity was assessed in the presence of 3 mM GAP, 100 μM 2PG and increasing concentrations of the coupled enzyme GPDH. The concentration of GPDH is in mg/mL. Increasing GPDH concentration did not restore c*Ath*TPI activity to negative 2PG control. 70 pM c*Ath*TPI was used and each measurement was performed in duplicate and normalised to 0 mM GAP. Bars represent mean ± SEM with dots representing individual measurements.

Appendix E, Summary statistics of the p*Cau*TPI X-ray crystal structure.

Data or statistic	p*Cau*TPI PDB ID: 7SKJ
Beamline	AS-MX2
Wavelength (Å)	0.9464
Resolution (Å) (outer shell)	44.97–1.90 (1.94–1.90)
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	74.95, 82.12, 154.25
α, *β*, *γ* (°)	90.0, 90.0, 90.0
R_{meas} (outer shell)	0.090 (0.646)
$I/\sigma I$ (outer shell)	16.3 (3.4)
Completeness (%) (outer shell)	99.9 (96.8)
Multiplicity (outer shell)	13.6 (13.3)
Total no. of reflections (outer shell)	1030194 (56601)
No. of unique reflections (outer shell)	75702 (7447)
Redundancy (outer shell)	13.6 (13.3)
Mean half-set correlation, $CC_{(1/2)}$ (outer shell)	0.999 (0.928)
Wilson B factor ($Å^2$)	24.2
Refinement statistics	
No. of reflections (no. of test reflections)	75695 (7447)
R_{work}	0.18 (0.22)
R_{free}	0.24 (0.28)
No. of atoms: protein/ligands/water	7063/11/892
B factors: protein/ligands/water ($Å^2$)	27.5/35.0/33.9
RMSD: bond lengths/bond angles (Å/°)	0.008/0.86
Ramachandran plot statistics	
Favoured regions	97.37 %
Allowed regions	2.42 %
Outliers	0.21 %

Appendix F, Full list of TPI structures used in Consurf conservation analysis and DALI correspondence analysis

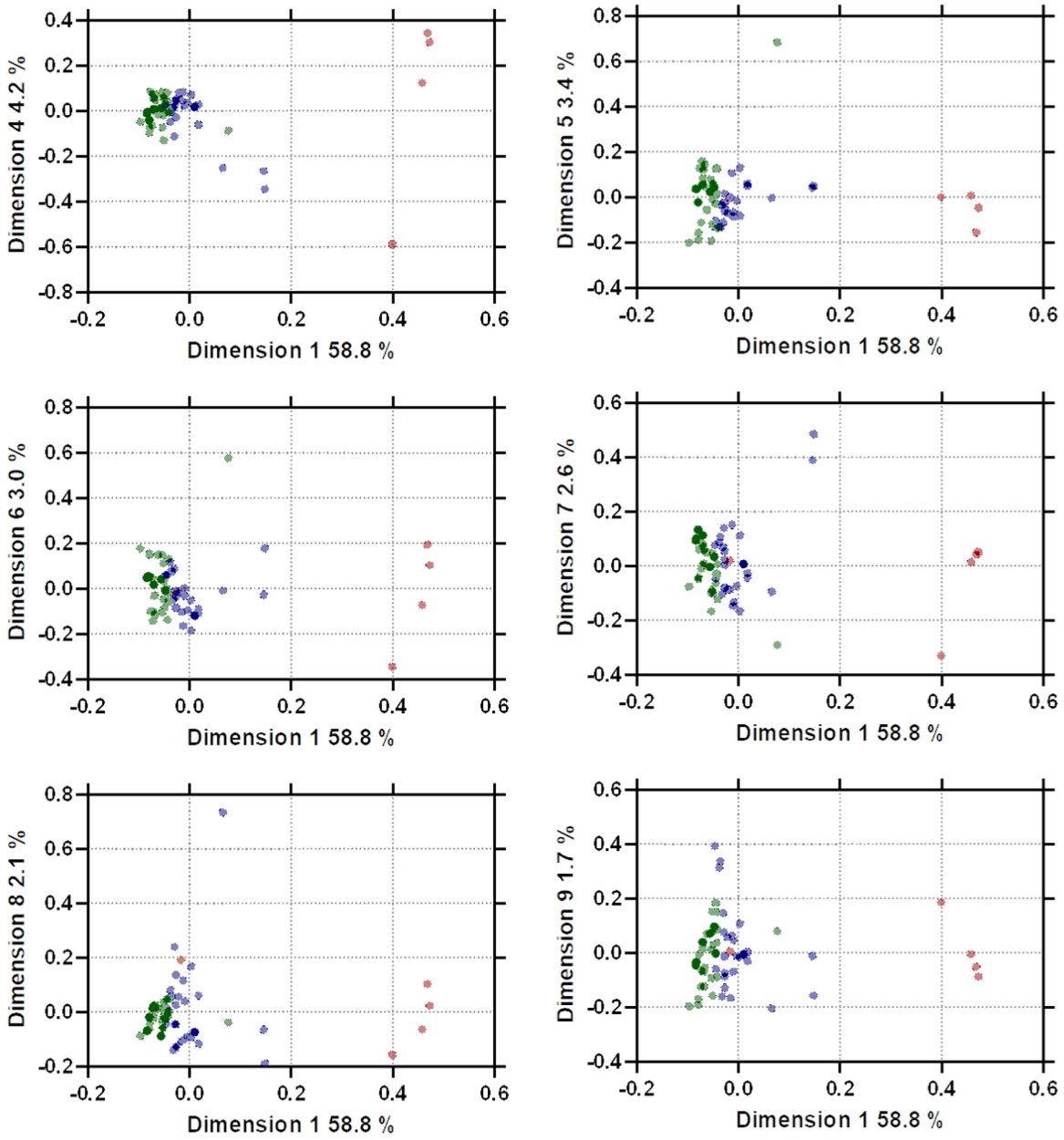
PDB ID	Organism	Domain	PDB ID	Organism	Domain
4OBT	*Arabidopsis thaliana* (cytoplasmic)	Eukaryota	1YDV	*Plasmodium falciparum*	Eukaryota
4OHQ	*Arabidopsis thaliana* (chloroplast)	Eukaryota	1HG3	*Pyrococcus woesei*	Archaea
3KXQ	*Bartonella henselae*	Bacteria	3TH6	*Rhipicephalus microplus*	Eukaryota
4NVT	*Brucella melitensis* ATCC 23457	Bacteria	1NEY	*Saccharomyces cerevisiae*	Eukaryota
4G1K	*Burkholderia thailandensis* E264	Bacteria	6OOI	*Schistosoma mansoni*	Eukaryota
1MO0	*Caenorhabditis elegans*	Eukaryota	6JOX	*Scylla paramamosain*	Eukaryota
4MKN	*Chlamydomonas reinhardtii*	Eukaryota	3M9Y	*Staphylococcus aureus subsp. aureus* MRSA252	Bacteria
4Y8F	*Clostridium perfringens* str. 13	Bacteria	6W4U	*Stenotrophomonas maltophilia* K279a	Bacteria
3S6D	*Coccidioides immitis* RS	Eukaryota	5IBX	*Streptococcus pneumoniae* D39	Bacteria
3KRS	*Cryptosporidium parvum* Iowa II	Eukaryota	4Y9A	*Streptomyces coelicolor* A3(2)	Bacteria
4Y90	*Deinococcus radiodurans* R1	Bacteria	6BVE	*Synechocystis* sp. PCC6803	Bacteria
1M6J	*Entamoeba histolytica*	Eukaryota	6OOG	*Taenia solium*	Eukaryota
4MVA	*Escherichia coli* BW2952	Bacteria	2I9E	*Tenebrio molitor*	Eukaryota
6R8H	*Fasciola hepatica*	Eukaryota	5CSR	*Thermoplasma acidophilum* DSM 1728	Archaea
5UJW	*Francisella tularensis subsp. tularensis* SCHU S4	Bacteria	1W0M	*Thermoproteus tenax*	Archaea
1TIM	*Gallus gallus*	Eukaryota	1B9B	*Thermotoga maritima*	Bacteria
4Y96	*Gemmata obscuriglobus*	Bacteria	1YYA	*Thermus thermophilus* HB8	Bacteria
4BI7	*Giardia intestinalis*	Eukaryota	5UPR	*Toxoplasma gondii* ME49	Eukaryota
1BTM	*Geobacillus stearothermophilus*	Bacteria	4WJE	*Trichomonas vaginalis*	Eukaryota
2JGQ	*Helicobacter pylori* 26695	Bacteria	6TIM	*Trypanosoma brucei brucei*	Eukaryota
4POC	*Homo sapiens*	Eukaryota	2V5B	*Trypanosoma cruzi*	Eukaryota
1AMK	*Leishmania mexicana*	Eukaryota	6CG9	*Zea mays* (cytoplasmic)	Eukaryota
4GNJ	*Leishmania* sp. 'siamensis	Eukaryota	*§	*Candidate division WWE3 bacterium*	Bacteria
4X22	*Leptospira interrogans serovar Icterohaemorrhagiae*	Bacteria	*§	*Ktedonobacter racemifer*	Bacteria
2H6R	*Methanocaldococcus jannaschii* DSM 2661	Archaea	*§	*Candidate division SR1 bacterium*	Bacteria
1AW2	*Moritella marina*	Bacteria	*§	*Candidatus Prometheoarchaeum syntrophicum*	Archaea
3GVG	*Mycobacterium tuberculosis*	Bacteria	*§	*Candidatus Roizmanbacteria bacterium*	Bacteria
5ZFX	*Opisthorchis viverrini*	Eukaryota	*§	*Verrucomicrobium spinosum*	Bacteria
1R2R	*Oryctolagus cuniculus*	Eukaryota	✠	*Cuscuta australis* (cytoplasmic)	Eukaryota
5EYW	*Penaeus vannamei*	Eukaryota	7SKJ*	*Cuscuta australis* (chloroplast)	Eukaryota

* Structures that are deposited in the PDB but are currently embargoed until publication.

§ Structures solved by Dr. Chelsea Vickers, *SR1* TPI structure was solved with the assistance of Jordan Compton

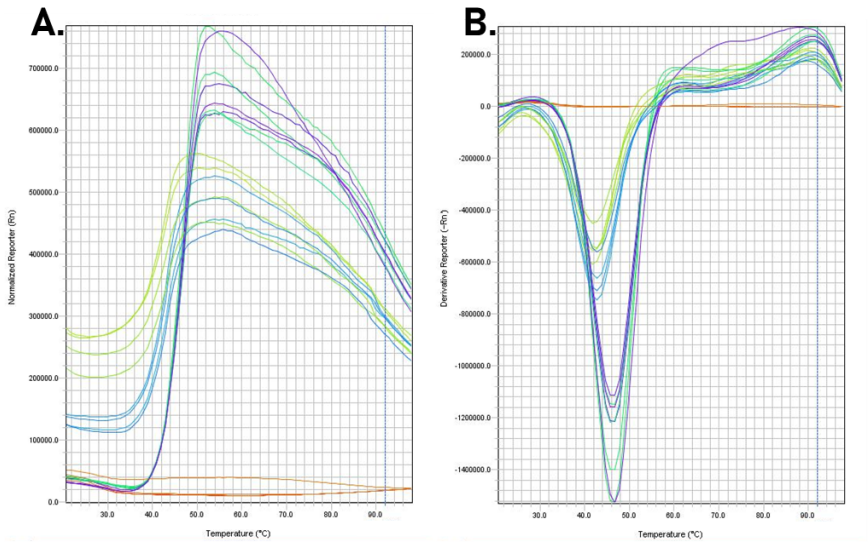
✠ Preliminary structure solved with the assistance of Emeritus Professor Dr. Geoff Jameson

Appendix G, The lesser dimensions (4-9) of TPI Dali Correspondence analysis supporting Figure 4.6



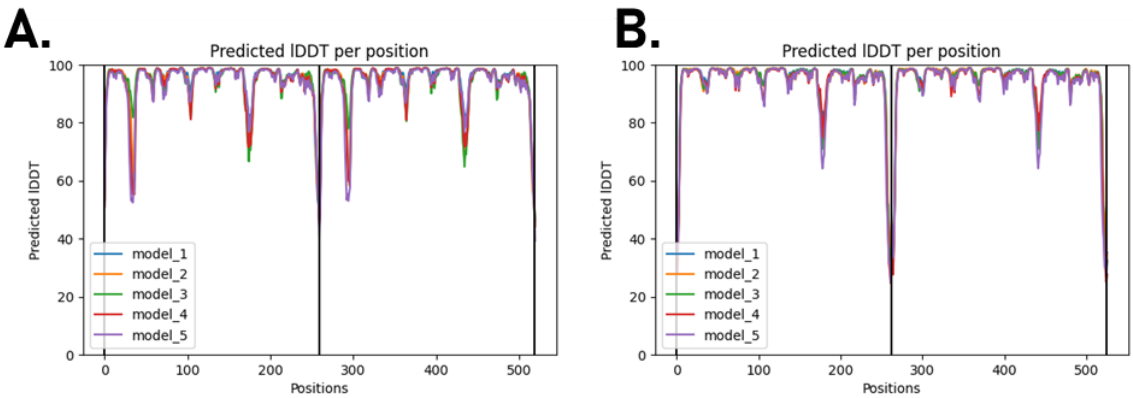
Green are eukaryotic TPIs. Blue are bacterial TPIs. Red are archaeal TPIs. Full list of TPI structures used in analysis can be found in appendix F.

Appendix H, Exemplar FTS data for c*Pum*TPI and p*Pum*TPI in presence or absence of DTT



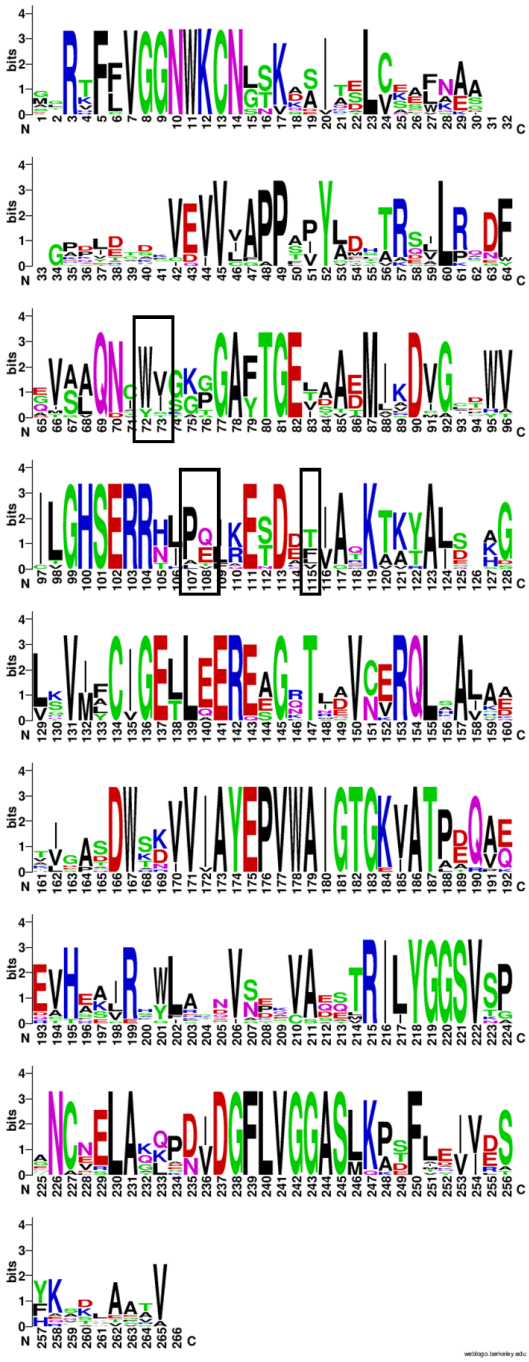
A. FTS melt curves of c*Pum*TPI and p*Pum*TPI. **B.** First derivative of data in **A.** Graphs were generated using StepOne™ software v2.1. Assay was carried out in the presence or absence of 1 mM DTT. Yellow-Green, p*Pum*TPI. Blue, p*Pum*TPI plus 1 mM DTT. Green c*Pum*TPI. Purple, c*Pum*TPI plus 1mM DTT. Orange, negative TPI control, Each condition was carried out in technical quadruplicate. T_m was estimated from the minima of each curve in **B.**

Appendix I, Predicted local distance difference test (pLDDT) plots for AlphaFold v2.1.0 structural models of c*Pum*TPI and p*Pum*TPI



A. pLDDT plot of c*Pum*TPI. **B.** pLDDT plot of p*Pum*TPI. Plots were generated using AlphaFold v2.1.0 (Jumper *et al.,* 2021). Interpretation of summary statistics was based on work from Tunyasuvunakool *et al.,* 2021. Model_1 from each plot was used to generate Figure 5.5. Vertical line at residue ~260 identifies separation of the monomers in each model.

Appendix J, Sequence logo of 15 Rhodophyta TPIs



15 TPIs were identified in the nine currency available Rhodophyta genomes. MSA of TPIs was generated using Clustal Omega (Sievers *et al.,* 2011) while sequence logo plot of the MSA was generated using WebLogo (Schneider and Stephens, 1990, Crooks *et al.,* 2004). Boxes identify residues discussed in text (Section 5.6.1). Note numbering in the sequence logo differs slightly from those used in text.

8. References

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., *et al.* (2011). The Phenix software for automated determination of macromolecular structures. Methods *55*, 94–106.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Alvarez, M., Zeelen, J.P., Mainfroid, V., Rentier-Delrue, F., Martial, J.A., Wyns, L., Wierenga, R.K., and Maes, D. (1998). Triosephosphate isomerase (TIM) of the psychrophilic bacterium *Vibrio marinus*: Kinetic and structural properties. J. Biol. Chem. *273*, 2199–2206.

Anderson, L.E. (1971). Chloroplast and cytoplasmic enzymes II. Pea leaf triose phosphate isomerases. BBA - Enzymol. *235*, 237–244.

Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. *37*, 420–423.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. *44*, W344–W350.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. *38*, W529–W533.

Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., *et al.* (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5Å resolution: Using amino acid sequence data. Nature *255*, 609–614.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D.S., and Milo, R. (2011). The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry *50*, 4402–4410.

Berney, C., and Pawlowski, J. (2006). A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. Proc. R. Soc. B Biol. Sci. *273*, 1867–1872.

Blat, Y. (2010). Non-competitive inhibition by active site binders. Chem. Biol. Drug Des. *75*, 535–540.

Brawley, S.H., Blouin, N.A., Ficko-Blean, E., Wheeler, G.L., Lohr, M., Goodson, H. V., Jenkins, J.W., Blaby-Haas, C.E., Helliwell, K.E., Chan, C.X., *et al.* (2017). Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). Proc. Natl. Acad. Sci. U. S. A. *114*, E6361–E6370.

Busch, F.A. (2020). Photorespiration in the context of rubisco biochemistry, CO2 diffusion and metabolism. Plant J. *101*, 919–939.

Cai, L., Arnold, B.J., Xi, Z., Khost, D.E., Patel, N., Hartmann, C.B., Manickam, S., Sasirat, S., Nikolov, L.A., Mathews, S., *et al.* (2021). Deeply altered genome architecture in the endoparasitic flowering plant *Sapria himalayana* Griff. (Rafflesiaceae). Curr. Biol. *31*, 1002-1011.e9.

Casteleijn, M.G., Alahuhta, M., Groebel, K., El-Sayed, I., Augustyns, K., Lambeir, A.M., Neubauer, P., and Wierenga, R.K. (2006). Functional role of the conserved active site proline of triosephosphate isomerase. Biochemistry *45*, 15483–15494.

Castro-Torres, E., Jimenez-Sandoval, P., Fernández-de Gortari, E., López-Castillo, M., Baruch-Torres, N., López-Hidalgo, M., Peralta-Castro, A., Díaz-Quezada, C., Sotelo-Mundo, R.R., Benitez-Cardoza, C.G., *et al.* (2018). Structural basis for the limited response to oxidative and thiol-conjugating agents by triosephosphate isomerase from the photosynthetic bacteria *Synechocystis*. Front. Mol. Biosci. *5*.

Castro-Torres, E., Jiménez-Sandoval, P., Romero-Romero, S., Fuentes-Pascacio, A., López-Castillo, L.M., Díaz-Quezada, C., Fernández-Velasco, D.A., Torres-Larios, A., and Brieba, L.G. (2019). Structural basis for the modulation of plant cytosolic triosephosphate isomerase activity by mimicry of redox-based modifications. Plant J. *99*, 950–964.

Chen, M., and Thelen, J.J. (2010). The plastid isoform of triose phosphate isomerase is required for the postgerminative transition from heterotrophic to autotrophic growth in *Arabidopsis*. Plant Cell *22*, 77–90.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A sequence logo generator. Genome Res. *14*, 1188–1190.

Cui, S., Kubota, T., Nishiyama, T., Juliane, K., Shigenobu, S., Shibata, T.F., Toyoda, A., Hasebe, M., Shirasu, K., and Yoshida, S. (2020). Ethylene signaling mediates host invasion by parasitic plants. Sci. Adv. *6*.

Dawson, J.H., Musselman, L.J., Wolswinkel, P., and Dörr, I. (1994). Biology and control of *Cuscuta*. Rev. Weed Sci. *6*, 265–317.

Dumont, S., Bykova, N. V., Pelletier, G., Dorion, S., and Rivoal, J. (2016). Cytosolic triosephosphate isomerase from *Arabidopsis thaliana* is reversibly modified by glutathione on cysteines 127 and 218. Front. Plant Sci. *7*, 1–16.

Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 2126–2132.

Feierabend, J., Kurzok, H.G., and Schmidt, M. (1990). Genetics and evolution of chloroplast isozymes of triosephosphate isomerase. Prog. Clin. Biol. Res. *344*, 665–682.

Ferla, M.P., Brewster, J.L., Hall, K.R., Evans, G.B., and Patrick, W.M. (2017). Primordial-like enzymes from bacteria with reduced genomes. Mol. Microbiol. *105*, 508–524.

Flamholz, A.I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y.M., Oltrogge, L.M., Alves, R., Savage, D., and Milo, R. (2019). Revisiting trade-offs between rubisco kinetic parameters. Biochemistry *58*, 3365–3376.

Flügel, F., Timm, S., Arrivault, S., Florian, A., Stitt, M., Fernie, A.R., and Bauwe, H. (2017). The photorespiratory metabolite 2-phosphoglycolate regulates photosynthesis and starch accumulation in *Arabidopsis*. Plant Cell *29*, 2537–2551.

Füssy, Z., Záhonová, K., Tomčala, A., Krajčovič, J., Yurchenko, V., Oborník, M., and Eliáš, M. (2020). The cryptic plastid of *Euglena longa* defines a new type of nonphotosynthetic plastid organelle. MSphere *5*.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In the Proteomics protocols handbook, pp. 571–607.

Go, M.K., Koudelka, A., Amyes, T.L., and Richard, J.P. (2010). Role of Lys-12 in catalysis by triosephosphate isomerase: A two-part substrate approach. Biochemistry *49*, 5377–5389.

Guiry, M.D., and Guiry, G.M. (2022). Phylum:Rhodophyta (Algaebase).

Henze, K., Schnarrenberger, C., Kellermann, J., and Martin, W. (1994). Chloroplast and cytosolic triosephosphate isomerases from spinach: purification, microsequencing and cDNA cloning of the chloroplast enzyme. Plant Mol. Biol. *26*, 1961–1973.

Herlihy, J.M., Maister, S.G., Albery, W.J., and Knowles, J.R. (1976). Energetics of triosephosphate isomerase: The fate of the 1(R)-3H label of tritiated dihydroxyacetone phosphate in the isomerase reaction. Biochemistry *15*, 5601–5607.

Holm, L., and Rosenström, P. (2010). Dali server: Conservation mapping in 3D. Nucleic Acids Res. *38*, W545–W549.

Jaffe, A.L., Castelle, C.J., Matheus Carnevali, P.B., Gribaldo, S., and Banfield, J.F. (2020). The rise of diversity in metabolic platforms across the Candidate Phyla Radiation. BMC Biol. *18*, 1–15.

Johnson, K.A. (2013). A century of enzyme kinetic analysis, 1913 to 2013. FEBS Lett. *587*, 2753–2766.

Johnson, M.P. (2016). Photosynthesis. Essays Biochem. *60*, 255–273.

Johnson, X., and Alric, J. (2013). Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: Metabolic constraints for carbon partitioning between oil and starch. Eukaryot. Cell *12*, 776–793.

Johnston, A.M., and Raven, J.A. (1987). The C4 -like characteristics of the intertidal macroalga *Ascophyllum nodosum* (L.) Le Jolis (Fucales, Phaeophyta). Phycologia *26*, 159–166.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., *et al.* (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

Kelly, G.J., and Latzko, E. (1976). Inhibition of spinach-leaf phosphofructokinase by 2-phosphoglycolate. FEBS Lett. *68*, 55–58.

Knowles, J.R. (1991). Enzyme catalysis: Not different, just better. Nature *350*, 121–124.

Knowles, J.R., and Albery, W.J. (1977). Perfection in enzyme catalysis: The energetics of triosephosphate isomerase. Acc. Chem. Res. *10*, 105–111.

Kovalenko, M., Rönnstrand, L., Heldin, C.H., Loubtchenkov, M., Gazit, A., Levitzki, A., and Böhmer, F.D. (1997). Phosphorylation site-specific inhibition of platelet-derived growth factor β-receptor autophosphorylation by the receptor blocking tyrphostin AG1296. Biochemistry *36*, 6260–6269.

Krämer, U. (2015). Planting molecular functions in an ecological context with *Arabidopsis thaliana*. Elife *4*.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. *33*, W299–W302.

Li, J., Weraduwage, S.M., Preiser, A.L., Tietz, S., Weise, S.E., Strand, D.D., Froehlich, J.E., Kramer, D.M., Hu, J., and Sharkey, T.D. (2019). A cytosolic bypass and G6P shunt in plants lacking peroxisomal hydroxypyruvate reductase. Plant Physiol. *180*, 783–792.

Liao, Q., Kulkarni, Y., Sengupta, U., Petrović, D., Mulholland, A.J., Van Der Kamp, M.W., Strodel, B., and Kamerlin, S.C.L. (2018). Loop motion in triosephosphate isomerase is not a simple open and shut case. J. Am. Chem. Soc. *140*, 15889–15903.

Lim, L., and McFadden, G.I. (2010). The evolution, metabolism and functions of the apicoplast. Philos. Trans. R. Soc. B Biol. Sci. *365*, 749–763.

Lin, M.T., and Hanson, M.R. (2018). Red algal rubisco fails to accumulate in transplastomic tobacco expressing *Griffithsia monilis* RbcL and RbcS genes. Plant Direct *2*.

Liu, S., Zeng, L.F., Wu, L., Yu, X., Xue, T., Gunawan, A.M., Long, Y.Q., and Zhang, Z.Y. (2008). Targeting inactive enzyme conformation: Aryl diketoacid derivatives as a new class of PTP1B inhibitors. J. Am. Chem. Soc. *130*, 17075–17084.

Lolis, E., and Petsko, G.A. (1990). Crystallographic analysis of the complex between triosephosphate Isomerase and 2-phosphoglycolate at 2.5-Å resolution: implications for catalysis. Biochemistry *29*, 6619–6625.

López-Castillo, L.M., Jiménez-Sandoval, P., Baruch-Torres, N., Trasviña-Arenas, C.H., Díaz-Quezada, C., Lara-González, S., Winkler, R., and Brieba, L.G. (2016). Structural basis for redox regulation of cytoplasmic and chloroplastic triosephosphate isomerases from *Arabidopsis thaliana*. Front. Plant Sci. *7*, 1–17.

Luong, L.T., and Mathot, K.J. (2019). Facultative parasites as evolutionary stepping-stones towards parasitic lifestyles. Biol. Lett. *15*.

Lyko, P., and Wicke, S. (2021). Genomic reconfiguration in parasitic plants involves considerable gene losses alongside global genome size inflation and gene births. Plant Physiol. *186*, 1412–1423.

McClain, A.M., and Sharkey, T.D. (2019). Triose phosphate utilization and beyond: From photosynthesis to end product synthesis. J. Exp. Bot. *70*, 1755–1766.

McPhillips, T.M., McPhillips, S.E., Chiu, H.J., Cohen, A.E., Deacon, A.M., Ellis, P.J., Garman, E., Gonzalez, A., Sauter, N.K., Phizackerley, R.P., *et al.* (2002). Blu-Ice and the distributed control system: Software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Radiat. *9*, 401–406.

Mercado, J.M., Niell, F.X., and Figueroa, F.L. (1997). Regulation of the mechanism for HCO3- use by the inorganic carbon level in *Porphyra leucosticta* Thur. in Le Jolis (Rhodophyta). Planta *201*, 319–325.

Minhas, A., Kaur, B., and Kaur, J. (2020). Genomics of algae: Its challenges and applications. In Pan-Genomics: Applications, Challenges, and Future Prospects, pp. 261–283.

Moriyama, T., Sakurai, K., Sekine, K., and Sato, N. (2014). Subcellular distribution of central carbohydrate metabolism pathways in the red alga *Cyanidioschyzon merolae*. Planta *240*, 585–598.

Nickbarg, E.B., and Knowles, J.R. (1988). Triosephosphate isomerase: Energetics of the reaction catalyzed by the yeast enzyme expressed in *Escherichia coli*. Biochemistry *27*, 5939–5947.

Nickrent, D.L. (2020). Parasitic angiosperms: How often and how many? Taxon *69*, 5–27.

Pelser, P.B., Nickrent, D.L., van Ee, B.W., and Barcelona, J.F. (2019). A phylogenetic and biogeographic study of *Rafflesia* (Rafflesiaceae) in the Philippines: Limited dispersal and high island endemism. Mol. Phylogenet. Evol. *139*, 106555.

Plaxton, W.C. (1996). The organization and regulation of plant glycolysis. Annu. Rev. Plant Physiol. Plant Mol. Biol. *47*, 185–214.

Ralph, S.A., van Dooren, G.G., Waller, R.F., Crawford, M.J., Fraunholz, M.J., Foth, B.J., Tonkin, C.J., Roos, D.S., and McFadden, G.I. (2004). Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nat. Rev. Microbiol. *2*, 203–216.

Read, B.A. (1994). High substrate specificity factor ribulose bisphosphate carboxylase/oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial rubisco containing “algal” residue modifications. Arch. Biochem. Biophys. *312*, 210–218.

Richard, J.P. (2008). Restoring a metabolic pathway. ACS Chem. Biol. *3*, 605–607.

Rozovsky, S., and McDermott, A.E. (2001). The time scale of the catalytic loop motion in triosephosphate isomerase. J. Mol. Biol. *310*, 259–270.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: A new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.

Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., McVeigh, R., O'Neill, K., Robbertse, B., *et al.* (2020). NCBI Taxonomy: A comprehensive update on curation, resources and tools. Database *2020*.

Schreier, T.B., and Hibberd, J.M. (2019). Variations in the Calvin-Benson cycle: Selection pressures and optimization? J. Exp. Bot. *70*, 1697–1701.

Sharkey, T.D. (1988). Estimating the rate of photorespiration in leaves. Physiol. Plant. *73*, 147–152.

Sharkey, T.D. (2019). Discovery of the canonical Calvin–Benson cycle. Photosynth. Res. *140*, 235–252.

Sharkey, T.D., and Weise, S.E. (2012). Autotrophic carbon dioxide fixation. In Photosynthesis, pp. 651–674.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*.

Stein, N. (2008). CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. *41*, 641–643.

Sun, G.L., Shen, W., and Wen, J.F. (2008). Triosephosphate isomerase genes in two trophic modes of euglenoids (Euglenophyceae) and their phylogenetic analysis. J. Eukaryot. Microbiol. *55,* 170–177.

Sun, G., Xu, Y., Liu, H., Sun, T., Zhang, J., Hettenhausen, C., Shen, G., Qi, J., Qin, Y., Li, J., *et al.* (2018). Large-scale gene losses underlie the genome evolution of parasitic plant *Cuscuta australis*. Nat. Commun. *9*, 4–11.

Thorogood, C.J., Teixeira-Costa, L., Ceccantini, G., Davis, C., and Hiscock, S.J. (2021). Endoparasitic plants and fungi show evolutionary convergence across phylogenetic divisions. New Phytol. *232*, 1159–1167.

Touloupakis, E., Rontogiannis, G., Silva Benavides, A.M., Cicchi, B., Ghanotakis, D.F., and Torzillo, G. (2016). Hydrogen production by immobilized *Synechocystis* sp. PCC 6803. Int. J. Hydrogen Energy *41*, 15181–15186.

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., *et al.* (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590–596.

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

Vickers, C.J., Fraga, D., and Patrick, W.M. (2021). Quantifying the taxonomic bias in enzymology. Protein Sci. 1–8.

Vogel, A., Schwacke, R., Denton, A.K., Usadel, B., Hollmann, J., Fischer, K., Bolger, A., Schmidt, M.H.W., Bolger, M.E., Gundlach, H., *et al.* (2018). Footprints of parasitism in the genome of the parasitic flowering plant *Cuscuta campestris*. Nat. Commun. *9*, 1–11.

von Heijne, G., Stepphun, J., and Herrmann, R.G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. Eur. J. Biochem. *180*, 535–545.

Walker, M.C., Kurumbail, R.G., Kiefer, J.R., Moreland, K.T., Koboldt, C.M., Isakson, P.C., Seibert, K., and Gierse, J.K. (2001). A three-step kinetic mechanism for selective inhibition of cyclo-oxygenase-2 by diarylheterocyclic inhibitors. Biochem. J. *357*, 709–718.

Wang, D., Yu, X., Xu, K., Bi, G., Cao, M., Zelzion, E., Fu, C., Sun, P., Liu, Y., Kong, F., *et al.* (2020). *Pyropia yezoensis* genome reveals diverse mechanisms of carbon acquisition in the intertidal environment. Nat. Commun. *11*, 1–11.

Wertheim, J.O., Murrell, B., Smith, M.D., Pond, S.L.K., and Scheffler, K. (2015). RELAX: Detecting relaxed selection in a phylogenetic framework. Mol. Biol. Evol. *32*, 820–832.

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

Wierenga, R.K., Kapetaniou, E.G., and Venkatesan, R. (2010). Triosephosphate isomerase: A highly evolved biocatalyst. Cell. Mol. Life Sci. *67*, 3961–3982.

Wierenga, R.K., Noble, M.E.M., Postma, J.P.M., Groendijk, H., Kalk, K.H., Hol, W.G.J., and Opperdoes, F.R. (1991). The crystal structure of the “open” and the “closed” conformation of the flexible loop of trypanosomal triosephosphate isomerase. Proteins Struct. Funct. Bioinforma. *10*, 33–49.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 235–242.

Woehle, C., Dagan, T., Landan, G., Vardi, A., and Rosenwasser, S. (2017). Expansion of the redox-sensitive proteome coincides with the plastid endosymbiosis. Nat. Plants *3*, 17066.

Yoshida, S., Cui, S., Ichihashi, Y., and Shirasu, K. (2016). The haustorium, a specialized invasive organ in parasitic plants. Annu. Rev. Plant Biol. *67*, 643–667.

Yoshida, S., Kim, S., Wafula, E.K., Tanskanen, J., Kim, Y.M., Honaas, L., Yang, Z., Spallek, T., Conn, C.E., Ichihashi, Y., *et al.* (2019). Genome sequence of *Striga asiatica* provides insight into the evolution of plant parasitism. Curr. Biol. *29*, 3041–3052.

Zaffagnini, M., Michelet, L., Sciabolini, C., Giacinto, N. Di, Morisse, S., Marchand, C.H., Trost, P., Fermani, S., and Lemaire, S.D. (2014). High-resolution crystal structure and redox properties of chloroplastic triosephosphate isomerase from *Chlamydomonas reinhardtii*. Mol. Plant *7*, 101–120.

Zakryś, B., Milanowski, R., and Karnkowska, A. (2017). Evolutionary origin of Euglena. In Advances in Experimental Medicine and Biology, pp. 3–17.

Zhang, B., Xie, X., Liu, X., He, L., Sun, Y., and Wang, G. (2020). The carbonate concentration mechanism of *Pyropia yezoensis* (Rhodophyta): evidence from transcriptomics and biochemical data. BMC Plant Biol. *20*, 1–12.

Zhang, Z., Sugio, S., Komives, E.A., Liu, K.D., Knowles, J.R., Petsko, G.A., and Ringe, D. (1994). Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-Å resolution. Biochemistry *33*, 2830–2837.