DETERMINATION OF THE ROLE OF *Synechococcus* RUBISCO RESIDUE VAL-425 IN CHIMERAS

By

CLARA WONG CHIA CI

A project report submitted to the Department of Chemical Science

Faculty of Science

Universiti Tunku Abdul Rahman

in partial fulfilment of the requirements for the degree of

Bachelor of Science (Hons) Biochemistry

September 2023

ABSTRACT

DETERMINATION OF THE ROLE OF *Synechococcus* RUBISCO RESIDUE VAL-425 IN CHIMERAS

CLARA WONG CHIA CI

Ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39), Rubisco is one of the most extensively studied enzymes due to its abundance on Earth and its significance in the food supply chain. As the rate-limiting enzyme in the Calvin cycle, the catalytic inefficiency of Rubisco has led to an undesirable side reaction known as photorespiration that compromises the carboxylase activity of CO_2 fixation, thus, making it the primary target for bioengineering to improve its catalytic efficiency. Innovative approaches were made by researchers to improve the carboxylation capacity of Rubisco which nonetheless was hampered by the unsuccessful heterologous formation of functional Rubisco enzyme, plausibly due to incompatible assembly pathways. Previous findings by Koay, Wong and Lim (2016) were extended in this study whereby the chimeric mutant bearing mutated residue 425 (Thr-425) is used as the target for site-directed mutagenesis (SDM) that reverts the Thr mutation to Val residue. The current study explores the potential role of Val-425 in the large subunit (RbcL) of *Synechococcus* Rubisco that is previously hypothesized to be critical for the chaperone-mediated assembly of the holoenzyme. SDM had successfully restored Val residue at position 425 of RbcL which was validated by restriction

digestion of PstI enzyme. Expression for both Rubisco subunits (RbcL and RbcS) upon SDM was detected on SDS-PAGE at 55 kDa and below 15 kDa, respectively. Furthermore, Val-425 is also found to be important in conferring stability in the native conformation of *Synechococcus* RbcL and its contribution to greater hydrophobicity suggests its associated role in chaperonin-binding.

Keywords: Rubisco, holoenzyme assembly, *Synechococcus elongatus*, *Escherichia coli*, site-directed mutagenesis, bioinformatics

ACKNOWLEDGMENT

I would like to take this opportunity to thank those who have assisted and encouraged me throughout making this thesis a reality. Foremost, I would like to express my deepest appreciation to my supervisor, Dr. Lim Boon Hoe for the continuous support and guidance throughout my final year project (FYP). His insightful guidance and constructive comments have been a great influence on the direction and quality of this study. It has been a pleasure to complete my FYP under his guidance and learn from his expertise.

Besides, I would like to extend my gratitude to the postgraduate, Mr. Steven Khoo Kah Liang for his patience and kindness in sharing his knowledge and skills during the FYP. His willingness to engage in discussions and to clarify my doubts has kept me motivated even in the most challenging times. I would also like to express my heartfelt gratitude to my bench mate, Ms. Ng Yi Wen for her constant comradeship and emotional support during hardships. The shared moments of triumph during collaborative problem-solving have been incredibly valuable.

Furthermore, my sincere thanks to the lab officers, Mr. Ooh Keng Fei and Ms. Wong Wei Seng from the Faculty of Science, who have ensured that I have access to all the laboratory equipment and chemicals necessary to complete my FYP. Last but not least, I would like to thank my family and friends for their care and constant encouragement in making this thesis possible.

DECLARATION

I hereby declare that this final year project is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

(CLARA WONG CHIA CI)

APPROVAL SHEET

This final year project entitled "**DETERMINATION OF THE ROLE OF *Synechococcus* RUBISCO RESIDUE VAL-425 IN CHIMERAS**" was prepared by CLARA WONG CHIA CI and submitted as partial fulfilment of the requirements for the degree of Bachelor of Science (Hons) Biochemistry at Universiti Tunku Abdul Rahman.

Approved by:

Lim Boon Hoe

(Dr. Lim Boon Hoe) Date: 2 Sept 2023

Supervisor

Department of Chemical Science

Faculty of Science

Universiti Tunku Abdul Rahman

FACULTY OF SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

Date: 02-09-2023

PERMISSION SHEET

It is hereby certified that **CLARA WONG CHIA CI** (ID No: **20ADB06159**) has completed this final year project entitled "**DETERMINATION OF THE ROLE OF *Synechococcus* RUBISCO RESIDUE VAL-425 IN CHIMERAS**" under the supervision of DR. LIM BOON HOE from the Department of Chemical Science, Faculty of Science.

I hereby give permission to the University to upload the softcopy of my final year project in PDF format into the UTAR Institutional Repository, which may be made accessible to the UTAR community and the public.

Yours truly,

(CLARA WONG CHIA CI)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Amp	Ampicillin
ApE	A plasmid Editor
APS	Ammonium persulfate
ATP	Adenosine triphosphate
β	Beta
BCL-2	B-cell lymphoma 2
BioEdit	Biological sequence editor
Bsd2	Bundle sheath defective 2
CBB	Calvin-Benson-Bassham
CD	Circular Dichroism
CES	Control by epistasy of synthesis
CO_2	Carbon dioxide
Cpn10	Chaperonin 10
Cpn20	Chaperonin 20
Cpn60α	Chaperonin 60 alpha
Cpn60β	Chaperonin 60 beta
ΔΔG/DDG	Free energy change
DF	Dilution factor

DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ExPASy	Expert Protein Analysis System
GMQE	Global Model Quality Estimation
GRAVY	Grand average of hydropathy
G3P	Glyceraldehyde-3-phosphate
IgG	Immunoglobulin G
IPTG	Isopropyl-β-D-thiogalactopyranoside
JSON	JavaScript Object Notation
K_M	Michaelis-Menten constant
LB	Luria Bertani
LSU	Large subunit of Rubisco
Mg^{2+}	Magnesium (II) ion
mmCIF	Macromolecular Crystallographic Information File
MUSCLE	Multiple Sequence Comparison by Log-Expectation
NADH	Nicotinamide adenine dinucleotide
Native-PAGE	Native polyacrylamide gel electrophoresis
O_2	Oxygen
ORF	Open reading frame

PCR	Polymerase chain reaction
PDB	Protein Data Bank
pE_{M1}	pE mutant of first mutagenesis
pE_{M2}	pE mutant of second mutagenesis (Val425)
QMEAN/QMEANDisCo	Qualitative Model Energy ANalysis
Raf1	Rubisco accumulation factor 1
Raf2	Rubisco accumulation factor 2
RbcL	Rubisco large subunit
RbcS	Rubisco small subunit
RbcX	Rubisco chaperone X
RLP	Rubisco-like protein
RMSD	Root-mean-square-deviation
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	D-ribulose-1,5-bisphosphate
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SMTL	SWISS-MODEL Template Library
SSU	Small subunit of Rubisco
SVM	Support vector machine

TAE	Tris acetate EDTA
TEMED	N,N,N', N'-Tetramethylethylenediamine
Tet	Tetracycline
θ	Theta
V_{max}	Maximum velocity
WT	Wildtype
2PG	2-Phosphoglycolate
3D	Three-dimensional
3PGA	3-Phosphoglycerate

CHAPTER 1

INTRODUCTION

Rubisco or ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39) is known to be the most ubiquitous enzyme found in the biosphere. Its role in CO_2 assimilation is pivotal for photosynthesis which takes place in most autotrophs as it catalyzes the rate-limiting step of the Calvin-Benson-Bassham (CBB) pathway. There are four naturally occurring forms (I, II, III, IV) of Rubisco possessing distinct holoenzyme structures and Form I is reported as the most abundant group among autotrophic organisms (Andersson and Backlund, 2008; Tabita et al., 2008). Most prokaryotes such as cyanobacteria and proteobacteria, and eukaryotes such as algae and higher plants express Form I Rubisco while Form II is commonly observed in eukaryotic dinoflagellates and non-sulfur phototropic bacteria. The third form of Rubisco is found exclusively in archaea where, unlike the two aforementioned forms, it utilizes the pentose-bisphosphate pathway for D-ribulose-1,5-bisphosphate (RuBP) carboxylation (Frolov et al., 2019). Diverse forms of Rubisco quaternary structures are found in the archaeal family, comprising different numbers of large subunits, ranging from 2 to 10 (Andersson and Backlund, 2008). Besides, Form IV Rubisco exists only in certain *Bacilli* sp. are also known as Rubisco-like proteins or RLPs, dissimilar to conventional Rubisco, it is involved in sulfur metabolism instead of RuBP-dependent CO_2 fixation. **Figure 1.1** shows the structural overview of the four Rubisco forms.

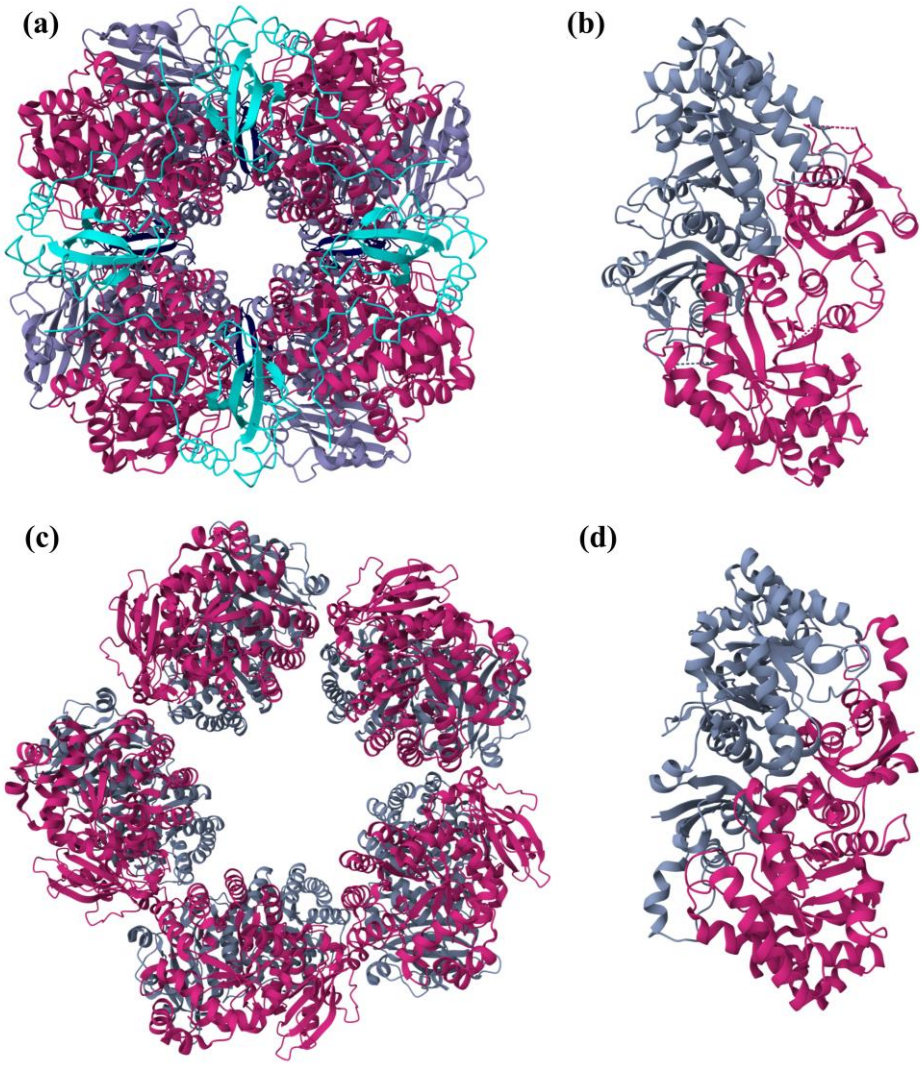


Figure 1.1: Structural representation of different forms of Rubisco. The large subunit dimers (L_2) are the building block for all forms, shown in violet and pink. Form I Rubisco (a) is comprised of four L_2 dimers decorated with eight small subunits in (cyan and dark blue), surrounding the L_8 core (PDB ID: 1RBL) (Newman, Brändén and Jones, 1993). Form II (PDB ID: 2RUS) and IV (PDB ID: 2ZVI) only consist of one L_2 dimer while Form III (PDB ID: 3KDN) is found in an $(L_2)_5$ arrangement (Lundqvist and Schneider, 1991; Tamura et al., 2009; Nishitani et al., 2010).

The bifunctional characteristic of Rubisco that catalyzes both carboxylation and oxygenation of RuBP results in the catalytic inefficiency of the enzyme in photosynthesis. Oxygenation to carboxylation ratio is heavily dependent on environmental conditions, particularly the temperature and CO_2: O_2 ratio (Heldt

and Piechulla, 2011). Despite Rubisco was reported by van Lun et al (2014) to have a greater preference for CO_2 over O_2 at equal concentrations of both gases, the higher composition of O_2 in the Earth's atmosphere relative to CO_2 downplays the substrate specificity in Rubisco. In the presence of RuBP oxygenation, a toxic compound named glycolate is formed which could lead to adverse effects if accumulates in the chloroplasts. Detoxification of glycolate via photorespiration leads to an approximate loss of 25% of the photosynthetically fixed carbon as CO_2 and the rest of 75% is recovered as 3-phosphoglycerate (3PGA) that is recycled into the CBB cycle for RuBP regeneration (Walker et al., 2016; Busch, 2020). The wasteful process of photorespiration takes place inevitably in plants, specifically C_3 plants as a protective mechanism against strong light intensity and in response to osmotic stress; nonetheless, with an implication of reduced crop yield which can be detrimental to agricultural productivity, raising food security concerns (Walker et al., 2016; Wada et al., 2020). The global population is expected to reach 9.7 billion by 2050 as reported by the United Nations Department of Economic and Social Affairs, Population Division (2022). However, the present rate of agricultural productivity is insufficient to accommodate the demand and at least a 60% increase in food production is necessary (Muhie, 2022). This has encouraged researchers to explore means to improve crop production instead of relying completely on the unsustainable expansion of agricultural lands.

Rubisco has been the main emphasis in genetic engineering and protein studies due to its abundance and significance in the food supply chain which makes it the major target for enhancement of crop productivity and biofuel production.

Initiatives were made by researchers to improve photosynthesis including Rubisco engineering, engineering of C_4 photosynthesis in C_3 plants, reconstitution of photorespiration pathway, and photorespiratory bypass (Mackinder, 2018; Kubis and Bar-Even, 2019). For instance, the bacteria Rubisco generally exhibits greater turnover rates relative to higher plants which display lower turnover rates with higher specificity values for substrate (Cummins, Kannappan and Gready, 2018). By combining the high turnover capability of bacteria Rubisco with the high specificity of plant Rubisco, an enhanced version of the Rubisco enzyme can be generated. However, these approaches have only met limited success due to the complexity of the folding and assembly pathway of the Form I holoenzyme which often fails to achieve in phylogenetically distant hosts, mainly owing to the incompatibility in the chaperone-mediated folding pathway. This hinders the progress of Rubisco engineering in crop improvement as a repercussion of the minimal diversity of Rubisco genes available for study.

Previous studies conducted by Koay et al. (2016) and Ong et al. (2023) addressed the aforementioned issue by the sequential swapping of RbcL coding sequence between *Synechococcus elongatus* PCC6301 and *Chlamydomonas reinhardtii* that resulted in ten chimeric mutants. This study hence extends from previous findings to assess one of the chimeric mutants that failed to assemble. The residue 425 in the chimera containing mutated segments of 398 to 447 was hypothesized to participate in the holoenzyme formation of Rubisco, particularly as a chaperone binding site. By reverting the mutation at residue 425 from Thr to Val via site-directed mutagenesis, we aimed to identify the potential role of

residue 425 in the large subunit of *Synechoccocus* Rubisco that ensures holoenzyme assembly in a phylogenetically distant host. *In vitro* and *in silico* approaches employed in this study complement each other in assessing the possibility of protein assembly, particularly via the protein structure and stability.

The objectives of this study include

i. To perform site-directed mutagenesis that changes Thr-425 to Val-425 in the large subunit of Rubisco.
ii. To evaluate the expression of Rubisco subunits in *E. coli* by SDS-PAGE.
iii. To assess the effect of Val-425 on the structural stability of Rubisco via I-Mutant2.0.
iv. To determine the role of Val-425 in the folding and assembly of *Synechococcus* Rubisco holoenzyme via *in silico* analysis.

CHAPTER 2

LITERATURE REVIEW

2.1 Ribulose-1,5-bisphosphate Carboxylase/Oxygenase (Rubisco)

2.1.1 Form I Rubisco Structure

Form I Rubisco as the most prominent carbon-fixing enzyme makes it a significant machinery in driving the carbon cycle, ultimately sustaining life (Banda et al., 2020). The fundamental structural anatomy of all forms of Rubisco comprised of a catalytically active large subunit dimer. The distinctive feature of Form I Rubisco include additional small subunits that are absent in others. Four large subunit dimers assemble into an octameric core (L_8) of RbcL that is capped by eight small subunits at each end, creating a hexadecameric (L_8S_8) holoenzyme. The large subunit of Rubisco is composed of two domains, namely the N-terminal domain containing approximately 150 residues and the C-terminal TIM barrel which is comprised of about 300 residues (Newman and Branden, 1993). The active site of Rubisco is formed from the specific interaction between two large subunits in a manner that gives rise to two active sites per dimer (Tabita et al., 2008). Precisely, the active site is located at the interface between the N-terminal domain of a subunit and the C-terminal domain of the adjacent subunit. The combination of the N-terminal domain in the second subunit and the barrel domain of the first subunit generates another catalytic site of Rubisco. A total of eight active sites facing outward are generated via the protein-protein interactions between the subunits in specific orientations. The small subunit consists of a single domain made up of a four-stranded antiparallel β-sheet with

two helices. Despite the small subunit is not directly involved in the catalytic reaction, it is experimentally validated that it might demonstrate a structural role in stabilizing the active site of Rubisco (Andrews, 1988). Besides, it is also responsible for ensuring the 60s loop of RbcL in order that results in the active conformation of Rubisco (Bracher et al., 2011). In addition to the structural function in assembling and concentrating the large subunits, the small subunit may also act as a CO_2 reservoir that enhances specificity for CO_2 (Andersson and Backlund, 2008; van Lun et al., 2014). The conformational transition of Rubisco from open state to closed state involves the movements of loops and the C-terminal strand which encloses the active site from access by the solvent during catalysis (van Lun et al., 2014).

2.1.2 Catalytic Mechanism

Rubisco demonstrates catalytic promiscuity via its imperfect discrimination between CO_2 and O_2 where it undergoes both carboxylation and oxygenation of RuBP at its active site. Atmospheric CO_2 fixation in carboxylase of Rubisco generates two molecules of 3-phosphoglycerate (3PGA) which is later reduced to glyceraldehyde-3-phosphate (G3P) in the CBB cycle for carbohydrate synthesis. The promiscuous enzyme follows a counterproductive oxygenation reaction in the presence of O_2 accumulation during sunny weather where one molecule of 2-phosphoglycolate (2PG) and 3PGA are released. Since 2PG is futile in the CBB cycle and its accumulation may cause toxicity to enzymes in the photosynthetic pathway, it has to be converted to 3PGA via additional processes known as the glycolate pathway which is associated with unnecessary

loss of carbon and ATP (Shi and Bloom, 2021). Fixation of CO_2 or O_2 begins with the activation of Rubisco via carbamylation of conserved Lys201 residue and stabilization by Mg^{2+} in the active site, forming a ternary complex (Sharwood, 2017; Tommasi, 2021). RuBP then binds tightly to the activated Rubisco which enables carboxylation or oxygenation reactions to take place. It is noteworthy that the CO_2 molecule that participates in Rubisco activation is independent of the subsequent carboxylation reaction. A more comprehensive mechanism for Rubisco activation was elucidated by Stec (2012) which suggested the presence of separate binding sites for activation and catalysis. This involves the conformational change in the side chain of His335 that creates the activation binding site followed by the relocation of the Mg^{2+} stabilized complex to a different site for the re-establishment of initial conformation in the His335 side chain. The conformational transition from an open state to a closed state is observed as the binding of RuBP occurs with the occlusion of the TIM barrel entrance and the active site by the cooperative movement of the N-terminal domain, the loop 6 and the C-terminal strand, thus making it solvent-free.

Furthermore, carboxylation of RuBP proceeds in a five-step reaction which also involves Bronsted acid-base catalytic mechanism where some active site residues (Lys, His, Glu) serve as Lewis bases for protein abstraction whereas Lys175 and Lys334 function to stabilize the negatively charged anions via hydrogen bonds (Andersson, 2008; Tommasi, 2021). Conversion of RuBP follows an overall pathway of enolization, carboxylation, hydration, scission C_2-C_3 bond and finally C_2 protonation. A gem-diolate intermediate formed during the hydration step is broken down into a 3PGA and an aci-acid intermediate. The

latter is stereospecifically protonated to yield another 3PGA molecule. Oxygenation of enolized RuBP, 2,3-enediol leads to the formation of peroxo intermediate that is split into 3PGA and 2PG as end products (Tcherkez, 2016). Bathellier et al. (2020) reported that the addition of O_2 to 2,3-enediol occurs through the transfer of a single electron that generates short-lived superoxide radicals that assemble into the peroxide intermediate. The mechanism greatly depends on the dielectric environment in the active site which also clarifies the impracticality to eradicate photorespiration in Rubisco via bioengineering.

2.1.3 Rubisco Assembly

The proper folding and assembly of Rubisco holoenzyme in a heterologous host is essential to prevent protein aggregation, thus ensuring the formation of functional enzymes for successful Rubisco engineering studies. The construction of Rubisco holoenzyme is a chaperone-dependent process that requires a stream of molecular chaperones. It is remarkable that dissimilar to cyanobacterial Rubisco that readily reconstitute in the presence of bacterial chaperones and chaperonins in *E. coli* or *ex vivo*, its eukaryotic homologs have been refractory on the aforementioned event, considering that all Form I Rubisco exhibit high homology. The phylogenetic relationship of Rubisco-assembly factors between green-type Rubisco was presented by Gruber and Feiz (2018) where auxiliary factors (Bsd2, Raf2) were found in the assembly of eukaryotic group of Form I RbcL subunits in addition to the fundamental chaperones and chaperonins (GroEL-GroES, Raf1, RbcX) reported for cyanobacterial Rubisco holoenzyme formation. In cyanobacteria, the newly synthesized RbcL polypeptide chain is

folded by GroEL-GroES chaperonin followed by the binding of chaperones Rubisco accumulation factor 1 (Raf1) and RbcX that can either individually assist or function in coordination for the formation of the L8 octameric core (Li et al., 2022). Generally, the simultaneous binding of Raf1 and RbcX which results in a ternary complex is the preferred mode of assembly as it favours the displacement of chaperones by RbcS in the final phase of holoenzyme formation. Despite the two aforementioned chaperones displaying functional similarity in *Synechococcus* sp., the distinct interaction sites on RbcL eliminate the preconception of RbcX as a functionally redundant factor for Rubisco biogenesis (Hauser et al., 2015). Unlike other TIM barrel proteins, the partially folded RbcL of the Form I Rubisco is relatively structurally unstable and prone to reassociate with the chaperonin which makes the chaperones indispensable for stabilizing the dimeric L8 (Hauser et al., 2015; Feiz and Gruber, 2018). This is consistent with the findings by Huang et al. (2020) where Raf1 is responsible for mediating dimer-dimer interactions in RbcL. Besides, Raf1 ensures the folded RbcL remain in an assembly-competent state by having a single Raf1 molecule associated with one RbcL dimer, resulting in an $RbcL_2$-Raf1 complex and subsequently at higher oligomeric ($RbcL_8$-$Raf1_4$) state. Given the larger RbcL-Raf1 interface relative to that of RbcL-RbcX, Li et al. (2022) suggested that Raf1 may play a central role in chaperone-mediated Rubisco assembly while RbcX serves as reinforcement at the later stage. Finally, dissociation of chaperones from the RbcL occurs following the docking of RbcS subunits, thus, giving rise to the L8S8 hexadecameric holoenzyme.

The eukaryotic large subunit of Rubisco is generally chloroplast-encoded while the small subunit is nuclear-encoded in higher plants and green algae which later folds in the chloroplast (Gruber and Feiz, 2018). The chaperones RbcX and Raf1 which are of cyanobacterial origin are conserved in higher plants and green algae (Bracher et al., 2015). Additional ancillary proteins such as bundle sheath defective 2 (Bsd2) and Raf2 were discovered to participate in eukaryotic-specific Rubisco assembly which supported the fact that biogenesis eukaryotic Rubisco follows a highly orchestrated pathway with the facilitation of different Rubisco-specific assembly factors (Fristedt et al., 2018; Conlan et al., 2019). The chloroplast chaperonin complex responsible for the folding of nascent RbcL is a GroEL-GroES ortholog comprised of Cpn60α and Cpn60β bound by cofactors Cpn20 and Cpn10 (Aigner et al., 2017). The overall folding and assembly pathway of eukaryotic Rubisco proceeds in a similar manner as that of cyanobacterial Rubisco with the exception that Raf1 is supplanted by Bsd2 as part of the end-state intermediate ($RbcL_8$-$Bsd2_8$) before RbcS docking (Aigner et al., 2017). Bsd2 was reported by Conlan et al. (2018) to function as an end-stage assembly factor in tobacco plants by displacing Raf1 after the association of $RbcL_8$-$Raf1_4$ units, yielding a more stable complex. Besides, Raf2 found in *Arabidopsis thaliana* was reported to support RbcS delivery to the RbcL oligomer for assembly into Rubisco holoenzyme (Fristedt et al., 2018). Furthermore, the synthesis of RbcL in the chloroplast is regulated via the control by epistasy of synthesis (CES) (Wietrzynski et al., 2021). The translation rate of chloroplast-encoded proteins is adjusted based on the availability of their assembly partners which in this case, the production of the RbcL is restricted by

the amount of RbcS present in the cell, which ultimately balances the formation of both subunits.

2.2 Heterologous Expression and Assembly of Rubisco

The distinctive degree of complexity in Rubisco folding and assembly mechanism across organisms as reviewed in **Section 2.1** has restricted the diversity of Rubisco isoforms that can employed in heterologous expression systems, especially the *E. coli*. This section thus reviews previous studies of Rubisco expression and assembly in phylogenetically distant organisms. Whitney et al. (2001) attempted to express Rubisco of non-green alga origin in tobacco plants. Rubisco subunits of both the rhodophyte *Galdiera sulphuraria* and the diatom *Phaeodactylum trocornutum* are conveniently expressed with no L_8S_8 complex observed, indicating unsuccessful folding. Another attempt was made by Lin and Hanson (2018) who successfully expressed that Rubisco originated from red alga *Griffithsia monilis* in plants but failed to produce functional Rubisco in the chloroplast, confirmed by the presence of RbcL detected as part of the insoluble fraction on immunoblot and SDS-PAGE. The non-assembly was explained as a consequence of the incompatibility of the holoenzyme folding pathway between tobacco and *Griffithsia monilis* as successful assembly of heterologous Rubisco is frequently constrained among closely associated plant species and simple bacteria (Lin and Hanson, 2018). Besides, given that cyanobacterium-derived Rubisco readily reconstitutes in *E. coli* due to its simpler mechanism, approximately 98% of the *Synechoccocus* RbcL expressed exist as misfolded aggregates in the bacterium system (Mueller-

Cajar and Whitney, 2008; Wilson et al., 2018). In contrast, Lin et al. (2014) were able to produce functional Rubisco by transforming the RbcL-knockout tobacco with cyanobacterial RbcL and RbcS coding genes along with their assembly factors. Despite successful assembly, the transgenic plants only survived under an elevated CO_2 atmosphere and grew at a slower rate than the wildtype (WT) tobacco which may be associated with the lower concentration of functional Rubisco synthesized. Occhialini et al. (2016) advanced their findings years later and proved that the assembly of cyanobacterial Rubisco in the chloroplast without the presence of the prokaryotic assembly factors is possible which is in agreement with the simpler folding and assembly requirements in cyanobacteria. Contrary to previous studies, the transgenic plants were able to achieve a near-WT rate of photosynthesis under raised CO_2 conditions. Expression of functional eukaryotic Rubisco in *E. coli* has remained a major challenge for directed evolution studies before Aigner et al. (2017) successfully expressed functional *Arabidopsis* Rubisco in *E. coli* with the co-expression of all the five essential auxiliary factors. This provided an insight into the chaperone-mediated folding pathway of eukaryotic Rubisco that was previously incomprehensible. Remarkably, an alternative approach made by Wilson et al. (2019) in expressing plant Rubisco in *E. coli* not only capable of synthesizing properly folded holoenzyme but also raised enzyme yield. This was achieved by the simultaneous expression of corresponding chaperones and Rubisco for folding-assembly efficiency and RbcS C-terminal tag for RbcS stability after formation.

2.3 Site Directed Mutagenesis of Rubisco Studies

Site directed mutagenesis (SDM) is a technique frequently used in the study of study-function relationships of proteins by modifying the nucleotide sequence of a particular location in the gene (Walker and King, 2023). Nakano et al. (2010) had employed the SDM strategy to elucidate the role of Lys122 in *Rhodospirillum* and *Bacillus subtilis* Rubisco active site. The K122M/E mutations in *Rhodospirillum* resulted in inactivation of catalytic function without disrupting the structure while these mutations disrupted the stable conformations of *B. subtilis* Rubisco, precluding the formation of catalytic dimer. Similar application was demonstrated by Rasineni, Loh and Lim (2017) in determining the significance of posttranslationally modified residues at four distinct locations, namely Pro-104, Pro-151, Cys-256, and Cys-369 within the large subunit of *Chlamydomonas* Rubisco via alanine substitution. Substitution of the Pro-104 and Cys-256 residues with Ala abolished Rubisco catalysis, suggesting their roles in regulation of catalysis. Furthermore, four critical amino acid residues responsible for thermostability and activity of *Synechococcus* Rubisco were identified in the study by Miller, McGuirl and Carvey (2013) using SDM approach. Substitution of these four residues altogether raised both the thermal stability and catalytic activity of the Rubisco. Kreel and Tabita (2015) had also utilized the SDM method to study the conserved Ser-363 residue in Rubisco of all variant which is found to involve in oxygen sensitivity of archaeal Rubisco.

2.4 Plasmid Transformation in *Escherichia coli*

Plasmid transformation refers to the process where exogenous or foreign DNA is introduced into the host organism for further replication. Chemical transformation or heat shock method and high-voltage electroporation are two common techniques used to insert foreign DNA into bacteria, yeast or plant cells. Electroporation demonstrates a 10-fold greater efficiency in transforming plasmids of all sizes compared to chemical transformation (Green and Sambrook, 2020). The high-voltage pulse that is administered briefly helps to permeabilize the cell membrane which enables the entry of DNA and other small molecules into the host cells (Lessard, 2013). Electrocompetent cells are kept at a low temperature prior to electroporation, preventing heating generated from the electric current which ensures the survival of cells. Standard settings of electroporation for most *Escherichia coli* strains were reported to be 1.8 kV, 25 μF, and 200 Ω (Lessard, 2013; Green and Sambrook, 2020). Dower, Miller and Ragsdale (1988) found that transformation efficiency is at its highest when electrical is supplied at the range from 1.5 kV to 2.0 kV. Besides, *E. coli* has remained the most prevalent host employed for mutagenesis and heterologous protein expression studies due to the characteristics of structural and genetic simplicity for culturing and manipulation, well-established physiological understanding, rapid protein expression capability, and economic viability (Fakruddin et al., 2013; Chang et al., 2023; Liu et al., 2023). *E. coli* is known to demonstrate exceptionally rapid growth kinetics which enables high-density cell culture to be achieved conveniently (Rosano and Ceccarelli, 2014). Pope and Kent (1996) reported that the transformation of plasmid in *E. coli* with high efficiency can be easily achieved within 5 minutes (mins) via electroporation.

Therefore, the plasmid transformation of *E. coli* using electroporation method would be a suitable method for this study due to time constraints.

2.5 Comparative Studies of Cyanobacterial and Eukaryotic Rubisco Genes in Species-Specific Assembly

The study conducted by Koay, Wong and Lim (2016) revealed the potential regions in the RbcL that might be essential for species-specific formation of Rubisco holoenzyme, specifically as a chaperone binding site. This addressed the long-existing problem associated with heterologous assembly of eukaryotic Rubisco in *E. coli* by unveiling its mechanism at the molecular level. Sequential exchange of RbcL sequences between cyanobacteria *Synechococcus elongatus* and green alga *Chlamydomonas reinhardtii* generated ten chimeric mutants and the respective ability to reconstruct in *E. coli* was assessed. Among the ten chimeras, only four tolerated the mutation, uncovering the non-critical regions (residues 1-47, 48-97, 198-247, 448-472) in species-related Rubisco holoenzyme assembly which can be used as targets for future genetic engineering without disturbing the assembly efficacy in Rubisco. Conversely, swapped sections that disrupted the reconstruction of the holoenzyme in *E. coli* suggested the functional role of residues that were mutated within the section in the Rubisco assembly. The rationale underlying the phenomenon has yet to be discovered. However, it was hypothesized that some of the residues might serve as the recognition site for chaperone binding whereby the mutation could abolish the binding affinity of chaperones to the RbcL hence resulting in improper folding or non-assembly of the holoenzyme. Some of the mutations disturb the

stability of Rubisco tertiary structure, thus collectively obstructing its assembly into the quaternary structure. Yeap et al. (2018) investigated the combinatorial effects of mutations that were previously reported as non-critical when swapped individually between *Synechococcus* and *Chlamydomonas* RbcLs. Collinear swapping was made between the RbcL and RbcS sequences of the two aforementioned organisms to produce five chimeras comprising different combinations of the mutational sites (residues 1-47, 48-97, 198-247, 448-472). Formation of holoenzyme was observed only in the chimeric holoenzyme that has both the first two sections swapped, indicating that the additive mutations on a single RbcL cause an overall destabilizing effect that abrogates the L_8S_8 assembly in other chimeric mutants.

Ong et al. (2023) extended the findings of Koay, Wong and Lim (2016) by examining three of the regions that were hypothesized to be crucial for L_8S_8 holoenzyme formation in *E. coli*. The study magnified the three sections (residues 248-297, 348-397, 398-447) by dividing them into six sections (residues 248-272, 273-297, 348-372, 373-397, 398-422, 423-447) prior to homologous switching of *Synechococcus* RbcL with that of *Chlamydomonas*. Two out of the six chimeric mutants resulted in successful assembly which successfully narrowed down the range of hypothesized residues important for chaperone-mediated folding, possibly for GroEL-RbcL interaction. Notably, chimeric mutants that contain mutation at residue 425 (V425T) were all reported for unsuccessful heterologous holoenzyme assembly of Rubisco (Koay, Wong and Lim, 2016; Ong et al., 2023). Therefore, the current study aims to explore

the potential role of residue 425 in the large subunit of Rubisco of *Synechococcus elongatus* for heterologous assembly in *E. coli*.

2.6 Bioinformatics Analysis

2.6.1 Sequence Alignment

Bioinformatics tools provide an economical, speedy, and accurate alternative to experimental approaches which are generally time-consuming, laborious, and expensive. With the increasing amount of biological data recorded in databases, computational analysis has been a great help in planning research direction and prediction of outcomes. It also allows high-throughput analysis of data in screening potential candidates for experimental validation. Visualization and alignment of biological sequences have been a norm in research, particularly in mutagenesis studies where the mutation functionality is determined based on the target sequences (Hall, 1999). BioEdit (Biological sequence editor) and ApE (A plasmid Editor) are both commonly used computational tools in sequence analysis. Sequence alignment via algorithms such as ClustalW and MUSCLE in BioEdit are employed by several comparative studies of proteins of different variants (Alshrari et al., 2021; Alam et al., 2023; Mosad et al., 2023). ApE software enables nucleotide sequence visualization in various forms including vector map, open reading frame (ORF) map, or text map (Davis and Jorgensen, 2022). Additional features such as simulation of PCR reaction and restriction digestion expedite the process of DNA analysis where the position of electrophoretic bands on agarose gel can be predicted (Britton et al., 2007). Balamurugan et al. (2021) successfully verified the phylogenetic relationship

between the wild and commercial strains of medicinal macrofungi via the comparative analysis of mismatches in genomic sequences on ApE software.

2.6.2 Protein Modeling and Structure Visualization

Protein structure plays an important role in functional studies where it provides insight at the molecular level and is often crucial in hypothesizing how it can be altered or modified to enhance or diminish its function. Molecular modeling software such as SWISS-MODEL allows the modelling of the three-dimensional (3D) structure of a protein based on previously identified homologous crystal structures deposited in the database (Duarte et al., 2018). Modelling of homologous proteins is frequently conducted in phylogenetic and mutational studies for further protein analyses (Saini et al., 2018; Alam et al., 2023; Kabiraj et al., 2023). SWISS-MODEL is a protein homology platform where the target sequence can be submitted in the form of FASTA, Clustal, plain text or even UniProtKB accession code while the output is available in several formats including PDB, mmCIF, JSON, and DeepView (Waterhouse et al., 2018). The input data will be used for template search against the SWISS-MODEL Template Library (SMTL) curated via the Protein Data Bank (PDB). In addition, information such as the oligo-state, Global Model Quality Estimation (GMQE), and Qualitative Model Energy ANalysis (QMEAN/QMEANDisCo) are displayed on the output page for quality evaluation of protein structure modelled. Protein modeling is generally coupled with visualization where the structure-activity relationship or the structure-based protein-protein interaction can be studied. Molecular viewer such as PyMOL provides the convenience of

structural visualization in the form of high-quality interactive graphics and animations. The studies conducted by Sevindik (2018, 2019) utilized PyMOL software for the comparative analysis of 3D structures of *Morus*- and *Sideritis*-derived RbcL proteins. Moreover, PyMOL is also capable of determining the root-mean-square-deviation (RMSD) of two protein chains via superimposition as performed in the study of Alam et al. (2023). RMSD value quantitatively portrays the degree of similarity between two superimposed proteins where a lower RMSD value implies a greater resemblance whereas a high RMSD value suggests a substantial structural deviation. Generally, the deviation is considered minor or negligible for RMSD value below 2 Å (Carugo, 2003).

2.6.3 Prediction of Change in Protein Stability and Physicochemical Properties

Experimental mutagenesis studies are often supported by the *in silico* prediction of mutational influences on protein stability in investigating the stabilizing or destabilizing potential of mutation on the native structure. This is reported in a study conducted by Kumari and Rameshwari (2022) where the I-Mutant2.0 server is used to reveal the detrimental effects of missense and point mutations on the native BCL-2 protein structure. A similar application of the I-Mutant2.0 server was reported by Shahzan, Girija and Priyadharsini (2019). I-Mutant2.0 employs a support vector machine (SVM)-based algorithm that allows automated forecasting of protein stability upon single amino acid change either in structure or sequence (Capriotti et al., 2005). Predicted stability is presented as the free energy change ΔΔG where a positive value indicates enforcement

while a negative value suggests destabilizing of structure. The physicochemical properties of a protein are closely linked to its physiological behavior and its interaction with the environment. Theoretically, mutated proteins are expected to have altered physicochemical characteristics and are often studied by researchers to understand the functional and evolutionary purpose of the mutation (Math et al., 2022). These properties are conveniently predicted via proteomic tools on the ExPASy server including “Compute pI/MW”, ProtParam, and ProScale (Gasteiger et al., 2005). The ProtParam tool was used in the study involving the functional characterization of date palm proteins by Subbiah, Maheswaran and Raja (2023). Overall, bioinformatics serves as a helpful analytical tool that complements experimental studies, taking advantage of extensive experimentally validated data that significantly accelerate experimental processes.

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials

All the chemicals and instruments are provided by the Department of Chemical Science, Faculty of Science of Universiti Tunku Abdul Rahman (UTAR), Perak. The chemicals and instruments used in this research are shown in **Table 3.1** and **Table 3.3**, respectively. The nucleotide sequences of the DNA primers used in this study are listed in **Table 3.2**.

Table 3.1: List of chemicals with respective manufacturers or sources.

Chemicals	**Manufacturer / Source**
Agarose	PhileKorea
Ammonium persulfate (APS)	ChemAR-SYSTERM®
Ampicillin (Amp)	Fisher Scientific
Bromophenol blue	Nacalai Tesque
BsmBI (Esp3I)	Thermo Scientific
Concentrated hydrochloric acid	QRëC™
Coomassie Brilliant Blue G-250	Nacalai Tesque
Dithiothreitol (DTT)	GoldBio
DNA primers	Apical Scientific
EconoTaq® PLUS GREEN 2X Master Mix	Lucigen
FavorPrep™ GEL/PCR Purification Mini Kit	Favorgen® Biotech Corp.
Floro^{+}Green Nucleic Acid stain	1st Base

Table 3.1: (continued)

Chemical	Manufacturer / Source
Gel loading dye, purple (6X)	New England BioLabs
Glacial acetic acid	Sigma-Aldrich
Glycine	Chemiz
HiYield™ Plasmid Mini Kit	Real Biotech Corporation
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Calbiochem®
Luria Bertani agar (Miller)	Sigma-Aldrich
Luria Bertani broth (Lennox)	Pronadisa
Magnesium chloride	QRëC™
Methanol	Sigma-Aldrich
N,N-Bis(2-hydroxyethyl)glycine (bicine)	Sigma-Aldrich
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Nacalai Tesque
PageRulerTM Unstained Protein Ladder	Fermentas Life Sciences
pE plasmid	From Koay, Wong and Lim (2016)
pTrcSynLS-MC (Wildtype plasmid)	From Mueller-Cajar and Whitney (2008)
PstI (10 U/μL)	Thermo Scientific
Sodium bicarbonate	R&M Chemicals
Sodium dodecyl sulfate (SDS)	Amresco
Sodium hydroxide	Sigma-Aldrich
Sucrose	Fisher Bioreagents
Tetracycline (Tet)	Calbiochem®

Table 3.1: (continued)

Chemical	Manufacturer / Source
Tris base	Fisher Bioreagents
Tris-HCl	Fisher Bioreagents
T4 DNA Ligase C301	Vazyme
VC 1 kb DNA ladder (ready-to-use)	Vivantis Technologies
XL1-Blue Supercompetent Cells	Agilent Technologies
40% Acrylamide/bis-Acrylamide 19:1	Bio Basic Canada Inc.
40% Acrylamide/bis-Acrylamide 37.5:1	Fisher Bioreagents

Table 3.2: DNA sequence of primers used in plasmid construction.

Primer	Sequence
425-PstI-Fwd	5'-CTTGAAGCTTGTA**CTGCAG**CTCGTAAC-3'
425-PstI-Rev	5'-GTTACGAG**CTGCAG**TACAAGCTTCAAG-3'
Syn50-Fwd-BsmBI	5'-TCAGCCGGGTCGTCTCGCTGACGAAGCTGGTG-3'
Syn50-Rev-BsmBI	5'-ATCGCCGCACCCGTCTCGTCAGCAGGGACAC-3'
Vec-Fwd	5'-CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTC-3'
Vec-Rev	5'-GACACATGCAGCTCCCGGAGACGGTCACAGCTTG-3'
Chl450-Rev-BsmBI	5'-CAGCAGCAAGTCGTCTCGACCATTTACAAGCTG-3'

The BsmBI recognition site is underlined and the PstI site is **bolded**; Sequence encodes for Val-425 are indicated in dashed underline.

Table 3.2: (continued)

Primer	Sequence
Syn450-Fwd-BsmBI	5’-CTTCGTGAAGCCGTCTCGTGGTCGCCTGAACTG-3’
425V-Fwd	5’-CTTGAAGCTTGTGTCCAAGCTCGTAAC-3’
425V-Rev	5’-GTTACGAGCTTGGACACAAGCTTCAAG-3’

The BsmBI recognition site is underlined and the PstI site is **bolded**; Sequence encodes for Val-425 are indicated in dashed underline.

Table 3.3: List of instruments with respective manufacturers or brands.

Instruments	Manufacturer / Brand
ChemiDoc™ XRS+ gel imaging system	Bio-Rad Laboratories
CP224S analytical balance	Sartorius
Maxi Horizontal Gel Electrophoresis System ME20 series	Major Science
Microfuge® 16 centrifuge	Beckman Coulter®
Micropipette	Eppendorf
MicroPulser™ electroporator	Bio-Rad Laboratories
Mini Vertical Gel Electrophoresis Apparatus MV-10DSYS	Major Science
MP250V power supply	BR Biochem
NanoDrop™ 2000 spectrophotometer	Thermo Scientific
NB-205L shaking incubator	N-BIOTEK
PCR SuperCycler Trinity	Kyratec
Shaking incubator	LabTech
VCX 130 ultrasonic processor	Sonics®
Velocity 14R refrigerated centrifuge	Dynamica

3.2 Experimental Design

The study involved *in vitro* and computational analysis of Rubisco assembly in *E. coli* as depicted in **Figure 3.1**. Site-directed mutagenesis (SDM) that changed Thr to Val at residue 425 of Rubisco was conducted in second mutagenesis step of the pE plasmid which was subsequently expressed in E. coli prior to analysis in SDS-PAGE. Computational analysis was conducted via multiple servers and software that mainly involved the prediction of change in protein stability and physicochemical properties upon the mutation.

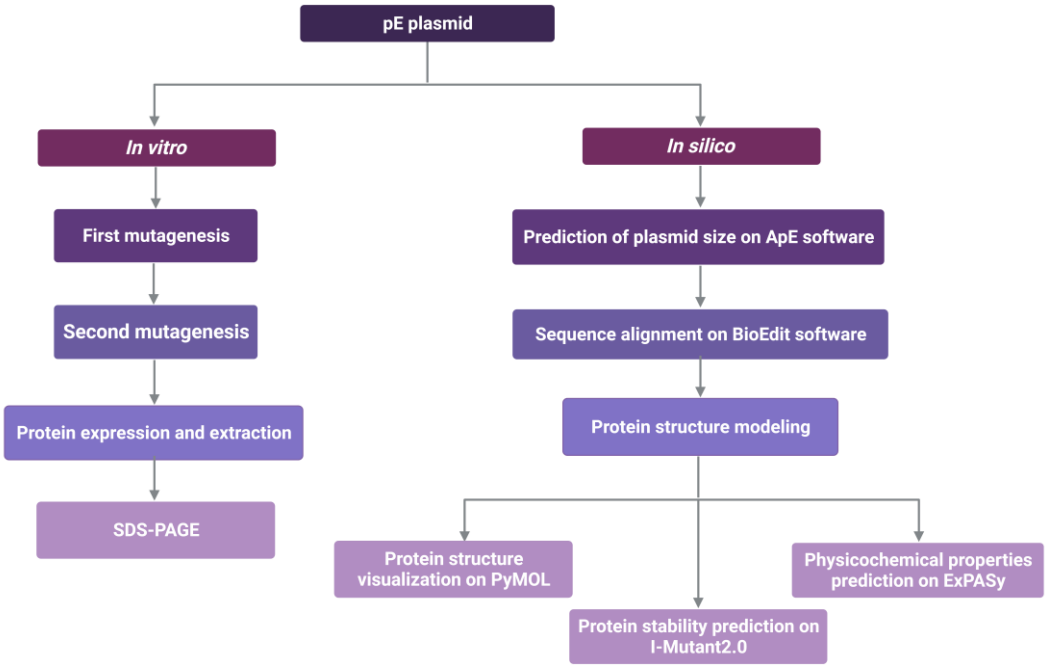


Figure 3.1: An overview of the experimental design for this study.

3.3 *In Vitro* Analysis

3.3.1 First Mutagenesis

3.3.1.1 Plasmid Construction via Fragment-Based Approach

Mutation of plasmid pE to introduce an additional PstI restriction site was conducted via polymerase chain reaction (PCR) using pE as the DNA template. Synthesis of mutated pE (pE_{M1}) plasmid was carried out by dividing the circular plasmid into four fragments (A_1, B_1, C_1, D_1) with the respective primers (**Table 2**) responsible for that DNA segment. 425-PstI-Rev and 425-PstI-Fwd were the two main primers responsible for mutagenesis that insert the PstI site in the plasmid. The PCR protocols used to synthesize the DNA fragments are shown in **Table 3.5**, modified from Koay et al. (2016). PCR products were validated with gel electrophoresis (100 V, 30 mins).

Table 3.4: Composition of each PCR tube responsible for constructing respective fragments.

Fragment	**A_1**	**B_1**	**C_1**	**D_1**
Forward primer (µL)	1 (Syn50-Fwd)	1 (425-PstI-Fwd)	1 (Vec-Fwd)	1 (Syn450-Fwd)
Reverse primer (µL)	1 (425-PstI-Rev)	1 (Chl450-Rev)	1 (Syn50-Rev)	1 (Vec-Rev)
pE (µL)	1	1	1	1

The primers used for amplification of each fragment are stated in bracket.

Table 3.4: (continued)

Fragment	**A_1**	**B_1**	**C_1**	**D_1**
Master Mix (µL)	25	25	25	25
Deionized water (µL)	32	32	32	32
Total (µL)	60	60	60	60

3.3.1.2 Joining of Fragments A_1 and B_1

The two short fragments, A_1 and B_1 were joined to form a single DNA fragment (AB) via PCR protocol as shown in **Table 3.5**. 25 µL of purified products from PCR tubes A_1 and B_1 were mixed with 1.5 µL of Syn50-Fwd and Chl450-Rev primers each, along with 40 µL of EconoTaq Plus Green 2X master mix and 7 µL of deionized water. PCR products were purified and verified via gel electrophoresis (100V, 30 mins).

Table 3.5: PCR protocol for DNA fragment amplification in mutagenesis.

Cycling steps	**Temperature (°C)**	**Time (mins)**	**Number of cycles**
Hold (Initial denaturation)	94	3	1
Denaturation	94	1	30
Renaturation	56	2	
Extension	72	2	
Final extension	72	5	1

3.3.1.3 BsmBI Digestion

In three separate tubes, purified fragments AB_1, C_1, and D_1, were mixed with 10X Tango buffer, BsmBI (Esp3I) restriction enzyme, and 20 mM dithiothreitol (DTT) for digestion to produce sticky ends prior to ligation. The mixtures were incubated overnight at 37°C and purified prior to gel electrophoresis (100 V, 30 mins). The composition of each tube is shown in **Table 3.6**.

Table 3.6: Composition of each PCR tube for BsmBI digestion.

Fragment	**AB_1**	**C_1**	**D_1**
Purified PCR product (µL)	25	25	25
10X Buffer Tango (µL)	3	3	3
BsmBI enzyme (µL)	1	1	1
20 mM DTT (µL)	1	1	1
Total (µL)	30	30	30

3.3.1.4 Ligation

The three DNA fragments (AB_1, C_1, D_1) were ligated by T4 DNA ligase to form a circular plasmid. Samples were incubated overnight at 16°C. Ligated products

were validated via gel electrophoresis (100 V, 30 mins). The components of the ligation mixture are shown in **Table 3.7**.

Table 3.7: Volume of each component added for ligation.

Component	**Volume (µL)**
DNA fragment AB	8.5
DNA fragment C	4.0
DNA fragment D	4.0
T4 DNA ligase	0.5
Buffer	3.0
Total	20.0

3.3.1.5 Transformation

1 µL of mutated pE plasmid (pE_{M1}) was added to 50 µL of electrocompetent tetracycline (Tet)-resistant XL1-blue *Escherichia coli* strain and electroporated at 1600 V using MicroPulser^{TM} electroporator. Transformed cells were immediately suspended in 100 µL Luria Bertani (LB) broth and plated onto ampicillin (Amp) supplemented LB agar for selection of positive transformant (Amp^r) colonies. The cultures were incubated for 16 hours at 37°C. Bacterial colonies were isolated and sub-cultured in 10 mL LB broth containing 30 µL of 100 mg/mL Amp. Subcultures were incubated overnight in the NB-205L shaking incubator (200 rpm) at 37°C.

3.3.1.6 Plasmid Extraction

Plasmids were extracted from the transformed *E. coli* colonies using the HiYield™ Plasmid Mini Kit. 1.5 mL of transformed *E. coli* culture was transferred to a 1.5 mL microcentrifuge tube. Cells were harvested at the bottom of the tube as cell pellets by centrifuging at 14000 xg for 1 min where supernatant was discarded. 200 µL of PD1 buffer (RNase) was added to the microcentrifuge tube to resuspend the cell pellet and the mixture was vortexed until the precipitate completely dissolved. 200 µL of PD2 buffer was added and mixed by 10 times inversion. The tube was allowed to stand at room temperature for 2 mins to achieve homogenous lysate. 300 µL of PD3 buffer was added and mixed by inverting the tube 10 times. The tube was centrifuged at 14000xg for 3 mins at room temperature. The supernatant was transferred to the PD column and placed in a 2 mL collection tube. The tube was then centrifuged at 14000xg for 30 seconds at room temperature and flow-through was discarded.

600 µL of wash buffer was added to the PD column and centrifuged at 14000xg for 30 seconds. Flow-through was discarded and centrifuged at 14000xg for 3 mins to dry the column matrix. The dried PD column was transferred to a new 1.5 mL microcentrifuge tube for elution. 50 µL of deionized water was added to the column matrix and allowed to stand for 2 mins for complete adsorption. The tube was centrifuged at 14000xg for 2 mins for elution of purified DNA. The eluted product was validated with gel electrophoresis (100 V, 30 mins). The concentration and purity of the extracted products were determined via NanoDrop™ 2000 spectrophotometer.

3.3.1.7 PstI Digestion

10 µL of extracted product was digested with 1 µL of 10X Buffer O and PstI restriction enzyme. The digestion mixture was incubated at 37°C overnight before being analyzed via gel electrophoresis (100 V, 30 mins).

3.3.1.8 DNA Purification

Purification of PCR products was performed using FavorPrep^{TM} GEL/PCR Purification Mini Kit. The PCR product was mixed and vortexed with five times the volume of the FADF buffer. The mixture was transferred to the FADF column and centrifuged at 11000xg for 30 seconds. Flow-throughs were discarded. 750 µL of wash buffer provided was added to the FADF column and centrifuged at 11000xg for another 30 seconds. Flow-throughs were discarded and the tubes were centrifuged again at 14000xg for 5 mins to remove any residual liquid. The FADF column was then transferred to a new microcentrifuge tube where 60 µL of deionized water was added to the membrane center and allowed to stand for 2.5 mins before eluting at 14000xg for 5 mins.

3.3.1.9 Gel Electrophoresis

1% (w/v) agarose gel was prepared by dissolving 0.6 g of agarose in 60 mL of Tris-acetate EDTA (TAE) buffer. The agarose mixture was heated in the oven three times to ensure all powder had been dissolved. 2 µL of Floro^{+}Green Nucleic Acid stain was added to the gel when it was warm. The gel was gently poured into the gel tank and allowed to solidify at room temperature. Solidified

agarose gel was transferred to an electrophoresis tank filled with TAE buffer and was ready to be used. 10 µL of sample was mixed with 3 µL of 6X loading dye before loading into the well. 5 µL of VC 1 kb DNA ladder was loaded in one of the wells for approximation of band size. Voltage was set at 100 V and allowed to run for 30 mins. Visualization of the band was performed via the ChemiDoc™ XRS+ system equipped with Image Lab™ software.

Protocol setup was as follow:

Application: Nucleic acid gels; SYBR green

Imaging area: Bio-Rad Ready Gel

Image Exposure: Intense Bands

Filter: Standard filter (Filter 1)

3.3.2 Second Mutagenesis

The procedure for the second mutagenesis followed the same as that of the first mutagenesis as mentioned in **Section 3.3.1**. Primers 425-PstI-Fwd and 425-PstI-Rev were replaced with 425V-Fwd and 425V-Rev which are responsible for site-directed mutagenesis of Thr-425 to Val-425. pE_{M1} was used as the DNA template for the construction of the four fragments via PCR. Ligated products were transformed into *E. coli* at different dilution factor (1X and 1.5X) and incubated overnight at 37 °C on Amp plate. Amp^r colonies were inoculated and subcultured overnight in 100 µg/mL Amp LB broth at 37 °C and 200 rpm. Plasmids were extracted using HiYield™ Plasmid Mini Kit as described in **Section 3.3.1.6**. The

concentration and purity of plasmid were assessed via NanoDrop^{TM} 2000 spectrophotometer.

3.3.3 Protein Expression and Extraction

Extracted second mutagenesis mutant (pE_{M2}) was transformed into electrocompetent *E. coli* and grown overnight on an Amp plate. Wildtype (WT) plasmids were transformed as positive control while *E. coli* without plasmid were grown on a Tet plate as the negative control. A single colony of Amp^r from each plate was picked and subcultured at 37 °C, 200 rpm in 10 mL LB broth (100 µg/mL Amp). The whole overnight culture was poured into 150 mL LB broth (100 µg/mL Amp) to stimulate the exponential growth of cells until an OD_{600} of 0.5 is reached. The culture was kept on ice for an hour before 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture for 16-hour protein induction at 37 °C and at 200 rpm.

Cells were pelleted via centrifugation (5000 g, 5 mins, 4 °C) and resuspended in chilled extraction buffer at 10% w/v. The composition of the extraction buffer includes 50 mM Bicine/NaOH at pH 8, 10 mM sodium bicarbonate, 2 mM DTT, and 10 mM magnesium chloride. Disruption of cells was performed by sonication at 40% amplitude with 10s pulses for an overall of 10 mins. The resulting protein lysate was used for further expression evaluation on SDS-PAGE.

3.3.4 Protein Expression Analysis

Expression of the Rubisco enzyme was validated on SDS-PAGE. The proteins in the crude lysate were denatured by diluting in 1X loading buffer (100 mM DTT, 0.05% bromophenol blue, 5% SDS, 30% sucrose) at 3:2 ratio and boiled for 5 mins at 95 °C. DTT was added freshly to the mixture to ensure separation efficacy. 4 µL of boiled lysates along with PageRuler^{TM} Unstained Protein Ladder were loaded individually onto the 1 mm-thick polyacrylamide gel comprised of 2 mL of 4% stacking gel and 5 mL of 12% resolving gel in Mini Vertical Gel Electrophoresis Apparatus MV-10DSYS filled with approximately 900 mL of 1X running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). The total cellular proteins were segregated at 40 V, 150 mA until the proteins were stacked at the stacking-resolving gels interface. Voltage was raised to 60 V for subsequent separation in the resolving gel until the dye front appeared near the bottom of the gel.

Table 3.8: Composition of stacking and resolving gels in SDS-PAGE.

Components	**4% Stacking gel**	**12% Resolving gel**
40% Acrylamide/bis-Acrylamide 37.5:1	200 µL	-
40% Acrylamide/bis-Acrylamide 19:1	-	1500 µL
1.5 M Tris-HCl (pH 8.45)	200 µL	1300 µL
10% APS	10 µL	25 µL
0.1% TEMED	2 µL	5 µL
10 % SDS	20 µL	50 µL
50% Sucrose	-	1360 µL
Deionized water	1568 µL	760 µL
Total	2000 µL	5000 µL

Resolving gel was first soaked in fixing solution (50% methanol, 10% acetic acid) for 45 mins prior to staining. Staining of resolving gel was carried out by soaking the gel in Coomassie Brilliant Blue solution (0.25% Coomassie Brilliant Blue G-250, 50% methanol, 7% acetic acid) with gentle agitation for 45 mins and subsequently destained in destaining solution (40% methanol, 7% acetic acid) for an hour. The gel was transferred to distilled water for the further destaining process until protein bands could be clearly seen. Distilled water was changed every 3 to 4 hours to ensure destaining efficiency. The gel was viewed on a white light conversion screen in the ChemiDoc^{TM} XRS+ system with the setup shown below.

Application: Protein gels; Coomassie Blue

Imaging area: Bio-Rad Ready Gel

Image Exposure: Intense Bands

Filter: Standard filter (Filter 1)

3.4 *In Silico* Analysis

3.4.1 Plasmid Size Prediction

Prediction of plasmid and DNA fragment sizes was conducted via ApE software (Davis and Jorgensen, 2022). The nucleotide sequence of plasmid pE (Koay et al., 2016) was uploaded to the software in ape format. The size of the plasmid was displayed on the top of the window "Sequence". Image of plasmid pE with indicated plasmid size, RbcL, and RbcS were obtained from "Graphic Map" in the "Enzymes" section. The approximate position of bands was predicted on the software by the "Digestion" function in the "Enzymes" section.

3.4.2 Sequence Alignment

Alignment of sequences between original and mutated RbcL was conducted via the ClustalW algorithm in BioEdit (Hall et al., 2011). The FASTA sequence of the RbcL subunit was first retrieved from the "Translate…" function in the "ORFs" section of the highlighted nucleotide sequence of RbcL on ApE software. The protein sequences of original and mutant RbcL were uploaded individually on the window via the "New sequence" function in the "Sequence" section on

BioEdit. On “Accessory Application”, “ClustalW Multiple Alignment” was selected for the alignment of the two sequences.

3.4.3 Protein Modeling and Structure Visualization

Modeling of three-dimensional (3D) structures of RbcL of pE and pE_{M2} origin was achieved via SWISS-MODEL server (https://swissmodel.expasy.org/) (Accessed: 22 June 2023) (Waterhouse et al., 2018). Protein sequences generated from ApE (**Section 3.4.2**) were uploaded to the server in FASTA format for modelling. RbcL models were downloaded from the output page in PDB format. The interactions between the mutated and wildtype RbcL were then visualized on PyMOL software (The PyMOL Molecular Graphics System, Version 0.99 Schrodinger, LLC) by superimposing the protein models via command line input of “pair_fit name ca and pE, name ca and pE-T425V”. The root-mean-square-deviation (RMSD) value computed by PyMOL during superimposition was recorded. Structural interaction particularly van der Waals forces between the target (Thr-425, Val-425) and surrounding residues were also assessed via specific command entry on PyMOL. The distance between residues was calculated using the “Measurement” tool in the “Wizard” task on the toolbar.

3.4.4 Protein Stability Prediction

The change in the structural stability of the RbcL subunit upon single amino acid mutation was evaluated via I-Mutant2.0 (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) (Accessed: 16 August 2023) (Capriotti, Fariselli and Casadio, 2005). Protein sequence in single-letter code was uploaded as input with settings

of position “425”, new residue “V”, default temperature and pH at 25 °C and 7, respectively. “Sign of free energy change value (DDG)” was selected as the prediction output.

3.4.5 Physicochemical Properties

Physicochemical properties of pE-derived and pE_{M2}-derived RbcL subunit were predicted via several servers under ExPASy (https://www.expasy.org/) (Gasteiger et al., 2005). Parameters include isoelectric point and molecular weight via “Compute pI/MW” (https://web.expasy.org/compute_pi/); *in vivo* half-life, instability index, aliphatic index, and grand average of hydropathy (GRAVY) via ProtParam (https://web.expasy.org/protparam/); lastly, hydropathy plot via ProtScale (https://web.expasy.org/protscale/) with the default amino acid scale of Kyte & Doolittle (Kyte and Doolittle, 1982). The aforementioned tools were accessed on 21 June 2023. The FASTA sequences of the RbcLs were uploaded separately to the server for analysis of respective parameters.

CHAPTER 4

RESULTS

4.1 First Mutagenesis

4.1.1 Construction of Plasmid Mutant (pE_{M1})

The first mutagenesis of the pE plasmid is conducted in multiple PCR steps with the initial construction of the mutated plasmid in four separate fragments, named A_1, B_1, C_1, and D_1. **Figure 4.1 (a)** shows the electrophoretic bands of the four fragments which correspond to the particular segment of the complete plasmid. The approximate band sizes showed on the gel relative to the 1 kb DNA ladder conformed with the predicted fragment sizes of 1165 bp (A_1), 103 bp (B_1), 1889 bp (C_1), and 3131 bp (D_1) on ApE software. The approximate band sizes add up to more than 6000 bp which is in agreement with the predicted plasmid size (6148 bp) for mutated pE (pE_{M1}) as shown in **Figure 4.2**. PCR of fragments A and B in the same mixture resulted in a single band on the gel at about 1250 bp (**Figure 4.1 (b)**) which closely coheres with the expected 1241 bp fragment named AB_1. To create overhangs for ligation, the three fragments (AB_1, C_1, D_1) were individually subjected to BsmBI digestion after purification and the respective electrophoretic bands were found at positions similar to the aforementioned sizes. This suggests insignificant or no change in fragment size after the digestion which ensures minimal to no effect on the final plasmid size and sequence synthesis after the construct is assembled. Lastly, the presence of ligated products is validated with the appearance of a faint band at approximately 6200 bp, indicating successful ligation of the four fragments.

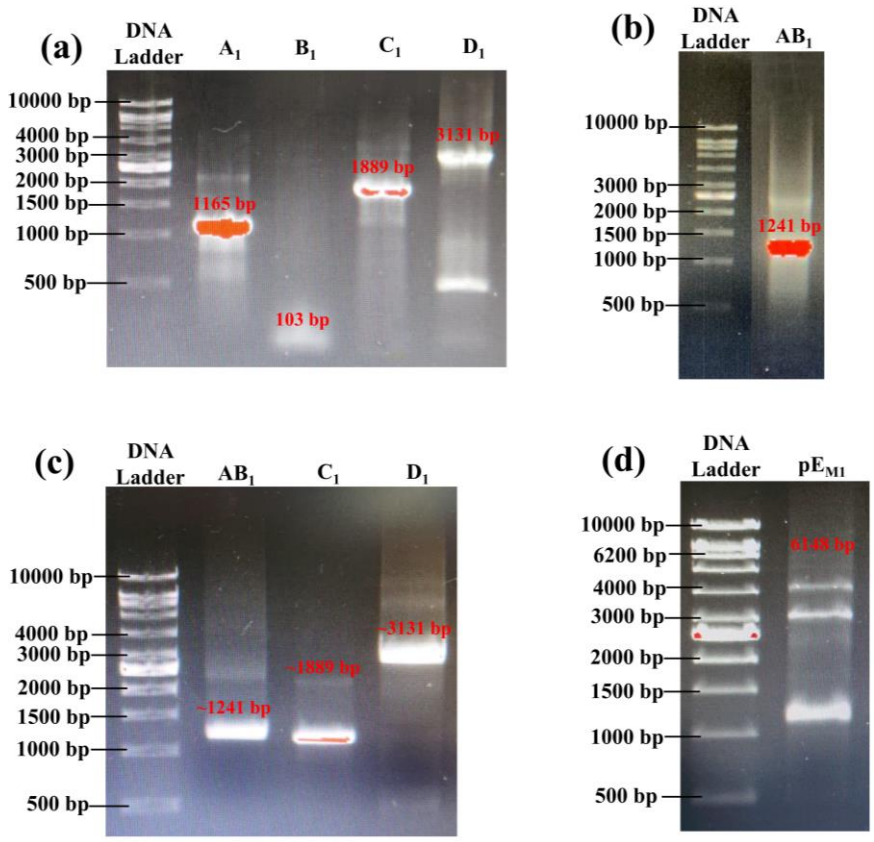


Figure 4.1: Gel images of the first mutagenesis. The images follow an order of four fragments construction (a), formation of AB_1 fragment (b), BsmBI digestion (c), and ligation (d). The predicted sizes of target fragments are indicated in red.

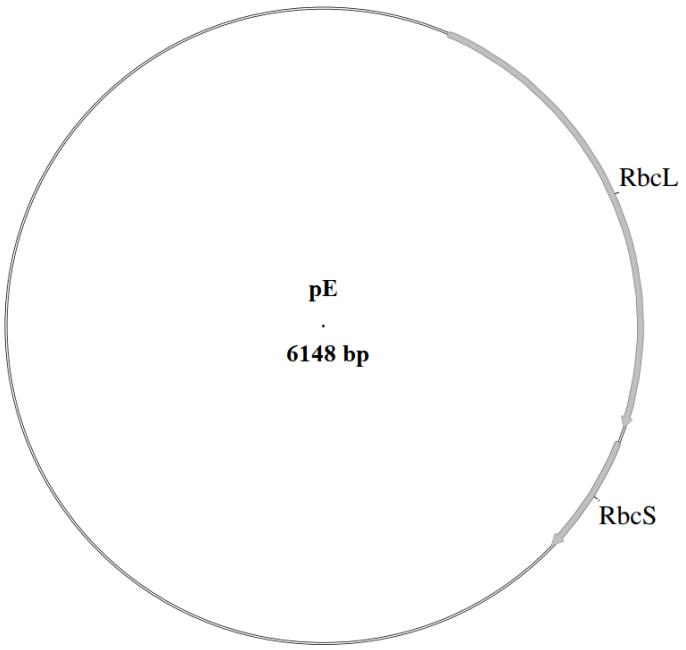


Figure 4.2: Graphical representation of pE plasmid generated in ApE software with its predicted size and locations of RbcL and RbcS genes.

4.1.2 Transformation and Validation

The mutagenesis process proceeds with the transformation of ligated products in electrocompetent Tet-resistant *E. coli* and grown on an Amp plate for positive selection. A time constant of 4.5 ms was reported during electroporation which falls within the acceptable range for efficient gene transfer. White single colonies were observed on agar plates implying successful uptake of plasmids by the bacterial cells (**Figure 4.3 (a)**). Any contamination or false positive is excluded with the absence of growth on the negative control plate. Single colonies picked and inoculated were subcultured in Amp LB broth and a milky suspension was observed while the negative control remained clear after the 16-hour incubation (data not shown). Plasmids were subsequently extracted from a portion of the culture and were confirmed by the well-defined band at 6000 bp (**Figure 4.3 (b)**). Plasmid concentrations of two different colony origins are measured at 83.5 ng/µL and 92.2 ng/µL whilst A_{260}/A_{280} ratios of 1.93 and 1.99 were computed, respectively, indicating relatively pure DNA obtained. The overall outcome of mutagenesis was validated with PstI digestion whereby two observable bands appeared at positions above 5200 bp and 500 bp which are both expected in the simulated digestion of ApE software (**Figure 4.3 (c)**). Another band at 33 bp that is computationally predicted may be too small to be detected on the gel. However, the slightly lower band observed in the second fragment compared to that of the digested original pE plasmid suggest a successful mutation.

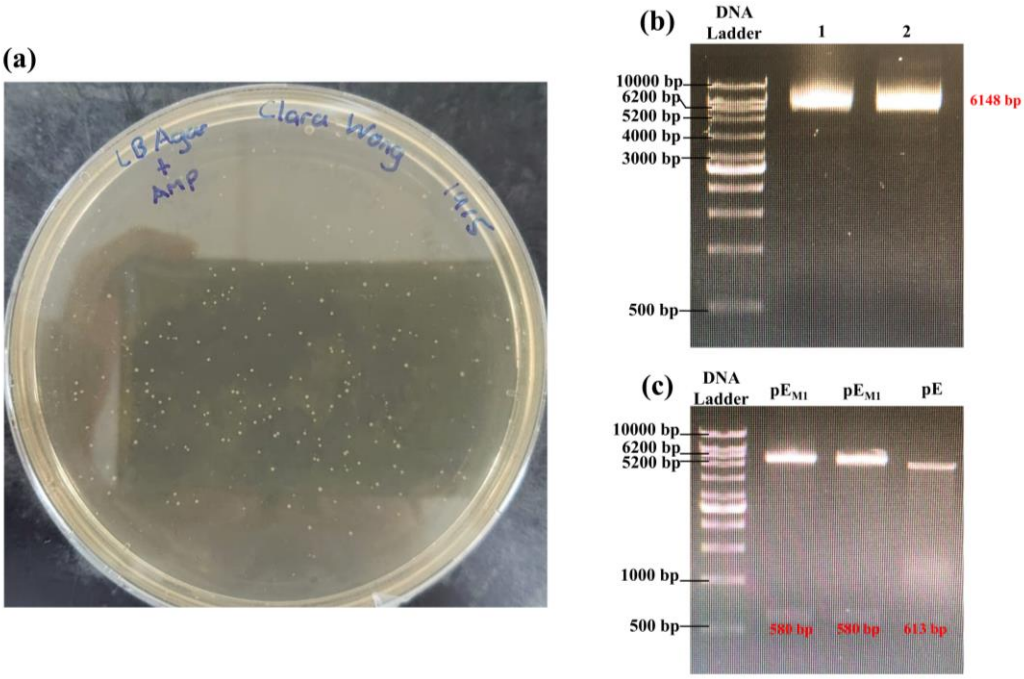


Figure 4.3: Results of pE$_{M1}$ plasmid transformation (a), extraction (b), and PstI digestion (c). Single colonies of transformed *E. coli* were observed on Amp-supplemented LB agar (a). The predicted size of each band in (b) and (c) is indicated in red.

4.2 Second Mutagenesis

As the second mutagenesis was carried out in a similar manner as the first mutagenesis, the results obtained were expected to resemble that of the first mutagenesis as supported by the prediction in ApE software (data not shown). The first four fragments (A_2, B_2, C_2, D_2) generated from pE$_{M1}$ were consistent with the predicted band sizes in ApE (**Figure 4.4 (a)**). Union of the first two fragments (A_2, B_2) via PCR yields a single DNA fragment of 1241 bp that corresponds with the approximate band size on gel which also appears to be located slightly higher than that of fragment A_2 (**Figure 4.4 (b)**). Similar to **Figure 4.1 (c)**, band sizes detected for BsmBI digestion do not deviate far from the band sizes prior to digestion. A band should be detected at approximately 6000 bp for ligated products which was not observed in **Figure 4.4 (d)**, however,

a more prominent band was observed at about 4000 bp, most probably due to supercoiling. Time constants of 4.7 ms and 4.5 ms were obtained for plasmid of 1X and 1.5X dilution factors (DFs), respectively. This also leads to a distinct growth rate in transformed bacterial cells where relatively more colonies are observed on the 1.5X plate (**Figure 4.5 (a), (b)**). Transformed colonies selected from all the Amp plates and grown in LB broth were subsequently subjected to plasmid extraction. As shown in **Figure 4.5 (c)**, plasmids were detected in all the extracts of 1.5X DF while only two out of three extracts were detected for plasmids (~6000 bp) on gel electrophoresis. 10 kb band appeared on the gel may be a consequence of inefficient purification during extraction where residual genomic DNA still remained. The concentration of DNA measured for 1.5X plasmid extract ranged between 84.9 ng/µL to 111.0 ng/µL while for 1.0X plasmid extract, concentrations were identified at 50.2 ng/µL and 107.2 ng/µL for the two samples with 6000 bp bands determined on the gel. The purity of the plasmid extracts is evaluated via the A_{260}/A_{280} ratio which all ranged between 1.91 to 1.97, suggesting relatively pure DNA in the extracts. Comparable band sizes detected for products from PstI digestion of pE_{M2} and pE indicate successful second mutagenesis (**Figure 4.5 (d)**).

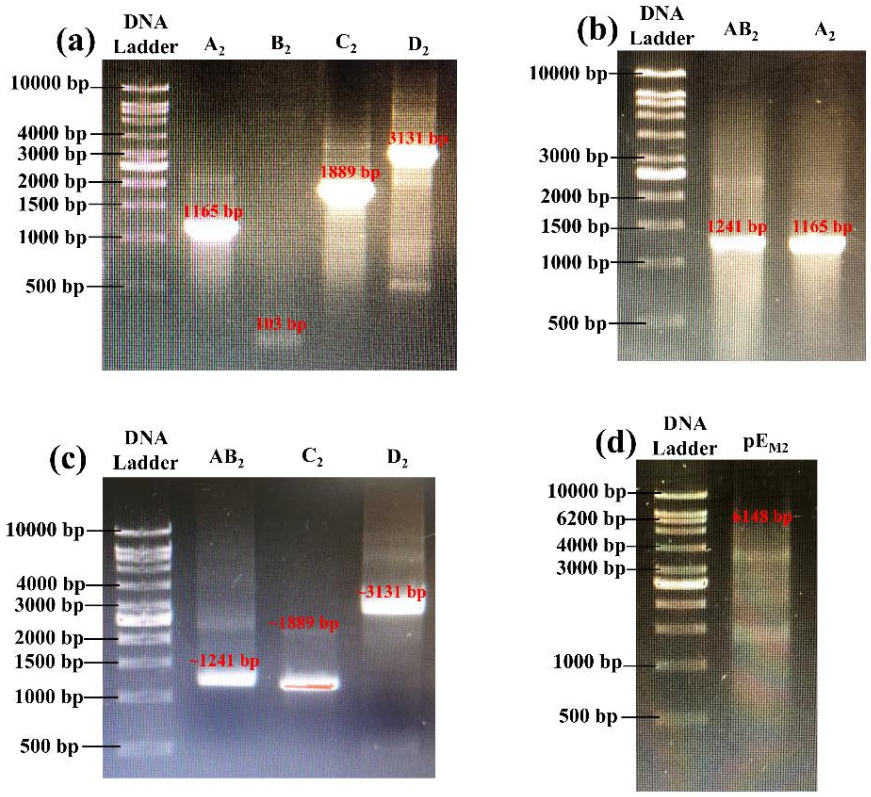


Figure 4.4: Gel images of the second mutagenesis. The images follow an order of four fragments construction (a), formation of AB_2 fragment (b), BsmBI digestion (c), and ligation (d). The predicted sizes of target fragments are indicated in red.

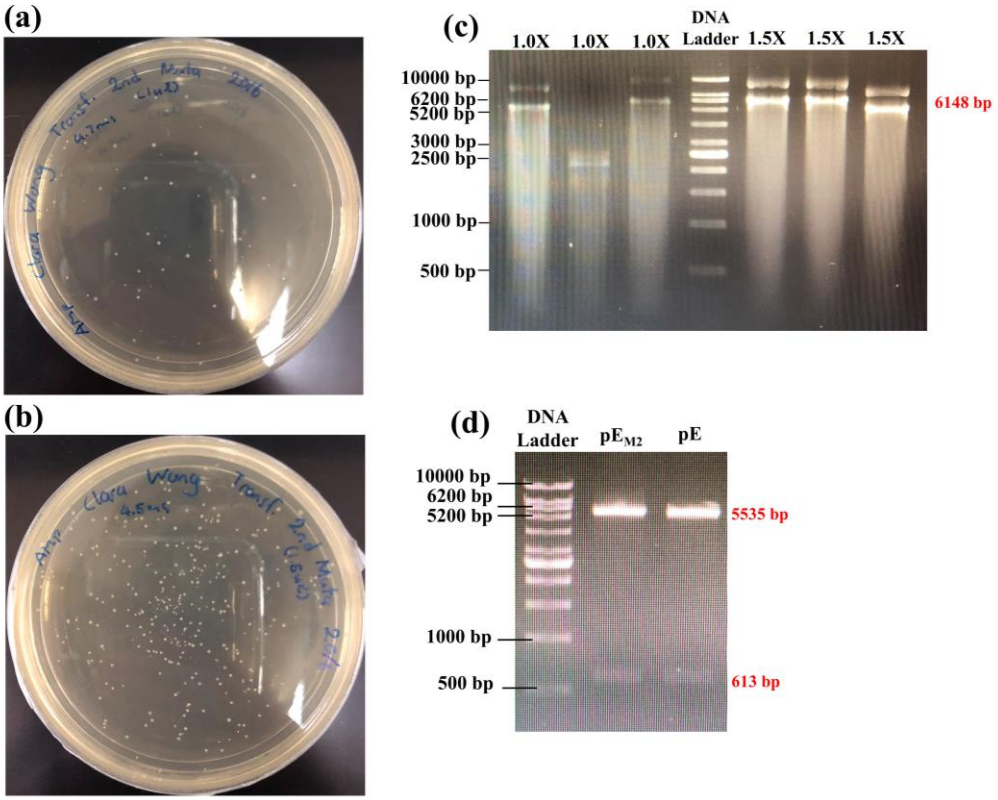


Figure 4.5: Results of pE$_{M2}$ plasmid transformation (a, b), extraction (c), and PstI digestion (d). Differential growth of colonies between transformations with plasmid at 1.0X (a) and 1.5X DF (b). Computationally estimated fragment sizes for (c) and (d) were indicated in red.

4.3 Protein Expression

Transformation of extracted plasmids from the second mutagenesis was conducted at a time constant of 4.8 ms. Single colonies on the Amp agar plate allow positive selection of transformed *E. coli* for subculture and subsequent exponential growth prior to protein induction. **Figure 4.6 (b)** shows the cultures of transformed bacterial cells (pE_{M2}, WT) compared to negative control (without *E. coli*). Cloudy suspension observed in Amp cultures containing transformed *E. coli* indicates positive selection without any biological contamination, given that the negative control (without *E. coli*) remains clear. Expression of Rubisco upon single amino acid mutation is validated on SDS-PAGE whereby the total cellular proteins extracted from the negative control (with untransformed *E. coli*), wildtype (WT), and mutants are resolved via protein electrophoresis as shown in **Figure 4.6 (c)**. The large subunits were detected in both wildtype and mutants where thick blots are observed at approximately 55 kDa when compared to the negative control. The lower bands that appeared at the bottom of the SDS-PAGE were suspected to be the small subunits of Rubisco, considering that WT and mutants appeared to have thicker blot compared to the negative control. Therefore, it can be inferred both RbcL and RbcS subunits were successfully expressed in *E. coli*.

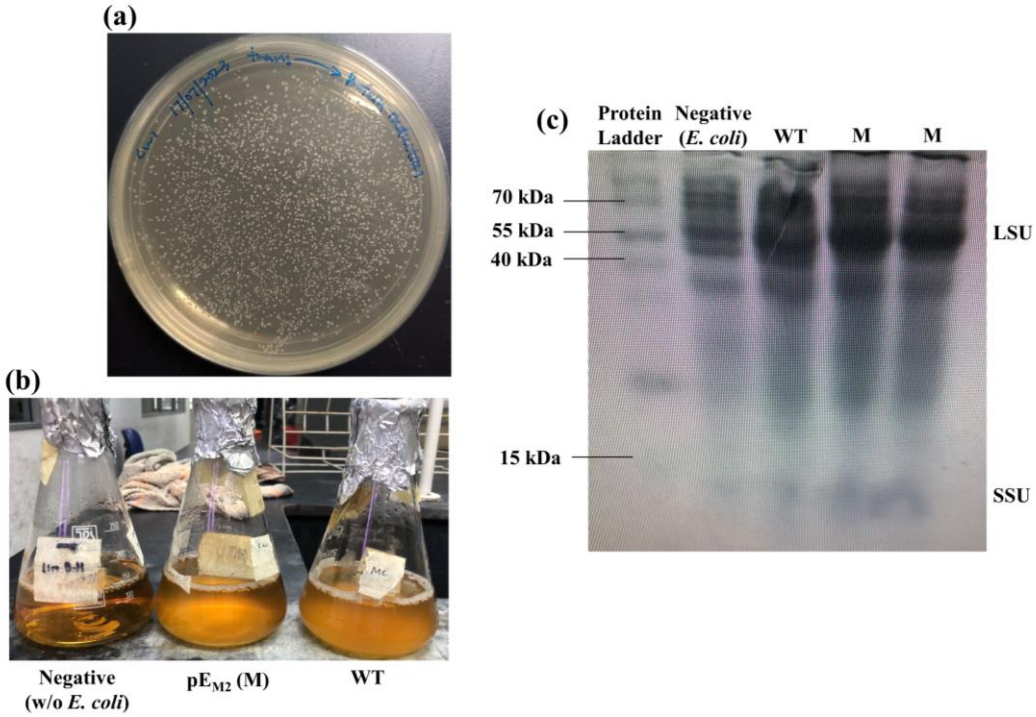


Figure 4.6: Results and observations of protein expression in *E. coli*. Colony growth of positive transformants using extracted pE_{M2} plasmids (a). Subcultures that were exponentially grown prior to IPTG protein induction (b). Total proteins extracted resolved on SDS-PAGE (c). LSU and SSU denote the large and small subunits of Rubisco, respectively.

4.4 Bioinformatics Analysis

Superimposition of RbcL subunit of the native and mutated protein in PyMOL generated an RMSD value of 0.014 Å which indicates minimal or insignificant structural change upon mutation. To study the effect of the mutation on the interactions between residues in the large subunit of Rubisco, structural visualization of the superimposed proteins was conducted. **Figure 4.7** and **Table 4.1** showed the residues that were predicted to be involved in van der Waals interaction with the target residue 425. An additional residue Pro-165 was found to interact with residue 425 after mutation from Thr to Val. Distance between the target residues (Thr-425, Val-425) and Pro-165 before and after mutation were computed (**Figure 4.7 (b)**) to be 5.34 Å and 3.95 Å, respectively with a cut-off

at 4 Å. Furthermore, the structural stability of the protein upon mutation was also assessed in I-Mutant2.0 where a DDG value of 0.48, suggesting an overall increase in stability. The physicochemical properties of the RbcL subunits predicted via the ExPASy server showed negligible change in molecular weight upon mutation. Besides, the pI value and *in vivo* half-life of the protein were unaffected upon single amino acid change (Thr → Val). Furthermore, a slight decrease in the *in vitro* instability index while a rise in aliphatic index and GRAVY value were also predicted. The hydropathy plots of native and mutated RbcL showed an upward shift of peaks at around position 425, suggesting an increased hydrophobicity. Overall, *in silico* analysis of the mutational effect on the RbcL subunit evaluates the possibility of the folding and assembly.

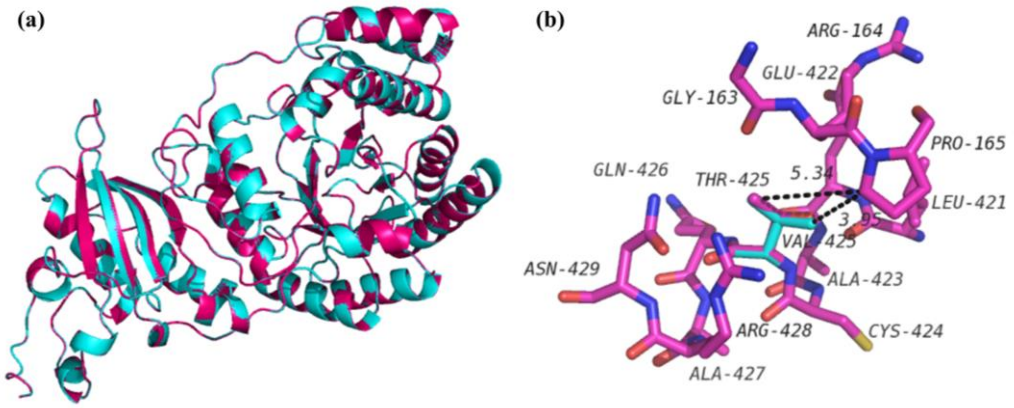


Figure 4.7: Superimposed structures between pE-derived and pE_{M2}-derived RbcLs (a). Eleven residues were predicted to interact with residue 425 via van der Waals force (b). RbcL bearing Thr-425 residue is shown in pink while RbcL with Val-425 residue is in cyan.

Table 4.1: Residues involved in van der Waals interaction (4 Å) with residue 425.

Surrounding residue	Thr-425 (pE)	Val-425 (pE_{M2})
Gly-163	√	√
Arg-164	√	√
Pro-165	-	√
Leu-421	√	√
Glu-422	√	√
Ala-423	√	√
Cys-424	√	√
Gln-426	√	√
Ala-427	√	√
Arg-428	√	√
Asn-429	√	√

Table 4.2: *In silico* prediction of physicochemical properties of RbcLs.

Properties	**pE-derived RbcL (Thr-425)**	**pE$_{M2}$-derived RbcL (Val-425)**
Molecular weight (MW)	52317.60	52315.63
Isoelectric point (pI)	6.26	6.26
In vivo half-life (*E. coli*)	>10 hrs	>10 hrs
Instability index	44.63 (unstable)	43.94 (unstable)
Aliphatic index	79.22	79.83
GRAVY	-0.256	-0.246

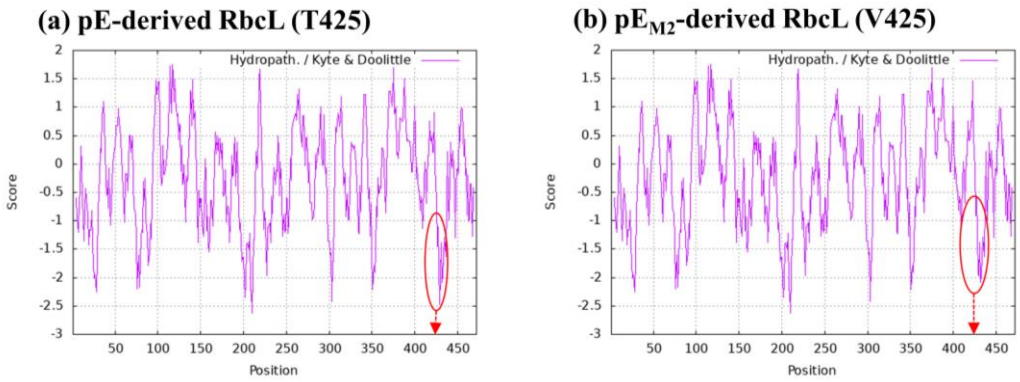


Figure 4.8: Hydropathy plots of RbcL derived from pE (a) and pE$_{M2}$ (b) plasmids.

CHAPTER 5

DISCUSSIONS

5.1 First Mutagenesis

The first mutagenesis served as a surveillance step in the mutational analysis of pE plasmid in this study by introducing a PstI recognition sequence at the targeted site within the double-stranded DNA. The two main primers, 425-PstI-Rev and 425-PstI-Fwd used in constructing fragments A_1 and B_1, respectively, are responsible for the incorporation of the recognition site. As site-directed mutagenesis (SDM) is often inefficient, especially for large plasmids, preliminary mutagenesis is required as assurance for both mutations in this study (Jia et al., 2022). Polymerase chain reaction (PCR) is employed as the method for SDM in this study and the multiple PCR steps approach is used to generate the plasmid construct by dividing the whole circular plasmid into four segments for amplification. This is mainly due to the large plasmid size of more than 6.0 kb which would be difficult to amplify within a single PCR step. Attempts were made, however, unsuccessful to introduce mutations in a straightforward manner via the conventional means of amplifying the plasmid template with a pair of primers containing the desired mutational sequence (data not shown). A highly efficient SDM method for large plasmids was recently developed by Zhang et al. (2021) which involved a fragment-based approach whereby PCR reactions were individually carried out for amplification of distinct segments within the plasmid and the fragments obtained were enzymatically ligated. The data shown in **Figure 4.1** verified that four fragments corresponding to the predicted sizes were

generated and were properly assembled into a circular plasmid of 6148 bp. Yang, Chen and Zhang (2022) successfully performed SDM on 8.3 kb and 11.0 kb plasmids by utilizing a similar multiple-fragment technique with a modification of *in vivo* assembly rather than *in vitro* ligation.

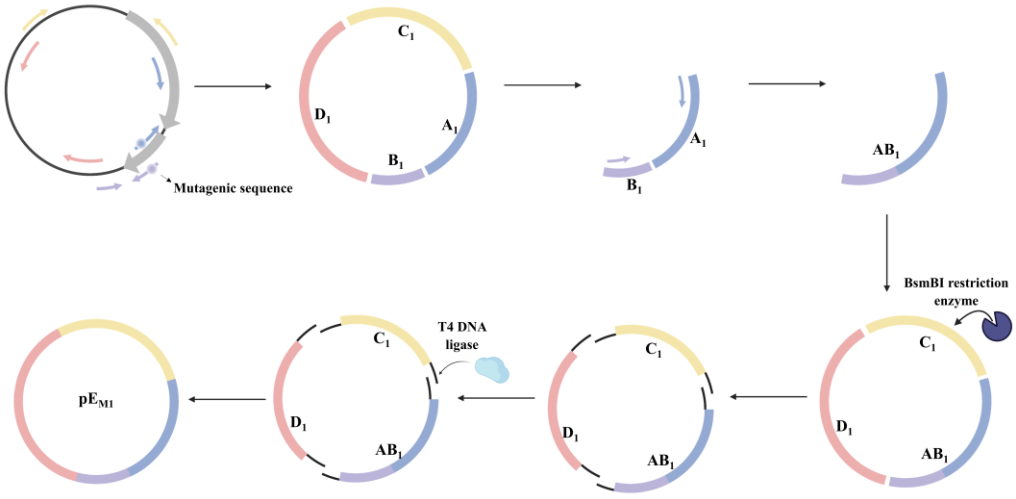


Figure 5.1: Graphical illustration of the mutagenesis approach employed in this study. Created with Biorender.com.

Joining fragments A and B before BsmBI digestion is essential to ensure both ends of the fragment consist of the BsmBI recognition site (5'-CGTCTCN-3'). Fragment A_1 was constructed with a BsmBI site at its 5' end, introduced by the forward primer Syn50-Fwd while fragment B_1 had its BsmBI sequence inserted at the 3' end by the reverse primer Chl425-Rev as shown in **Table 3.2**. Therefore, combining the two fragments resulted in a single fragment AB_1 that consists of a BsmBI site at both 5' and 3' ends. BsmBI is a type-II restriction endonuclease that generates 5'-overhangs by staggered cleaving downstream of the recognition sites. Similar sizes obtained for each fragment (AB_1, C_1, D_1) pre- and post-digestion (**Figures 4.1 (a) – (c)**) assured the absence of star activity of the restriction enzyme. The target plasmid construct was detected on gel above

6.0 kb as a faint band which might be due to the low concentration of product or inadequate ligation. Successful ligation and synthesis of plasmid can be further verified by the transformation of the plasmid into Tet-resistant *E. coli* as only transformed cells will exhibit growth with the Amp resistance gene furnished by the plasmid. In addition, gel electrophoresis of the miniprep products was identified at 6.0 kb which is in accordance with the estimated size of the plasmid. Thus, it can be inferred that both ligation and transformation in this study were successful given the positive results in colony formation and plasmid extraction steps.

The mutational insertion of an additional PstI site into the pE plasmid acts as a crucial validation step for successful mutagenesis in this study. The original pE plasmid carries two PstI sites (5'-CTGCA↓G-3') at 1653^{th} and 2266^{th} bases, respectively and is expected to generate two DNA fragments at 5535 bp and 613 bp (**Figure 5.2 (a)**). On the contrary, pE_{M1} possesses a total of three PstI restriction sites with the newly added restriction site located between the two original sites and will generate three fragments of 5535 bp, 580 bp, and 33 bp (**Figure 5.2 (b)**). Despite only two bands being visible on the gel as shown in **Figure 4.3 (c)**, the second fragment generated from the digestion of pE_{M1} was slightly lower than that of the original pE. The results were consistent with the predicted bands in the simulated enzyme digestion on ApE software (**Figure 5.3**). Hence, it can be assured that the first mutagenesis of pE plasmid is successful.

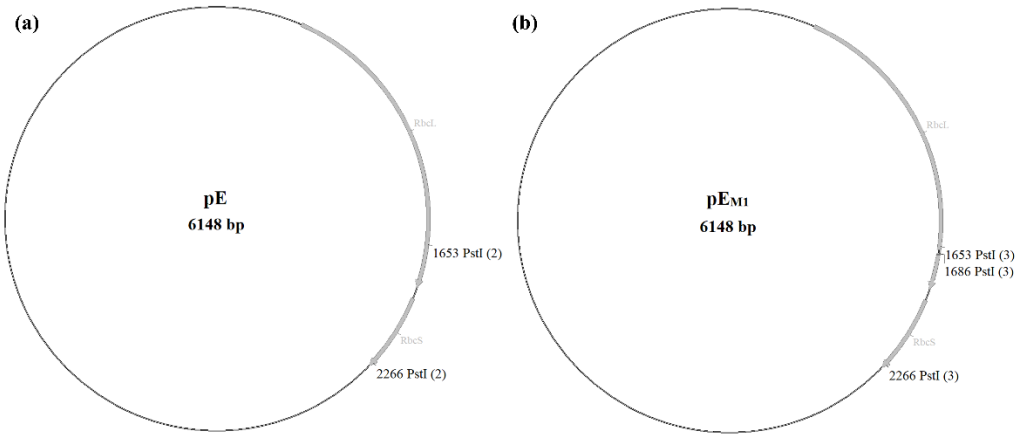


Figure 5.2: Graphical representation of pE (a) and pE_{M1} (b) plasmids with the respective PstI sites generated on ApE software. Two PstI sites are present in the original pE plasmid while an additional PstI is inserted between the originally present sites in pE_{M1}.

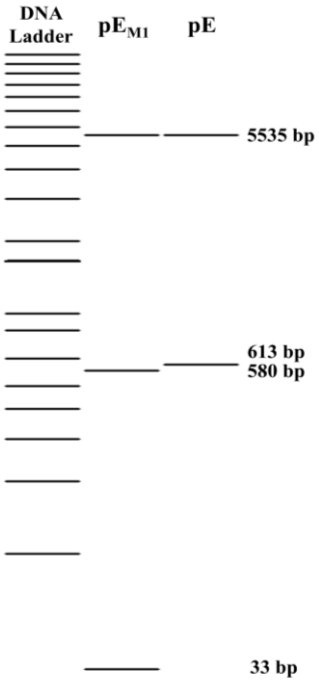


Figure 5.3: Simulated PstI digestion of pE and pE_{M1} plasmids on ApE software. Two bands are predicted at 5535 bp and 613 bp for pE plasmid while three bands at 5535 bp, 580 bp, and 33 bp are expected for pE_{M1} plasmid.

5.2 Second Mutagenesis

The second SDM of this study removes the PstI recognition site that was previously introduced in the first SDM and simultaneously inserts the sequence responsible for substituting Thr for Val at the target site. The nucleotide sequence in mutagenic primer that encodes for Val residue is located around the region where the PstI site was introduced in the first mutagenesis as shown in **Table 3.2**. This mutagenesis step reverts the mutation at position 425 that was introduced in the *Synechococcus* RbcL by a previous study (Koay, Wong and Lim, 2016). Similar to the first SDM, the second mutagenesis began with the four-fragment construction by supplanting the forward and reverse 425-PstI primers with corresponding 425V primers. However, the expected band at 6.0 kb was not observed on the gel but ligation was confirmed with the growth of positive transformants on the Amp plate. This phenomenon is plausibly due to the low concentration of ligated products that are insufficient to be visible on the gel even though they are present. Plasmid extracted from different transformation DF (1X, 1.5X) displayed distinct outcomes in terms of yield and purity. Plasmid extracted from 1.5X DF resulted in more prominent bands (**Figure 4.5 (c)**) and greater DNA concentration along with more consistent purity. Besides, the presence of the mutation was confirmed via restriction digestion with the PstI enzyme of the extracted plasmid. As the PstI site was replaced by the nucleotide sequence responsible for Val residue, the number of available PstI sites in the plasmid was restored to only two and thus will produce only two fragments with the same size as the product of PstI-digested pE plasmid (**Figure 5.4**). Since bands generated by pE_{M2} and pE plasmids upon PstI digestion were analogous

(**Figure 4.5 (d)**), this might imply that the nucleotide sequence encoding Val residue has been substituted into the segment responsible position 425.

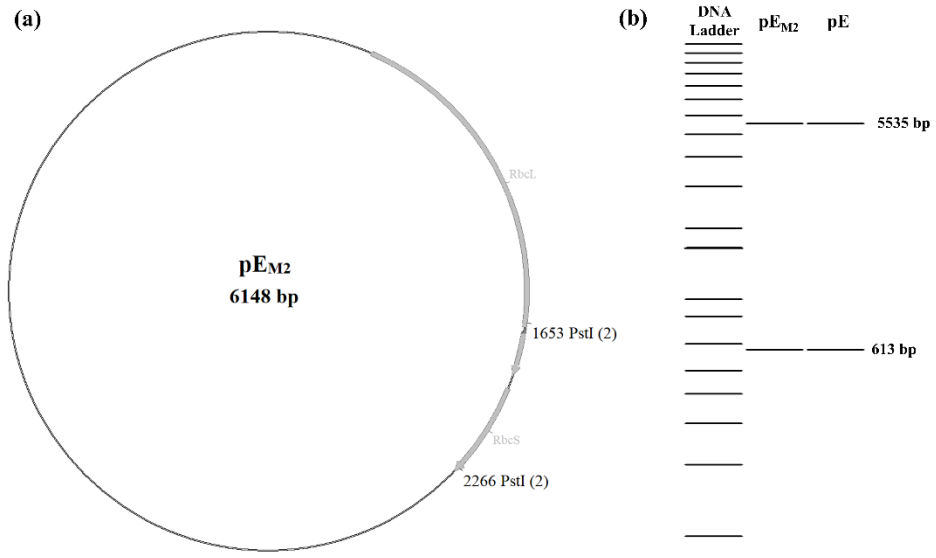


Figure 5.4: Graphical representation of pE$_{M2}$ plasmid with the locations of PstI sites (a). Simulated PstI digestion of pE$_{M2}$ and pE plasmids. Both are predicted to generate similar sizes of fragments.

5.3 Protein Expression

To assess the expression of Rubisco subunits in *E. coli*, total cellular proteins were resolved on SDS-PAGE with larger blots observed at approximately 55 kDa in WT and mutants compared to negative control, indicating the expression of large subunits. Similar results were reported by Koay, Wong and Lim (2016), Yeap et al. (2018) and Ong et al. (2023) where the RbcL was detected at 55 kDa. In addition, the RbcS is expected to appear at position between 10 and 15 kDa which was vaguely observed at the bottom of the gel, below 15 kDa. Similar results were observed in the study conducted by O'Donnelly et al. (2014) where the small subunits were faintly detected below 15 kDa. Expression of the

Rubisco subunits was expected with the reversion of mutation from Thr to Val at position 425 since the previous substitution of *Synechococcus* RbcL involving residue 425 with corresponding *Chlamydomonas* sequence resulted in successful expression regardless of holoenzyme assembly (Koay, Wong and Lim, 2016; Ong et al., 2023).

5.4 Bioinformatics Analysis

In silico analysis complements this study by predicting the influence of the mutation on Rubisco protein in terms of hydrophobicity and stability which are both critical determinants for protein folding and assembly (Dijk, Hoogeveen and Abeln, 2015). Superimposed structures of pE-derived and pE_{M2}-derived Rubisco large subunits showed great resemblance. Structural similarity is important in mutational studies as this suggests that the mutation does not affect the functional properties of the protein (Zhou, Chou and Wong, 2006). The change from Thr residue to Val residue conferred an additional van der Waals force with Pro-165 that provides added support to the overall structure. This might be due to the substitution of the hydroxyl group by the methyl group at the side chain which significantly reduces the distance between Pro-165 and Val-425 as shown in **Figure 4.7** (Li et al., 2018). This is also supported by the prediction made in I-Mutant2.0 where an increase in stability is computed with respect to $\Delta\Delta G$ that takes both the energy change between unfolded and folded states and the aforementioned energy change in the presence of a point mutation into consideration. This indicates that the mutation of Thr to Val at residue 425 is favorable in terms of Rubisco stability which may also imply the possibility

of Rubisco holoenzyme assembly that requires further validation via Native-PAGE.

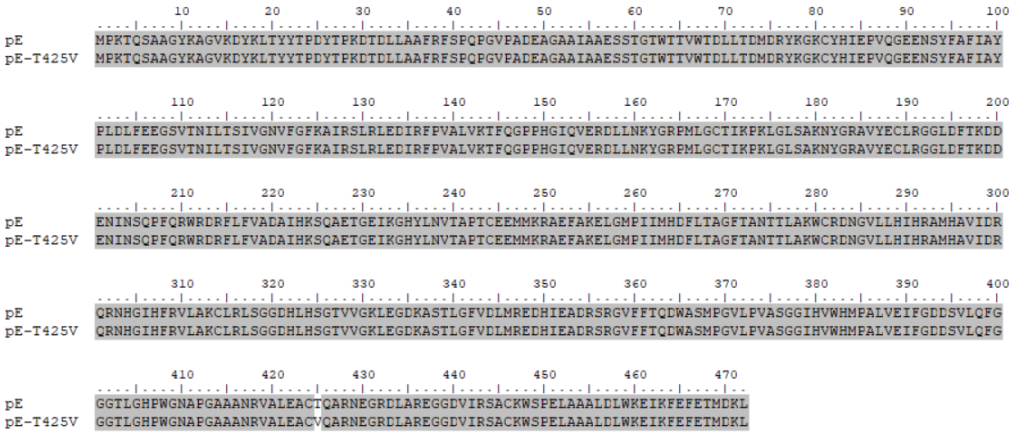


Figure 5.5: Sequence alignment of RbcL translated from pE and pE_{M2} (pE-T425V) plasmids on BioEdit (Hall et al., 2011).

Based on the results shown in **Table 4.2**, the mutation resulted in an overall rise in test-tube stability (instability index), thermostability (aliphatic index), and hydrophobicity (GRAVY). Despite a lower instability index predicted after the mutation, the Rubisco protein is still considered unstable *in vitro* (>40) which might suggest the relative susceptibility to degradation after extraction (Gamage et al., 2019). Besides, the protein also exhibits high thermostability prior to (79.22) and after the mutation (79.83), evaluated based on the number of aliphatic side chains occupying the protein. Studies conducted by Jaiswal et al. (2011), Panda and Chandra (2012), and Enany (2014) reported thermally stable proteins with aliphatic indices that ranged between 61.09 to 83.59, 66.5 to 84.33, and 71.13 to 143.54, respectively. The negative GRAVY values (<0) predicted for both pE-derived and pE_{M2}-derived RbcL suggested the hydrophilic properties of the proteins that are generally found in the chloroplast stroma. The values also conform to the findings by Sevindik (2019) where negative GRAVY values were

also reported for *Sideritis*-derived RbcL proteins. However, RbcL protein bearing the Val residue at position 425 exhibited a relatively high GRAVY value (-0.246), indicating a reduced hydrophilicity which is plausible with the nonpolar side chain of Val in contrast to Thr. The hydropathy plots further reinforced the statement whereby the negative peaks around position 425 displayed reduced peak height when Val residue is present in place of Thr. Yeap et al. (2018) also reported a higher hydropathy index for the sequence with Val at residue 425. Thus, it can be inferred that Val-425 might play an important role in chaperonin binding since hydrophobicity is an essential property for the proper interaction between GroEL and its substrate peptide (Brinker and Hartl, 2001). The nonpolar amino acid residues in the central cavity of GroEL formed the hydrophobic patches that are responsible for the binding of nascent RbcL which also explains the significance of hydrophobicity in RbcL folding (Hayer-Hartl, 2017). Furthermore, Val-425 also serves to confer stability in *Synechococcus* RbcL which favors holoenzyme assembly. Therefore, Val-425 in *Synechococcus* sp. should be avoided as the target residue for future engineering to ensure Rubisco assembly upon modification.

5.5 Future Research Work

5.5.1 *In Vitro* Rubisco Assembly Analysis

In vitro assembly of Rubisco holoenzyme can be assessed by Native-PAGE which separates proteins according to the respective native size and overall surface charge. Due to its non-denaturing nature, Native-PAGE is abundantly employed in protein assembly studies where interactions between subunits

remain intact in a multimeric protein. This technique is suitable for the study of Rubisco L_8S_8 holoenzyme assembly. Studies conducted by Koay, Wong and Lim (2016), Yeap et al. (2018) and Ong et al. (2023) have successfully detected the assembled Rubisco in the extracted proteins on Native-PAGE. Western blotting was also employed in the studies to confirm the presence of Rubisco proteins. Western blot is a common protein separation technique for detecting specific proteins based on the binding of labelled antibodies to the target protein (Mahmood and Yang, 2012). For detecting assembled Rubisco on Western blotting, probing of resolved proteins on the membrane can be done via anti-*Synechococcus* PCC6301 Rubisco IgG of rabbit origin (Parikh et al., 2006).

5.5.2 Rubisco Enzymatic Assay

The presence of Rubisco activity in the sample is a direct approach to determining whether a functional Rubisco has been expressed in the host upon mutation. In addition to assembly, kinetic characterization of the mutant Rubisco can also be conducted by analyzing the kinetic data such as the Michaelis-Menten constant (K_M) and maximum velocity (V_{max}). A non-radioactive method of Rubisco activity assay was proposed by Sulpice et al. (2007) which involved the continuous measurement of the change in nicotinamide adenine dinucleotide (NADH) spectrophotometrically at 340 nm. The rate of reaction can be calculated from the decreasing absorbance values and by using the 3PGA calibration curve as a reference (O'Donnelly et al., 2014). Besides, the initial and maximal velocity in addition to K_M for RuBP of the Rubisco enzyme can also be determined.

5.5.3 Protein Stability Assay

Circular dichroism (CD) is a frequently used technique in the evaluation of conformational stability in purified proteins (Greenfield, 2006). This makes it a possible choice of method for further validation of the computational prediction of change in stability in the mutant RbcL for this study. However, the downside of this method is that protein purity of at least 95% is necessary for CD spectroscopy application. O'Donelly et al (2014) included CD spectroscopy in the study of conformational changes in Rubisco proteins of distinct populations. The mutational effects on the stability of protein can be evaluated via thermal denaturation and renaturation of the mutant and native proteins that are linked to the CD spectrometer (Greenfield, 1999). Temperature for incubation of samples can range between 25 °C and 80 °C. The molar ellipticity (θ) on the CD spectra can be used to calculate the concentration of proteins present in the sample.

CHAPTER 6

CONCLUSIONS

6.1 Overview

The first mutagenesis of this study successfully introduced an additional PstI site to the pE plasmid that originally bore two PstI sites. This is essential as verification steps for successful mutations in both mutagenesis studies in this research. Side-by-side comparison of PstI digested pE and pE_{M1} plasmids showed differences in the lower bands which conformed to the predicted sizes, 613 bp and 580 bp, respectively. This suggests the successful construction of a mutant pE_{M1} plasmid. The mutant plasmid was employed as the template for the second mutagenesis which served to remove the newly inserted PstI site in pE_{M1} and to concurrently introduce the nucleotide sequence responsible for changing Thr to Val at residue 425 in the large subunit of Rubisco. The products of PstI-digested pE_{M2} resolved on agarose gel were the same as that of pE plasmid which suggests that the previously added PstI site had been removed. This may also imply that the nucleotide sequence accounts for Thr residue at position 425 has been substituted by nucleotide sequence encodes for Val residue. Furthermore, both large and small subunits of Val-425-bearing Rubisco were expressed in *E. coli*, supported by the similar band positions detected in wildtype on SDS-PAGE.

Computational analysis of Rubisco stability and change in physicochemical properties when Thr-425 is reverted to Val-425 suggested that Val residue provides additional support to the overall large subunit by engaging in van der

Waals interaction with Pro-165 in addition to other surrounding residues. This in turn raises the stability of the protein as predicted in I-Mutant2.0 which favors folding of the subunit into its native form. Moreover, the elevated hydrophobicity of the protein in the presence of Val-425 as indicated in the GRAVY value and hydropathy plots may explain its significance in chaperonin binding which is an essential process in Rubisco holoenzyme assembly. In conclusion, Val-425 is important in maintaining the stability of *Synechococcus* RbcL and possibly serves as a chaperonin binding site during holoenzyme assembly. Thus, this study suggests that future Rubisco engineering involving *Synechococcus* sp. should avoid supplanting or removing Val-425 residue to ensure the folding and assembly of the subunits into a functional Rubisco.

6.2 Limitations of Study

One of the limitations of this study is the validation of mutation in the mutant plasmid. Despite differences in position observed between the bands generated by PstI digestion of the plasmids (pE, pE_{M1}), it might be exiguous to give a concrete confirmation of the mutation. This may be resolved by carrying out gel electrophoresis with a DNA ladder of a smaller size range (25 bp to 800 bp) which can be used to detect the 33 bp fragment that was predicted to present as one of the products of PstI-digested pE_{M1}. However, this may also require further optimization in the gel electrophoresis procedure including gel percentage, run time, supply voltage and current which was limited by the time constraint for this study. Besides, in the second mutagenesis, the pE plasmid may sometimes be expressed in the *E. coli* instead of the expected pE_{M2} plasmid which both

demonstrate the same outcome in PstI digestion. Definitive mutation can be confirmed by DNA sequencing of the mutant plasmid extracted from transformed *E. coli*. Although DNA sequencing has been the most reliable approach for sequence determination, it is nonetheless a time-consuming and costly procedure.

REFERENCES

Aigner, H. et al., 2017. Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science*, 358, pp. 1272 – 1278.

Alam, M.M. et al., 2023. Beta, delta, and omicron, deadliest among SARS-CoV-2 variants: A computational repurposing approach. *Evolutionary Bioinformatics*, 19, pp. 1 – 12.

Alshrari, A.S. et al., 2021. Bioinformatics analysis of rhinovirus capsid proteins VP1-4 sequences for cross-serotype vaccine development. *Journal of Infection and Public Health*, 14, pp. 1603 – 1611.

Andersson, I., 2008. Catalysis and regulation in Rubisco. *Journal of Experimental Botany*, 59(7), pp. 1555 – 1568.

Andersson, I. and Backlund, A., 2008. Structure and function of Rubisco. *Plant Physiology and Biochemistry*, 46, pp. 275 – 291.

Balamurugan, J.P. et al., 2021. Efficient biomass-endopolysaccharide production from an identified wild-Serbian *Ganoderma applanatum* strain BGS6Ap mycelium in a controlled submerged fermentation. *Biocatalysis and Agricultural Biotechnology*, 37, 102166.

Banda, D.M. et al., 2020. Novel bacterial clade reveals origin of form I Rubisco. *Nature Plants*, 6, pp. 1158 – 1166.

Bathellier, C., Yu, L.-J., Farquhar, G.D. and Tcherkez, G., 2020. Ribulose 1,5-bisphosphate carboxylase/oxygenase activates O_2 by electron transfer. *Proceedings of the National Academy of Sciences*, 117(39), pp. 24234 – 24242.

Bracher, A. et al., 2015. Structural analysis of the Rubisco-assembly chaperone RbcX-II from *Chlamydomonas reinhardtii*. *PLoS ONE*, 10(8), e0135448.

Bracher, A., Starling-Windhof, A., Hartl, F.U. and Hayer-Hartl, M., 2011. Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nature Structural & Molecular Biology*, 18(8), pp. 875 – 881.

Brinker, A. and Hartl, F.U., 2001. Chaperonins. In: Brenner, S. and Miller, J.H. (eds.). *Encyclopedia of Genetics*. New York: Academic Press Inc., pp. 324 – 325.

Busch, F.A., 2020. Photorespiration in the context of Rubisco biochemistry, CO_2 diffusion and metabolism. *The Plant Journal*, 101, pp. 919 – 939.

Carugo, O., 2003. How root-mean-square distance (r.m.s.d.) values depend on the resolution of protein structures that are compared. *Journal of Applied Crystallography*, 36, pp. 125 – 128.

Chang, S. et al., 2023. Engineering aldehyde dehydrogenase PaALDH70140 from *Pseudomonas aeruginosa* PC-1 with improved catalytic properties for 5-hydroxymethyl-2-furancarboxylix acid synthesis. *Biochemical Engineering Journal*, 192, 108835.

Conlan, B. et al., 2019. BSD2 is a Rubisco-specific assembly chaperone, forms intermediary hetero-oligomeric complexes, and is nonlimiting to growth in tobacco. *Plant Cell Environment*, 42, pp. 1287 – 1301.

Cummins, P.L., Kannappan, B. and Gready, J.E., 2018. Directions for optimization of photosynthetic carbon fixation: RuBisCO's efficiency may not be so constrained after all. *Frontiers in Plant Science*, 9, 183.

Davis, M.W. and Jorgensen, E.M., 2022. ApE, A plasmid Editor: A freely available DNA manipulation and visualization program. *Frontiers in Bioinformatics*, 2, 818619.

Dower, W.J., Miller, J.F. and Ragsdale, C.W., 1988. High efficiency transformation of *E.coli* by high voltage electroporation. *Nucleic Acids Research*, 16(13), pp. 6127 – 6145.

Duarte, A.J., Ribeiro, D., Moreira, L. and Amaral, O., 2018. In silico analysis of missense mutations as a first step in functional studies: Examples from two sphingolipidoses. *International Journal of Molecular Sciences*, 19, 3409.

Enany, S., 2014. Structural and functional analysis of hypothetical and conserved proteins of *Clostridium tetani*. *Journal of Infection and Public Health*, 7(4), pp. 296 – 307.

Fakruddin, M. et al., 2013. Critical factors affecting the success of cloning, expression, and mass production of enzymes by recombinant *E. coli*. *ISRN Biotechnology*, 590587.

Fristedt, R. et al., 2018. RAF2 is a RuBisCO assembly factor in *Arabidopsis thaliana*. *The Plant Journal*, 94, pp. 146 – 156.

Frolov, E.N., Kublanov, I.V., Toshchakov, S.V. and Chernyh, N.A., 2019. Form III RuBisCO-mediated transaldolase variant of the Calvin cycle in a chemolithoautotrophic bacterium. *Biological Sciences*, 116(37), pp.18638 – 18646.

Gamage, D.G., Gunaratne, A., Periyannan, G.R. and Russell, T.G., 2019. Applicability of instability index for in vitro protein stability prediction. *Protein & Peptide Letters*, 26(5), pp. 339 – 347.

Gasteiger, E. et al., 2005. Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (ed). *The Proteomics Protocols Handbook*. New Jersey: Humana Press Inc., pp. 571 – 607.

Green, M.R. and Sambrook, J., 2020. Transformation of *Escherichia coli* by electroporation. *Cold Spring Harbor Protocols*, 6, pp. 232 – 238.

Greenfield, N.J., 1999. Applications of circular dichroism in protein and peptide analysis. *Trends in Analytical Chemistry*, 18(4), pp. 236 – 244.

Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, 1, pp. 2876 – 2890.

Gruber, A.V. and Feiz, L., 2018. Rubisco assembly in the chloroplast. *Frontiers in Molecular Biosciences*, 5(24).

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp. 95 – 98.

Hauser, T. et al., 2015. Structure and mechanism of the Rubisco-assembly chaperone Raf1. *Nature Structural & Molecular Biology*, 22, pp. 720 – 728.

Hayer-Hartl, M., 2017. From chaperonins to Rubisco assembly and metabolic repair. *Protein Science*, 26(12), pp. 2324 – 2333.

Heldt, H.-W. and Piechulla, B., 2011. 6 - The Calvin cycle catalyzes photosynthetic CO_2 assimilation. In: *Plant Biochemistry*. 4th ed. New York: Academic Press, pp. 163 – 191.

Huang, F. et al., 2020. Rubisco accumulation factor 1 (Raf1) plays essential roles in mediating Rubisco assembly and carboxysome biogenesis. *Proceedings of the National Academy of Sciences*, 117(29), pp. 17418 – 17428.

Jaiswal, A. et al., 2011. Comparative analysis of human matric metalloproteinases: Emerging therapeutic targets in diseases. *Bioinformation*, 6(1), pp. 23 – 30.

Jia, H. et al., 2022. Highly efficient and simple SSPER and rrPCR approaches for the accurate site-directed mutagenesis of large and small plasmids. *New Biotechnology*, 72, pp. 22 – 28.

Kabiraj, A., Laha, A., Panja, A.S. and Bandopadhyay, R., 2023. In silico comparative structural and functional analysis of arsenite methyltransferase from bacteria, fungi, fishes, birds, and mammals. *Journal of Genetic Engineering and Biotechnology*, 21, 64.

Koay, T.W., Wong, H.L. and Lim, B.H., 2016. Engineering of chimeric eukaryotic/bacterial Rubisco large subunits in *Escherichia coli*. *Genes & Genetic Systems*, 91(3), pp. 139 – 150.

Kreel, N.E. and Tabita, F.R., 2015. Serine 363 of a hydrophobic region of archaeal ribulose-1,5-bisphosphate carboxylase/oxygenase from *Archaeoglobus fulgidus* and *Thermococcus kodakaraensis* affects CO_2/O_2 substrate specificity and oxygen sensitivity. *PLoS ONE*, 10(9), e0138351.

Kubis, A. and Bar-Even, A., 2019. Synthetic biology approaches for improving photosynthesis. *Journal of Experimental Botany*, 70(5), pp. 1425 – 1433.

Kumari, P. and Rameshwari, R., 2022. In silico mutational analysis to identify the role and pathogenicity of BCL-w missense variants. *Journal of Genetic Engineering and Biotechnology*, 20, 120.

Kyte, J. and Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), pp. 105 – 132.

Lessard, J.C., 2013. Chapter twenty seven - transformation of *E. coli* via electroporation. In: Lorsch, J. (ed). *Methods in Enzymology*. New York: Academic Press Inc., pp. 321 – 327.

Li, J. et al., 2018. Direct observation of CH/CH van der Waals interactions in proteins by NMR. *Journal of the American Chemical Society*, 140(9), pp. 3194 – 3197.

Li, Q. et al., 2022. Structural insights into cyanobacterial RuBisCO assembly coordinated by two chaperones Raf1 and RbcX. *Cell Discovery*, 8, 93.

Lin, M.T. et al., 2014. A faster Rubisco with potential to increase photosynthesis in crops. *Nature*, 513, pp. 547 – 550.

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*, pp. 1 – 10.

Liu, F. et al., 2023. Site-directed mutagenesis improves the practical application of L-glutamic acid decarboxylase in *Escherichia coli*. *Engineering in Life Sciences*, 23, e2200064.

Lundqvist, T. and Schneider, G., 1991. Crystal structure of the ternary complex of ribulose-1,5-bisphosphate carboxylase, magnesium (II) and activator carbon dioxide at 2.3-.ANG. resolution. *Biochemistry*, 30(4), pp. 904 – 908.

Mackinder, L.C.M., 2018. The *Chlamydomonas* CO_2-concentrating mechanism and its potential for engineering photosynthesis in plants. *New Phytologist*, 217, pp. 54 – 61.

Mahmood, T. and Yang, P.-C., 2012. Western blot: Technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, 4(9), pp. 429 – 434.

Math, R.K., Mudennavar, N., Javaregowda, P.K. and Savanur, A., 2022. In silico comparative analysis of the functional, structural, and evolutionary properties of SARS-CoV-2 variant spike proteins. *JMIR Bioinformatics and Biotechnology*, 3(1), e37391.

Miller, S.R., McGuirl, M.A. and Carvey, D., 2013. The evolution of RuBisCO stability at the thermal limit of photoautotrophy. *Molecular Biology and Evolution*, 30(4), pp. 752 – 760.

Mosad, S.M. et al., 2023. Molecular and pathological investigation of avian reovirus (ARV) in Eqypt with the assessment of the genetic variability of field strains compared to vaccine strains. *Frontiers in Microbiology*, 14, 1156251.

Mueller-Cajar, O. and Whitney, S.M., 2008. Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli*. *Biochemical Journal*, 414(2), pp. 205 – 214.

Muhie, S.H., 2022. Optimization of photosynthesis for sustainable crop production. *CABI Agriculture and Bioscience*, 3, 50.

Nakano, T., Ashida, H., Mizohata, E., Matsumura, H. and Yokota, A., 2010. An evolutionally conserved Lys122 is essential for function in *Rhodospirillum rubrum* bona fide RuBisCO and *Bacillus subtilis* RuBisCO-like protein. *Biochemical and Biophysical Research Communications*, 392, pp. 212 – 216.

Newman, J., Brändén, C.I. and Jones, T.A., 1993. Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Synechococcus* PCC6301. *Acta Crystallography*, D49, pp. 548 – 560.

Nishitani, Y. et al., 2010.Structure-based catalytic optimization of a Type III Rubisco from a hyperthermophile. *Protein Structure and Folding*, 285(10), pp. 39339 – 39347.

Occhialini, A. et al., 2016. Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. *The Plant Journal*, 85(1), pp. 148 – 160.

O'Donnelly, K. et al., 2014. Isolation and kinetic characterization of hydrophobically distinct populations of form I Rubisco. *Plant Methods*, 10, 17.

Ong, W.C., Yap, M.L., Wong, H.L. and Lim, B.H., 2023. Pinpointing *Synechococcus* Rubisco large subunit sections involved in heterologous holoenzyme formation in *Escherichia coli*. *Malaysian Journal of Microbiology*, 19(1), pp. 47 – 54.

Panda, S. and Chandra, G., 2012. Physicochemical characterization and functional analysis of some snake venom toxin proteins and related non-toxin proteins of other chordates. *Bioinformation*, 8(18), pp. 891 – 896.

Parikh, M.R., Greene, D.N., Woods, K.K. and Matsumura, I., 2006. Directed evolution of RuBisCO hypermorphs through genetic selection in engineered *E.coli*. *Protein Engineering, Design and Selection*, 19(3), pp. 113 – 119.

Pope, B. and Kent, H.M., 1996. High efficiency 5 min transformation of *Escherichia coli*. *Nucleic Acids Research*, 24(3), pp. 536 – 537.

Rasineni, G.K., Loh, P.C. and Lim, B.H., 2017. Characterization of *Chlamydomonas* Ribulose-1,5-bisphosphate carboxylase/oxygenase variants mutated at residues that are post-translationally modified. *Biochimica et Biophysica Acta*, pp. 79 – 85.

Rosano, G.L. and Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5, 172.

Saini, S. et al., 2018. In silico mutational analysis and identification of stability centers in human interleukin-4. *Molecular Biology Research Communications*, 7(2), pp. 67 – 76.

Sevindik, E., 2018. Genome-wide identification and analysis of Rubisco large subunit proteins in *Morus* L. (Moraceae) species. *Genetika*, 50(2), pp. 603 – 616.

Sevindik, E., 2019. Comparative and phylogenetic analysis of Rubisco large subunit (*rbc*L) proteins in some *Sideritis* L. (Lamiaceae) species. *Genetika*, 51(1), pp. 69 – 80.

Shahzan, M.S., Girija, A.S.S., Priyadharsini, J.V., 2019. A computational study targeting mutated *L321F* of *ERG11* gene in *C. albicans*, associated with fluconazole resistance with bioactive compounds from *Acacia nilotica*. *Journal de Mycologie Médicale*, 29, pp. 303 – 309.

Sharwood, R.E., 2017. Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. *New Phytologist*, 213(2), pp. 494 – 510.

Shi, X. and Bloom, A., 2021. Photorespiration: the futile cycle?. *Plants (Basel)*, 10(5), 908.

Stec, B., 2012. Structural mechanism of RuBisCO activation by carbamylation of the active site lysine. *Proceedings of the National Academy of Sciences*, 109(46), pp. 18785 – 18790.

Subbiah, T., Maheswaran, G. and Raja, V.K., 2023. In silico protein analysis of *Phoenix dactylifera* (date palm) by a bioinformatics approach. *ACS Agricultural Science & Technology*, 3, pp. 528 – 534.

Sulpice, R. et al., 2007. Description and applications of a rapid and sensitive non-radioactive microplate-based assay for maximum and initial activity of D-ribulose-1,5-bisphosphate carboxylase/oxygenase. *Plant, Cell & Environment*, 30(9), pp. 1163 – 1175.

Tabita, F.R. et al., 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *Journal of Experiment Botany*, 59(7), pp. 1515 – 1524.

Tamura, H. et al., 2009. Structure of the apo decarbamylated form of 2,3-diketo-5-methylthiopentyl-1-phosphate enolase from *Bacillus subtilis*. *Acta Crystallography*, D65, pp. 942 – 951.

Tcherkez, G., 2016. The mechanism of Rubisco-catalysed oxygenation. *Plant, Cell and Environment*, 38, pp. 983 – 997.

Tommasi, I.C., 2021. The mechanism of Rubisco catalyzed carboxylation reaction: chemical aspects involving acid-base chemistry and functioning of the molecular machine. *Catalysts*, 11, 813.

United Nations Department of Economic and Social Affairs, Population Division, 2022. *World population prospects 2022: Summary of results* (UN DESA/POP/2022/TR/NO. 3). New York: United Nations Publication.

van Dijk, E., Hoogeveen, A. and Abeln, S., 2015. The hydrophobic temperature dependence of amino acids directly calculated from protein structures. *PLoS Computational Biology*, 11(5), e1004277.

van Lun, M., Hub, J.S., van der Spoel, D. and Andersson, I., 2014. CO_2 and O_2 distribution in Rubisco suggests the small subunit functions as a CO_2 reservoir. *Journal of the American Chemical Society*, 136, pp. 3165 – 3171.

Wada, S., Miyake, C., Makino, A. and Suzuki, Y., 2020. Photorespiration coupled with CO_2 assimilation protects photosystem I from photoinhibition under moderate poly(ethylene glycol)-induced osmotic stress in rice. *Frontiers in Plant Science*, 11, 1121.

Walker, B.J., VanLoocke, A., Bernacchi, C.J. and Ort, D.R., 2016. The costs of photorespiration to food production now in the future. *Annual Review of Plant Biology*, 67, pp. 107 – 129.

Walker, K.W. and King, J.D., 2023. Site-directed mutagenesis. In: Bradshaw, R.A., Hart, G.W. and Stahl, P.D. (eds.). *Encyclopedia of Cell Biology Second Edition*. New York: Academic Press Inc., pp. 161 – 169.

Waterhouse, A. et al., 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46 (Web Server issue), pp. W296 – W303.

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

Wietrzynski, W., Traverso, E., Wollman, F.-A. and Wostrikoff, K., 2021. The state of oligomerization of Rubisco controls the rate of synthesis of the Rubisco large subunit in *Chlamydomonas reinhardtii*. *The Plant Cell*, 33, pp. 1706 – 1727.

Wilson, R.H., Martin-Avila, E., Conlan, C. and Whitney, S.M., 2018. An improved *Escherichia coli* screen for Rubisco identifies a protein-protein interface that can enhance CO_2-fixation kinetics. *Journal of Biological Chemistry,* 293(1), pp. 18 – 27.

Wilson, R.H., Thieulin-Pardo, G., Hartl, F.-U. and Hayer-Ha, M., 2019. Improved recombinant expression and purification functional plant Rubisco. *FEBS Letters*, 593, pp. 611 – 621.

Yang, Z., Chen, Z. and Zhang, Y., 2022. A simple and economical site-directed mutagenesis method for large plasmids by direct transformation of two overlapping PCR fragments. *BioTechniques*, 73(5), pp. 227 – 233.

Yeap, Y.H., Koay, T.W., Wong, H.L. and Lim, B.H., 2018. Dissection of *Synechococcus* Rubisco large subunit sections involved in holoenzyme formation in *Escherichia coli* by combinatorial section swapping and sequence analyses. *Sains Malaysiana*, 47(10), pp. 2269 – 2289.

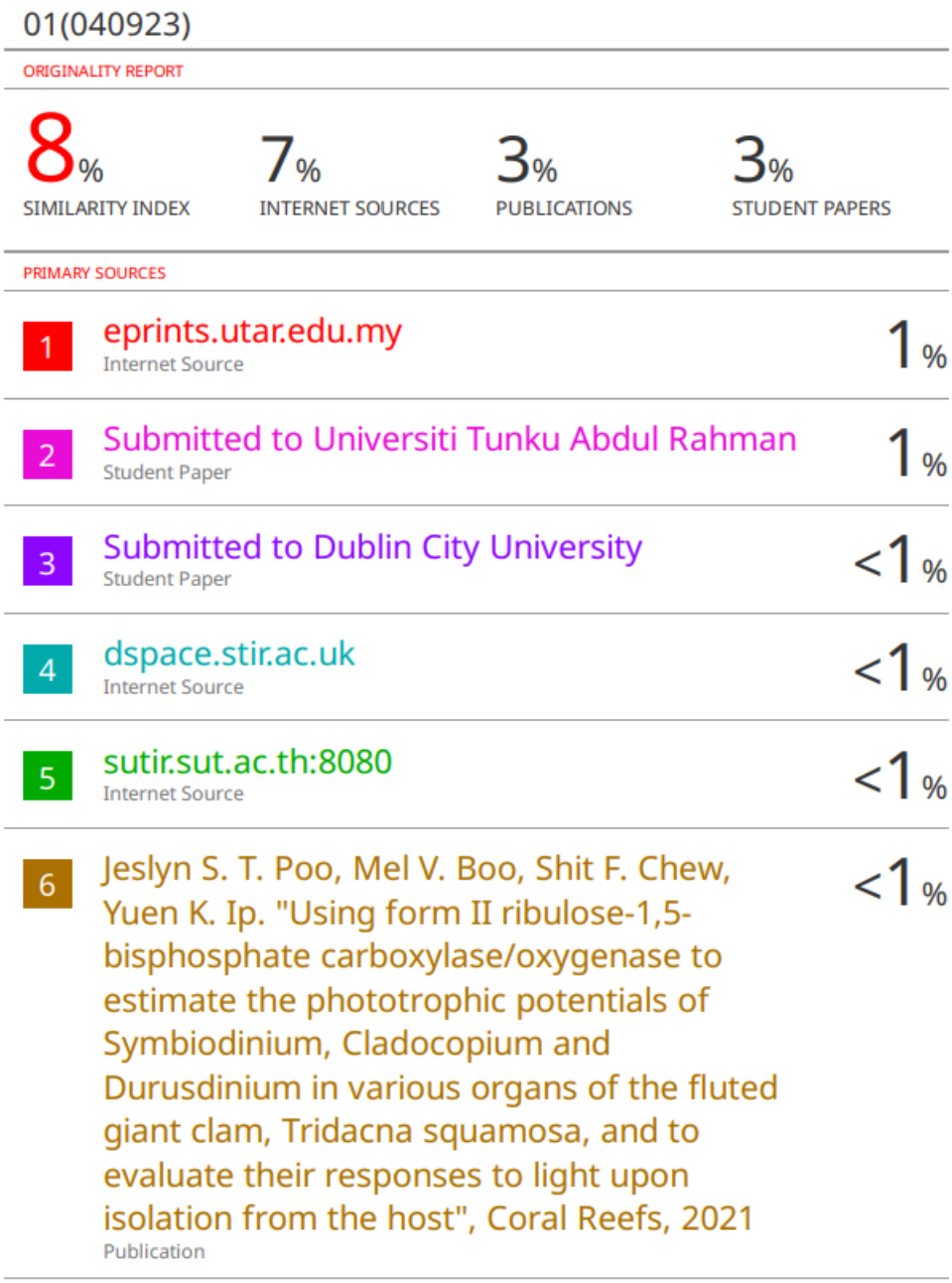
Zhang, K. et al., 2021. A high-efficiency method for site-directed mutagenesis of large plasmids based on large DNA fragment amplification and recombinatorial ligation. *Scientific Reports*, 11, 10454.

Zhou, X., Chou, J. and Wong, S.T.C., 2006. Protein structure similarity from principle component correlation analysis. *BMC Bioinformatics*, 7, 40.

APPENDICES

Appendix A

Turnitin originality report



01(040923)

ORIGINALITY REPORT

8%	7%	3%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	eprints.utar.edu.my Internet Source	1%
2	Submitted to Universiti Tunku Abdul Rahman Student Paper	1%
3	Submitted to Dublin City University Student Paper	<1%
4	dspace.stir.ac.uk Internet Source	<1%
5	sutir.sut.ac.th:8080 Internet Source	<1%
6	Jeslyn S. T. Poo, Mel V. Boo, Shit F. Chew, Yuen K. Ip. "Using form II ribulose-1,5-bisphosphate carboxylase/oxygenase to estimate the phototrophic potentials of Symbiodinium, Cladocopium and Durusdinium in various organs of the fluted giant clam, Tridacna squamosa, and to evaluate their responses to light upon isolation from the host", Coral Reefs, 2021 Publication	<1%

Appendix B

Supervisor's comments form

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 1	Effective Date: 3/10/2019	Page No.: 1of 1

UTAR
UNIVERSITI TUNKU ABDUL RAHMAN
Wholly owned by UTAR Education Foundation

FACULTY OF SCIENCE

Full Name(s) of Candidate(s)	CLARA WONG CHIA CI
ID Number(s)	2006159
Programme / Course	BACHELOR OF SCIENCE (HONS) BIOCHEMISTRY
Title of Final Year Project	DETERMINATION OF THE ROLE OF *SYNECHOCOCCUS* RUBISCO RESIDUE VAL-425 IN CHIMERAS

Similarity	**Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)**
Overall similarity index: 8 % **Similarity by source** Internet Sources: 7 % Publications: 3 % Student Papers: 3 %	
Number of individual sources listed of more than 3% similarity: -	

Parameters of originality required and limits approved by UTAR are as follows:
(i) Overall similarity index is 20% and below, and
(ii) Matching of individual sources listed must be less than 3% each, and
(iii) Matching texts in continuous block must not exceed 8 words
Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Lim Boon Hoe
Signature of Supervisor
Name: DR LIM BOON HOE

Date: 06 SEPT 2023

Signature of Co-Supervisor
Name: _____

Date: _____