CONSTRUCTION OF CHIMERIC RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT-CHLORAMPHENICOL ACETYLTRANSFERASE (*rbc*L-*CAT*) GENE

LOW YIN HUIE

BACHELOR OF SCIENCE (HONS) MICROBIOLOGY

FACULTY OF SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

MAY 2014

CONSTRUCTION OF CHIMERIC RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT-CHLORAMPHENICOL ACETYLTRANSFERASE (*rbc*L-*CAT*) GENE

By

LOW YIN HUIE

A project report submitted to the Department of Biological Science

Faculty of Science

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements for the degree of

Bachelor of Science (Hons) Microbiology

May 2014

ABSTRACT

CONSTRUCTION OF CHIMERIC RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT-CHLORAMPHENICOL ACETYLTRANSFERASE (*rbc*L-*CAT*) GENE

LOW YIN HUIE

Rubisco catalyzes the primary step of photosynthesis that produces usable sugar for plant growth. However, due to its slow and non-specific catalysis, extensive efforts have been made to re-engineer the enzyme to enhance agricultural productivity. Directed evolution or engineering experiment requires a rapid screening method to determine functionality that would enhance mutants screening. This study aims to construct a *rbc*L-*CAT* chimeric gene which can be used for the establishment of a model system to distinguish between functional rbcL from non-functional form by inferring from the functionality of CAT protein that confers chloramphenicol resistance. In this study, *rbc*L from *Chlamydomonas* and S*ynechococcus* were used as eukaryotic and prokaryotic *rbc*L, respectively. Unlike prokaryotic rbcL, eukaryotic rbcL protein is unable to assemble properly in *Escherichia coli*. Sequences encoding rbcL of *Chlamydomonas* and *Synechococcus*, and CAT were separately amplified by PCR, and the *rbc*L-*CAT* fusions were constructed by overlap extension PCR. The chimeric *rbc*L-*CAT* constructs were separately ligated into the cloning

vector, pENTR/D-TOPO and transformed into competent *E. coli*. Colonies on selective plates were screened by colony PCR. Plasmids isolated from positive clones were confirmed with DNA sequencing and the inserted genes further transferred into the expression vector, pDEST^{TM}17. Recombinant plasmid pENTR-*rbc*L-*CAT* derived from *Chlamydomonas* was obtained. Recombinant plasmids

pENTR-*rbc*L derived from both *Chlamydomonas* and *Synechococcus* were also successfully obtained, and the inserted *rbc*L gene subcloned into pDEST^{TM}17. Chimeric *rbc*L-*CAT* derived from *Synechococcus* was constructed, but the cloning of *rbc*L-CAT into pENTR^{TM}/D-TOPO was unsuccessful. In future study, recombinant pDEST^{TM}17 carrying the chimeric gene constructs will be expressed in *E. coli* to determine the concentration of chloramphenicol that can distinguish between functional and non-functional rbcL. When this functional selection system is established, sequence swapping between DNA libraries of *Chlamydomonas* and *Synechococcus rbc*L genes can be performed and selection of functional rbcL can be performed based on chloramphenicol resistance.

ACKNOWLEDGEMENTS

This project would not have been possible without the support and guidance of everyone around me. In particular, I would like to express my very great appreciation to my supervisor Associate Professor Dr. Wong Hann Ling and co-supervisor Assistant Professor Dr. Lim Boon Hoe for their willingness to give their time so generously on guiding me throughout the course of FYP project. I would also like to take this opportunity to express my gratitude to beloved Mum for constant encouragement and support, love you.

DECLARATION

I hereby declare that the project report 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 degree at UTAR or other institutions.

Low Yin Huie

APPROVAL SHEET

This project report entitled **"CONSTRUCTION OF CHIMERIC RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT-CHLORAMPHENICOL ACETYLTRANSFERASE (*rbc*L-*CAT*) GENE"** was prepared by LOW YIN HUIE and submitted as partial fulfilment of the requirements for degree of Bachelor of Science (Hons) in Microbiology at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Assoc. Prof. Dr. WONG HANN LING)

Date:……..

Supervisor

Department of Biological Science

Faculty of Science

Date: _____

PERMISSION SHEET

It is hereby certified that **LOW YIN HUIE** (ID No: **11ADB04711**) has completed this final year project entitled **"CONSTRUCTION OF CHIMERIC RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT–CHLORAMPHENICOL ACETYLTRANSFERASE GENE"** under the supervision of Assoc. Prof. Dr. Wong Hann Ling (supervisor) and Assist. Prof. Dr. Lim Boon Hoe (co-supervisor) from the Department of Biological Science and Department of Chemical Science, respectively, of 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 into to the UTAR community and public.

Yours truly,

_____

(LOW YIN HUIE)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

µL	Microlitre
^{14}C	Radioactive isotope of carbon
2x YT	2x Yeast extract and Tryptone medium
∞	Infinity
bp	Base pair
CAT	Chloramphenicol acetyltransferase
CFU	Colony forming unit
CO_2	Carbon dioxide
ddH_2O	Deionized water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleotide acid
dNTPs	Deoxynucleotide triphosphates
EDTA	Ethylnediaminetetraacetic acid
g	Acceleration of gravity
GapA	Glyceraldehyde-3-phosphate dehydrogenase
L	Litre
luxCt	Chloroplast luciferase gene
M	Molar
Mg^{2+}	Magnesium ion
$MgCI_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
mL	Millilitre
mM	Millimolar
NaCl	Sodium chloride

$NaH^{14}CO_3$	Radiolabelled sodium bicarbonate
ng	Nanogram
O_2	Oxygen molecule
OD	Optical Density
PCR	Polymerase Chain Reaction
PRK	Phosphoribulokinase
psi	Pounds per square inch
rbcL	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
rbcS	Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
RNA	Ribonucleotide acid
rpm	Revolution per minute
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	Ribulose-1,5-bisphosphate
S.O.C	Super Optimal broth with Catabolic repressor
sec	Seconds
T_a	Annealing temperature
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA

CHAPTER 1

INTRODUCTION

1.1 Rubisco and Its Limitation

Ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39, Rubisco), the most abundant protein on earth (Ellis 1979) can comprise up to fifty percent of total soluble protein of certain leaf or microbes (Andersson and Backlund 2008). It is found in all photosynthetic organisms within all three domains of life (Tabita 1999) and responsible for global assimilation of atmospheric carbon dioxide into biosphere by catalyzing carboxylation of ribulose-1,5-bisphosphate (RuBP), first step of Calvin cycle. In fact, virtually every carbon containing molecules in biosphere are derived from Rubisco catalysis. However, despite of its crucial role in sustaining life, Rubisco is perhaps the slowest metabolic enzyme because it catalyzes only few RuBP per second at atmospheric CO_2 (Parikh et al 2006), while most of the enzyme able to catalyze about one thousand per second (Alberts 2002). Besides, carboxylation activity of Rubisco is compromised by competing oxygen, which drives photorespiration, the opposite reaction of photosynthesis, where oxygen is used but carbon dioxide is released.

1.2 Engineering Rubisco

Identification of Rubisco had caught more than usual interest as its enzymatic activity is regarded as the rate limiting step on photosynthesis (Greene, Whitney and Matsumura 2007). Hence, numerous studies had been carried out with the ultimate aim of engineering Rubisco to alter the ratio in favor of carboxylase in order to increase agricultural productivity. Furthermore, extensive efforts of researchers on manipulating Rubisco had been sparked by the finding that Rubisco carboxylation kinetic varies among species (Esquível , Anwaruzzaman and Spreitzer 2002). Kinetic of Rubisco is measured based on its ability in discriminating CO_2 from O_2 which also termed as specificity factor (Greene, Whitney and Matsumura 2007). It was found that Rubisco enzyme in thermophilic red algae *Galderia parita* have the highest noted specificity factor of 328 (Sugawara et al 1999) whereas the rest of reported Rubisco specificity varies from 5 to 100 (Andersson and Backlund 2008). Researchers had also suggested that substituting land-plant Rubisco with kinetically superior variant which already exists in nature would boost carbon fixation of C3 plant by 27% (Zhu, Long and Ort 2010; Whitney et al 2001).

1.3 Functional Screening Method for Rubisco

Although numerous crystals structure or site-directed mutants had been determine and form a basis attempts to understand functional details of Rubisco, a great deal of progress is still needed in order to elucidate the molecular basis that govern variant specificity in Rubisco. However, these efforts have so far been stiffed by the inability to reconstitute the enzyme *in vitro* and hurdle by

the current slow mutant screening methods. A highly sensitive, precise and dynamic screening assay is important in determining the outcome of any mutagenesis study (Parikh et al 2006). Photosynthetic microbial functional selection system for Rubisco had been developed and widely used based upon complementation of auxotrophic mutants such as *Chlamydomonas reinherdti* (Du and Spreitzer 2000), *Anacystis nidulans* (*Synechococcus)* (Greene, Whitney and Matsumura 2007) and *Rhodobacter capsulatus* (Paoli, Vichivanives and Tabita 1998). Recently, two groups of researchers (Parikh et al 2006; Mueler-Cajar, Morell and Whitney 2007) had demonstrated a high throughput and sensitive screening assay which is based on heterotrophic Rubisco-dependent *E. coli*. *E. coli*-based screening assay is used in directed evolution study of Rubisco because of the superior transformation efficiency of *E. coli*, which is at least three orders of magnitude higher than photosynthetic microbes (Parikh et al 2006).

1.4 Significance and Experimental Design of This Study

Instead of employing a screening method that select for metabolically active Rubisco mutants, it is suggested that a method that utilize selectable marker may be useful for screening of functional Rubisco. Rubisco large subunit, rbcL derived from *Chlamydomonas* and *Synechococcus* were used as eukaryotic and prokaryotic rbcL, respectively. Although both prokaryotic and eukaryotic rbcL can be expressed in *E. coli*, the bacterial chaperone only folds and assembles rbcL derived from prokaryotes. In this project, *CAT* gene, which codes for the enzyme chloramphenicol acetyltransferase (EC 2.3.1.28) that detoxifies

antibiotic chloramphenicol, was used as selectable marker. *CAT* gene was fused in frame at downstream of *rbc*L gene with a GS rich linker. A 5'CACC was also placed at the upstream of *rbc*L gene to facilitate subsequent cloning step. The nucleotide sequence of *rbc*L derived from *Chlamydomonas* and *Synechococcus*, *CAT* genes and GS linker were shown in Appendix A, B and C, respectively. It is postulated that *E. coli* expressing a functional rbcL-CAT would confer chloramphenicol resistance. On the other hand, rbcL derived from eukaryotes are unable to be folded properly and might have distorted structures, which may subsequently affect the folding of the CAT that is downstream. Denatured proteins with altered conformations are also more sensitive to proteases and have increased susceptibility to proteolysis (Goldberg 1972). Therefore, *E. coli* carrying the non-functional rbcL-*CAT* are expected to have limited resistance or are susceptible to chloramphenicol.

Apart from constructing Rubisco-dependent *E. coli*, there are other components required in developing the functional assay, i.e. a positive and negative experimental controls and optimum parameters for the selection. *rbc*L gene derived from both *Chlamydomonas* and *Synechococcus* were amplified and expressed in *E. coli*. Both *rbc*L construct are important as experimental control to reaffirm that eukaryotic rbcL are unable to be folded properly in this system, thus are insoluble and that prokaryotic rbcL is able to be folded properly into soluble enzyme by *E. coli*. Once the experimental hypothesis that *E. coli* expressing the chimeric genes derived from *Chlamydomonas* and *Synechococcus* possess different extents of susceptibility or resistance towards chloramphenicol has been established, an optimum concentration of

chloramphenicrol that can distinguish functional and non-functional rbcL can be determined. *E. coli* expressing chimeric *rbc*L-*CAT* construct will be plated on selective media comprising different concentrations of chloramphenicol. When this functional selection system is established, sequence swapping between *Chlamydomonas* and *Synechococcus rbc*L sequences can be performed, thereby forming libraries of chimeric *rbc*L-*CAT* that can be rapidly identified and selected.

1.5 Objectives of This study

As preliminary step in developing this new functional screening assay, this project focused on construction and molecular cloning of *rbc*L-*CAT*.

The objectives of this project include:

i. Amplification of *rbc*L gene derived from *Chlamydomonas* and *Synechococcus*
ii. Amplification of *CAT* gene from pYES-DEST52 (Invitrogen)
iii. Fusion of PCR-amplified *rbc*L gene with *CAT* gene via overlap extension PCR
iv. Directional cloning of *rbc*L and *rbc*L-*CAT* into entry vector, pENTR^{TM}/D-TOPO
v. Subcloning of inserted *rbc*L from pENTR-*rbc*L into expression vector, pDEST^{TM}17

CHAPTER 2

LITERATURE REVIEW

2.1 Rubisco Enzyme and Agricultural Productivity

Rubisco catalyzes the fixation of atmospheric carbon dioxide by carboxylation of RuBP, the primary step in Calvin cycle. Global biological cycle relies on Rubisco catalysis and it is estimated Rubisco catalyzed more than 10^{11} tons of atmospheric carbon annually and is a key player in the geochemistry of our biosphere (Andersson 2008). Of equal important, Rubisco presence in oceanic phytoplankton catalyzes dissolved carbon dioxide in ocean and provides almost half (>45%) of global net primary production annually (Andersson and Backlund 2008). Carbon fixation into biosphere not only sustains all life on Earth, earliest life derived carbon from CO_2 and carbon fixation accounts for the foundation of most major early divergence in the tree of life (Braakman and Smith 2012). Furthermore, the assimilated carbon dioxide is converted into 3-phosphoglycerate which is crucial in building biomass, produce energy for growth and development, hence responsible for plant productivity which directly impact on origin of food sources.

2.2 Structure of Rubisco

All Rubisco are multimeric. Each Rubisco comprises identical even number of large subunits (rbcL) and small subunits (rbcS). There are several variant forms of Rubisco with distinct composition of rbcS and rbcL. Most Rubisco studies focus on plant Rubisco which comprises of eight rbcL and eight rbcS in hexadecameric structure (known as Form-I Rubisco). Form-I Rubisco not only found in higher plants, it is also present in most chemoautotrophic bacteria, cyanobacteria, rhodophyta and phaeophytes (Andersson and Backlund 2008). This enzyme have a core of four rbcL dimers which are capped by eight rbcS at each end (Andersson and Backlund 2008). Each rbcL dimer forms two active sites, whereas rbcS is not necessary for catalysis (Spreitzer 2003). However, in the absence of rbcS, catalysis of rbcL reduces substantially and suggested that rbcS does has profound influence on catalysis (Andrews 1988). The presence of rbcS in Form-I Rubisco distinguishes it from Form-II and Form-III Rubisco protein complexes which only consist of rbcL. The simpler form II Rubisco in some cynabacteria and dinoflagellates is a dimer of rbcL subunits, whereas Form III Rubisco which comprised of pentamer of rbcL dimers is found in thermophilic archaeon (Saschenbrecker et al 2007).

2.3 Rubisco Catalysis and Limitation

Despite the pivotal role of Rubisco, it is extremely ineffective with characteristic slow catalytic activity, perhaps the slowest metabolic enzyme (Parikh et al 2006). While most enzymes able to catalyze over thousand per second, Rubisco only able to carboxylate few RuBP per second (Alberts 2002).

Rubisco carboxylation activity is mainly compromised by its non-specific properties which also able to oxygenate RuBP concomitant with release of assimilated CO_2 (as shown in Figure 2.1). Moreover, this catalysis of competing O_2 at the same active site yields phosphoglycolate which had limited use to most organisms (Andersson 2008; Raghavendra and Sage 2010). The recycling of phosphoglycolate consumes energy and 20-50% of fixed carbon is again released as CO_2 (Raghavendra and Sage 2010), which considerably reduces CO_2 fixation by up to fifty percent (Andersson and Backlund 2008).

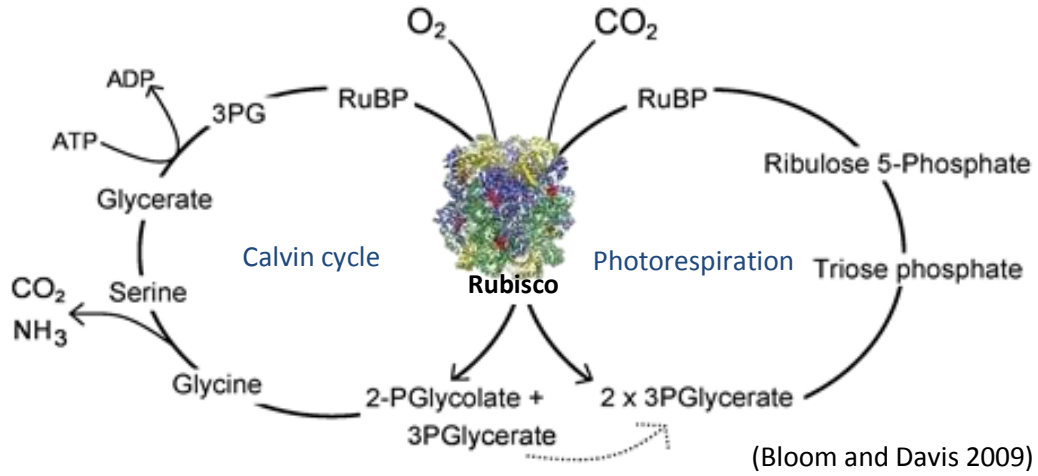


Figure 2.1 Photorespiration and Calvin cycle drives by Rubisco catalysis.

2.4 Engineering Rubisco to Improve Photosynthetic Efficiency

Increasing agricultural production had been one of the challenges to keep pace with increasing food demand and support forecast growth in population (Evans 2013). Although there is a remarkable increase of yield during green revolution through better fertilization and enhance disease resistance through genetic engineering, there is little potential for further improvement (Zhu, Long and Ort 2010). Hence, researchers had drawn new attention into metabolic engineering

approaches and the most identified target for manipulation in order to secure food demands in coming decades is Rubisco, enzyme which catalyzes cornerstone of photosynthesis (Zhu, Long and Ort 2010). Photosynthesis is the basis of plant growth and evidence had shown that increasing photosynthesis efficiency in crop plants able to raise yield potential (Evans 2013). Moreover, engineered Rubisco with better catalysis enable plant to use and convert CO_2 more efficiently and thus reduce the rate of global climatic change due to growing levels of greenhouse gases. Rising greenhouse gases particularly CO_2 have resulted in 0.76 ˚C increases in global temperature since 1800s, and it is predicted that there will be an additional increase of 1.3 to 1.8 ˚C by 2050 (Ainsworth, Rogers and Leakey 2008).

2.5 Rubisco Screening Applications

Site-directed mutagenesis and directed evolution had been frequently applied in order to generate a library of mutants with the ultimate aim of understanding the novel structure-function interactions (Parry et al 2003). Subsequently, development of screening assay is required to link changes in genotype with a measurable or selectable phenotype (Mueller-Cajar and Whitney 2008b). The common Rubisco selection system involved transformation of mutagenic libraries into host cells, followed by expression of interest gene and measuring of particular protein activity separately. *E. coli* had been the model host in expressing and assembly of engineered prokaryotic Rubisco since 1980s (Gatenby 1988). The transformed *E. coli* cells expressing engineered Rubisco were lysed and Rubisco holoenzymes were purified from cell extracts by

gradient centrifugation (Andrews 1988). Quantification of Rubisco is able to achieve by immunolabeling using anti-Rubisco polyclonal antibodies or dye-binding method (Andrews 1988; Parikh et al 2006). Subsequently, activity of Rubisco was quantified as the incorporation of acid-stable $^{14}CO_2$ from $NaH^{14}CO_3$ and radioactivity was measured in a scintillation counter (Lee, Read and Tabita 1991). The overall screening process by classical mutagenesis and enzyme assay was tedious, labor extensive, time consuming and hence slow mutant screening. Moreover, the key to the activity and efficiency of any enzyme lies hidden in the fine minutia of its three-dimensional structure and several lines of evidence had indicate slight alteration in primary structure may yield denatured protein (Goldberg 1972). Thus, when feasible, it is favorable to use functional screening method which selects enzymes based on activity by linking desired phenotype with the survival of host (Reyes-Duarte et al 2012).

2.5.1 Rubisco Functional Selection Strategies

The first reported engineered host for Rubisco bio-selection, which was created by Paoli, Vichivanives and Tabita (1998), was a Rubisco-deletion strain of *Rhodobacter capsulatus* mutant. *R. capsulatus*, a nonsulfur purple facultative bacterium, grow aerobically by utilizing hydrogen as electron source, photoheterotrophically with presence of reduced carbon substrate or chemolitotrophically on CO_2-H_2-O_2 (Hallenbeck and Kaplan 1987). Phototrophic growth of *R. capsulatus* occurs only under anoxic condition (Paoli and Tabita 1998). Complementation of mutants with functional *Synechococcus* Rubisco enables photoautotrophic growth of mutant *R. capsulatus* can be

screened and characterized with ease by plating on growth medium followed by altering culturing conditions, such as concentration of CO_2, incubation temperature etc. This *in vivo* study of Rubisco had identified several novel Rubisco mutants and demonstrated alteration of amino acids remote from active site able to affect Rubisco specificity and function (Smith and Tabita 2003). Other comparable auxotrophic photosynthetic microbial mutants had also been developed in the subsequent years and being widely used, including *Chlamydomonas reinhardtii* (Du and Spreitzer 2000) and *Synechococcus* (Greene, Whitney and Matsumura 2007). While these selection methods provide insight into development of functional screening assay for Rubisco, limitation of its throughput diminish its validity for screening large mutagenic libraries (Mueller-Cajar and Whitney 2008b). Thus, in order to enhance directed evolution schemes, a highly throughput screening method is essential.

2.5.2 Functional Screening Method for Rubisco in *Escherichia coli*

E. coli based screening assay is attractive mainly contributed by its high transformation efficiency, of up to 10^{10} transformants per microgram of plasmid (Mueller-Cajar and Whitney 2008b). The superior transformable property is at least three orders of magnitude greater compared to photosynthetic microbes and exploits the potential to screen millions of mutants each day (Parikh et al 2006). However, rate of carbon fixation in photosynthetic cells is not solely depends on Rubisco catalysis, but interaction with an array of biochemical accessories (Parikh et al 2006). Moreover, some photosynthetic

microorganisms had developed unique biological adaptation to compensate slow Rubisco catalysis, including CO2-concentrating mechanism found in algae (Moroney and Ynalvez 2007) and terrestrial plants (Parry et al 2003). This accessories and mechanisms obviate the further evolution of its intrinsic catalytic parameters (Parikh et al 2006). In contrast, heterotrophic *E. coli* does not ordinarily possess carbon fixation and hence Rubisco variants evolved within it will be under greater selective pressure. Moreover, the measured rate of carbon fixation in *E. coli* could be contributed by Rubisco characteristic solely.

The first reported *E. coli*-based strategy for selection of functional Rubisco was based on cloning of phosphoribulokinase (PRK) and *rbc*L/S genes into *E. coli* which subsequently cultured in medium supplemented with L-arabinose. Wild type *E. coli* converts L-arabinose into D-ribulose-5-phosphate and the intermediate is subsequently being converted into RuBP (Parikh et al 2006). *E. coli* intrinsically does not utilize RuBP and its accumulation diverts carbon flux from phosphogluconate shunt into metabolic dead end. However, co-expression of functional Rubisco rescue prkA-expressing *E. coli* from growth arrest and provided a relatively simple yet effective selection system to screen shuffled DNA libraries of Rubisco gene that able to form functional enzyme. The demonstrated *E. coli*-based screening method is sensitive and highly throughput, yet low validity due to frequency of false positive.

Another comparable Rubisco-dependent *E. coli* was demonstrated by Mueller-Cajar, Morell and Whitney (2007). A glycolytic blockage *E. coli* mutant was developed by removing glyceraldehyde-3-phosphate dehydrogenase (gapA), enzymes which converts D-glyceraldehyde-3-phosphate into 1,3-diphosphoglycerate and crucial in both glycolytic and gluconeogenesis pathway (Charpentier and Branlant 1994). Thus, co-expression of PRK and functional Rubisco is crucial in order to by-pass the blockage, particularly when cell are provided with minimal medium containing carbon source upstream of glycolytic lesion. Accumulation of RuBP is toxic in *E. coli* (Greene, Whitney and Matsumura 2007). *E. coli* with ΔgapA mutation is more susceptible towards toxic RuBP compared with wild type *E. coli* and had further prevents false positive (Mueller-Cajar, Morell and Whitney 2007). However, due to the stringency of selective pressure towards *E. coli* mutants, a longer incubation period is needed (4 to 10 days according to incubation parameter).

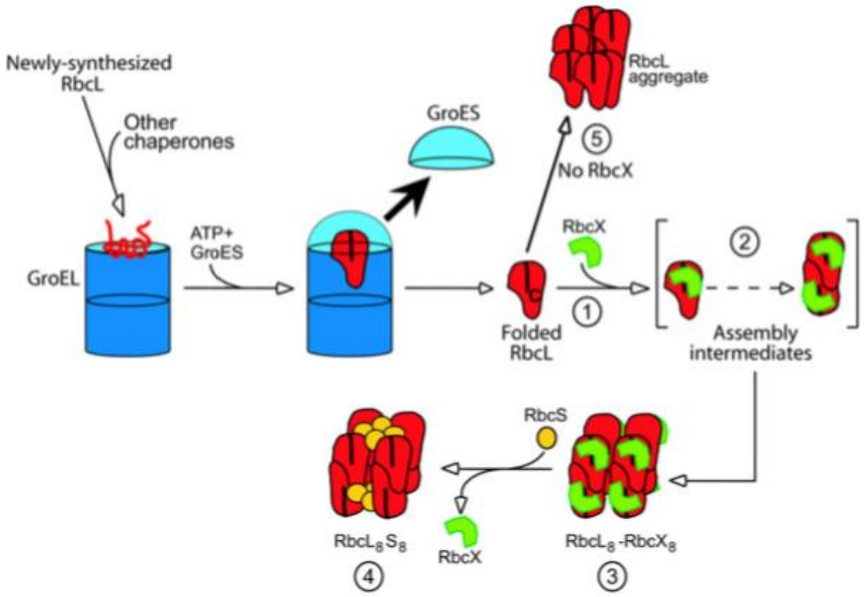
2.6 Species Specific Assembly Barrier of Rubisco

Factors influencing the folding and subsequent assembly of Rubisco into biologically active oligomeric complexes had been studied extensively with prokaryotic system, particularly *E. coli* (Tabita 1999). Expression and production of functional prokaryotic Rubisco in *E. coli*, such as *Rhodospirillum rubrum* Rubisco enables convenient assay and had aids the preliminary studies on structure and function of Rubisco enzyme (Somerville and Somerville 1984; Gutteridge et al 1984). In contrast, plant Rubisco genes, which have 80% identical in sequence to cyanobacteria (Cloney, Bekkaoui and Hemmingsen

1993), are readily expressed in *E. coli* due to high similarities between chloroplast and bacterial transcription-translation signals (Gatenby 1988), yet properly assembled functional eukaryotic Rubisco in *E. coli* has not been possible (Spreitzer 2003). In fact, despite the likelihood that all plastids belong to common lineage, the folding and assembly mechanism between higher plants and non-green algal had also diverged substantially that properly assembled algal Rubisco has not been recovered from transformed plant (Whitney et al 2001). Thus, it is suggested that the folding requirements vary between different origins and is mainly due to incompatible chaperonin system.

Molecular chaperone primarily prevents the aggregation of polypeptides chain and mediates folding followed by assembly of newly-synthesized polypeptides (Saschenbrecker 2007). Eukaryotic plastid molecular chaperone are structurally distinct from bacterial (Cloney, Bekkaoui and Hemmingsen 1993). GroEL, with its cofactor GroES from *E. coli* recognize as one of the best characterized chaperone system and plays pivotal role in assisting assembly of Rubisco (Liu et al 2010). Studies had also reported that expression of rbcX, genes juxtaposed between rbcL and rbcS in several cyanobacteria had efficiently enhances the production of functional enzymatic Rubisco in *E. coli* (Tabita 1999). Subsequent to GroEL/GroES-mediated folding, RbcX specifically recognize the C-terminal of RbcL and assists the formation of stable $RbcL_8$ followed by the displacement of RbcX with RbcS (Saschenbrecker et al 2007). However, co-expression of plastid chaperone system with plant Rubisco in *E. coli* did not aid the assembly of plant Rubisco, hence it is suggested that additional factors

had participate in the post-translational processes (Cloney, Bekkaoui and Hemmingsen 1993).



(Saschenbrecker et al 2007)

Figure 2.2 Proposed model of RbcX role in assembly of cyanobacterial Rubisco. RbcX dimer recognizes C-terminal of RbcL and protects it from undergoing aberrant interactions by bounding to RbcX central binding cleft, subsequent to GroEL/GroES-mediated folding. Role of RbcX in Rubisco assembly is not essential in some Rubisco homologs.

2.7 Recombinant Protein Consisting *rbc*L and Reporter Gene

*rbc*L promoter and translation elements had been used in driving expression of exogenous reporter gene due to its intrinsic abundance in photosynthetic microbes (Salvador, Klein and Bogorad 1993). Muto, Henry and Mayfield (2009) had demonstrated expression of *rbc*L-*lux*Ct (chloroplast lucriferase gene) in *Chlamydomonas reinhardtii* and exploit its potential to enhance accumulation of poorly-expressed protein. The study used *C. reinhartii* mutant with *rbc*L knocked out, and hence mutants depend on incorporated *rbc*L gene for photosynthetic growth. For construction of chimeric gene, *lux*Ct gene was fused and framed with downstream of *rbc*L with a GS linker and enzymatic

cleavage site. After cleavage of protease site by endogenous enzyme, rbcL with short linker peptide on carboxy terminus are found able to assemble into functional multimeric holoenzyme and support photoautotrophic growth. Western blot analysis also demonstrated predominant processing of rbcL to mature length and production of stable Rubisco. This study had also suggested *Chlamydomonas* rbcL with a short linker on carboxy terminal does not substantially affect its assembly or catalysis.

2.8 Degradation of Unstable Protein in *Escherichia coli*

Protein catabolism plays an important role in physiology of living cells. It involves the cleavage at multiple sites within target protein followed by complete degradation into amino acids, where intermediates and degradation products are generally not found (Gottesman and Maurizi 1992). This statement is supported as researchers found that while wild type β-galactosidase and lac repressor are stable in *E. coli*, both β-galactosidase fragments and denatured lac repressor produced by *E. coli* mutant are both rapidly degraded (Nath and Koch 1971). Proteolysis generally occurs in two conditions: (i) degradation of abnormal protein and denatured proteins (Varshavsky 2003) resulting from mistakes in transcription or translation, or genetic mutation (Goldberg 1972); (ii) degradation of cell proteins during starvation for new protein synthesis (Strauch and Backwith 1988). Intracellular proteolysis is also crucial in modulating the levels of metabolic enzymes.

2.9 Overlap Extension PCR

Overlap extension PCR, which also known as restriction-free cloning is one of the PCR-mediated cloning which is relatively straightforward, efficient and reliable (Bryksin and Matsumura 2010). In contrary to classical Ligation-Dependent Cloning approach, which constructs recombinant plasmid by cleaving DNA fragment with restriction enzyme followed by ligation of respective fragment into cloning vector (Unger et al 2010), overlap extension PCR is able to clone any DNA insert into desired plasmid without restriction endonucleases and DNA ligase (Bryksin and Matsumura 2013). Overlap extension PCR enables the cloning of any interest sequence, without any extra sequences, into any desired destination vector at any location (Bond and Naus 2012). Furthermore, this method is also known to be reliable, due to its simplicity in monitoring and optimization where pre-treatment of fragment and destination vector is omitted (Unger et al 2010). Moreover, its simplicity enables inexperience users to construct recombinant DNA without mastering idiosyncrasies of restriction enzymes. Two desired DNA fragments are amplified separately using overlapping primers encoding homologous sequence at the 5' ends and 3' ends, respectively. Hence, after PCR-amplification, two DNA fragments with homologous region are obtained. Both amplified fragments are mixed and use as DNA template for second round of PCR amplification. During the denaturation phase of PCR, both templates denatured into single-stranded and ligated at homologous region at annealing phase. The recombined fragment is then amplified at full length during elongation phase (Nelson and Fitch 2011).

2.10 Gateway® Recombination Cloning Technology

Construction of transgene and generation of expression clone are often hampered by laborious conventional cloning method and difficulty in generate appropriate expression clone which enables protein production in high amount (Curtis and Grossniklaus 2003). Trial and error is fundamental in construction of suitable expression clone and it usually requires repeatedly designing new experimental setup which may involves divergent multi-cloning sites to move gene from one vector to another (Esposito, Garvey and Chakiath 2009). Gateway® Technology, which patented by Invitrogen in year 2000 provides a simple way to generate constructs and enables moving gene of interest into multiple vector systems (Invitrogen Corporation 2006). Among recombination vectors available, Gateway® Technology poses the highest flexibility with amenable parallel generation of expression clones (Esposito, Garvey and Chakiath 2009). It is a relatively efficient method which utilizes 3' to 5' exonuclease activity of T4 DNA polymerase to cleave specific site and yield complementary overhang in target vector and insert that ligate together when mixed (Dortay et al 2011). The cloning vector is then transformed into *E. coli* to replicate and subjected to subsequent construct of compatible expression vector by performing LR recombination (Dortay et al 2010). Introduction of resulting expression construct into appropriate host enable expression of protein of interest in high amount (Invitrogen Corporation 2006).

2.10.1 pENTR™/D-TOPO Cloning Vector

pENTR™ TOPO Cloning enables directionally clone of blunt-end PCR product into entry vector by exploiting the reaction of topoisomerase I from *Vaccinia* virus (Invitrogen Corporation 2006). Respective topoisomerase cleaves one of the strands of pENTR™/D-TOPO vector and leaving 3' GTGG overhang followed by formation of bonding between topoisomerase and cleaved phosphodiester bond. Insert which possess complementary sequence reversed the reaction and releasing topoisomerase by binding to the cleavage site. The homologous region is created by adding 5' CACC to forward primer while amplifying insert (Invitrogen Corporation 2006). The *att*L1 and *att*L2 sites in pENTR vector enable cloning of insert into a choice of destination vector comprising *att*R sites via LR recombination (refer Appendix D). Besides, this vector also consists of kanamycin resistance gene for screening *in E. coli*.

2.10.2 pDEST™ 17 Expression Vector

pDEST™ vectors consists of *att*R1 and *att*R2 (refer Appendix E) which enables cloning of gene of interest from entry vector which consists attL into expression clone via LR recombination (Invitrogen Corporation 2012). Expression clone consists of ampicillin resistance gene and transformed *E. coli* are plated on ampicillin plate for selection. Purified plasmid from positive clone is further transform into BL21(DE3) *E. coli* for protein production. Bacteriophage T7 promoter tightly regulates the expression of insert and is inducible with IPTG (Invitrogen Corporation 2012).

CHAPTER 3

MATERIALS AND METHODS

3.1 Summary of Procedure for Molecular Cloning of *rbc*L

Preparation of media and solutions required was the preliminary step of experiment. Next, *E. coli* Top10 cell culture was revived from glycerol stock prior proceeding to preparation of electocompetent cells. Gradient PCR was then carried out in order to determine optimum parameter for amplification of both *rbc*L and *CAT*. PCR-amplified *rbc*L and *CAT* gene were fused together via overlap extension PCR. Purified PCR products were cloned into cloning vector and transformed into electrocompetent *E. coli* through electroporation. Positive clones grown on selective plate which carrying recombinant plasmid consisting inserted gene in accurate orientation was screened through colony PCR. Plasmids isolated from positive clones were further verified through amplification of respective inserted gene and followed by DNA sequencing. Verified recombinant cloning vector was further shuffled into expression vector and transformed into *E. coli*. Essential screening and verification steps were carried out in order to identify the desired recombinant plasmid.

3.2 Experimental Design for Amplification of *rbc*L and *rbc*L-*CAT*

The ultimate goal of this project is to amplify and clone *rbc*L gene derived from *Chlamydomonas* and *Synechococcus* with and without fusion of *CAT* gene. The molecular cloning of *rbc*L started off with the amplification of *rbc*L derived from *Chlamydomonas* and *Synechococcus* using pLS-H (Du and Spreitzer 2000) and pTrcSynLS-MC (Mueller-Cajar and Whitney 2008a), respectively as template. Meanwhile, *CAT* gene was amplified from pYES-DEST52 (Invitrogen). For construction of *rbc*L-*CAT*, amplification of both *rbc*L and *CAT* gene respectively were carried out by using one of the primers of each gene consists of homologous sequence. Subsequently, ligation of *rbcL* with *CAT* gene at the short homologous region and amplification of the resulting chimeric gene was performed via overlap extension PCR. Sequences of all primers used in this project were tabulated in Table 3.1. A schematic diagram showing the amplification of both *rbc*L and *rbc*L-*CAT* were also summarized in Figure 3.1 and 3.2.

Table 3.1 List of primers used in amplification of *rbc*L and *rbc*L-*CAT*.

Primer	Sequence (5' – 3')
F-RbcL-Syn	CACCATGCCCAAGACGCAATCTGC
R-RbcL-Syn	GAGCTTGTCCATCGTTTCGAATTCG
F-RbcL-Chl	CACCATGGTTCCACAAACAGAAACTAAAG
R-RbcL-Chl	TTAAAGTTTGTCAATAGTATCAAATTCG
F-RbcL-Cat	GAAACGATGGACAAGCTCGGTGGAGGGTCTGGCGG AGGTGAGAAAAAAATCACTGGATATACC
R-RbcL-Cat	GGTATATCCAGTGATTTTTTTCTCACCTCCGCCAGAC CCTCCACCGAGCTTGTCCATCGTTTC
F-RbcL-Chl-Cat	CGAATTTGATACTATTGACAAACTTGGTGGAGGGTC TGGCGGAGGTGAGAAAAAAATCACTGGATATACC
R-RbcL-Chl-Cat	GGTATATCCAGTGATTTTTTTCTCACCTCCGCCAGAC CCTCCACCAAGTTTGTCAATAGTATCAAATTCG
R-cat	TTACGCCCCGCCCTGCCA

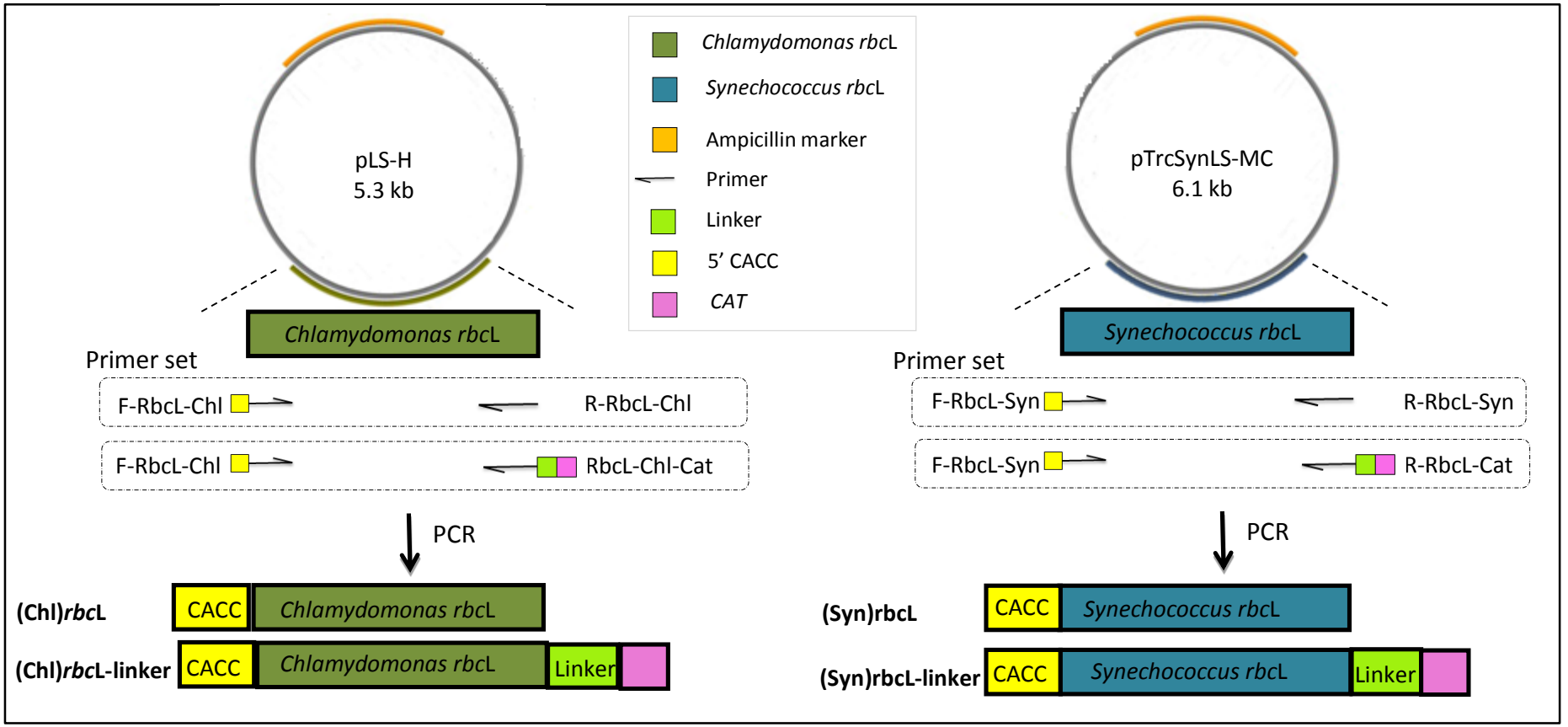


Figure 3.1 Schematic diagram showed the amplification of *rbc*L derived from both *Chlamydomonas* and *Synechococcus* using different primers. Two primer sets are used for amplification of each gene in order to yield full sequences of *rbc*L with 5'CACC and 5'CACC *rbc*L with linker and *CAT* gene homolog sequence (end product termed as *rbc*L-linker). The latter acts as one of the template for subsequent construction of *rbc*L-*CAT* chimeric gene, while amplified 5'CACC *rbc*L were subjected to TOPO cloning and LR recombination.

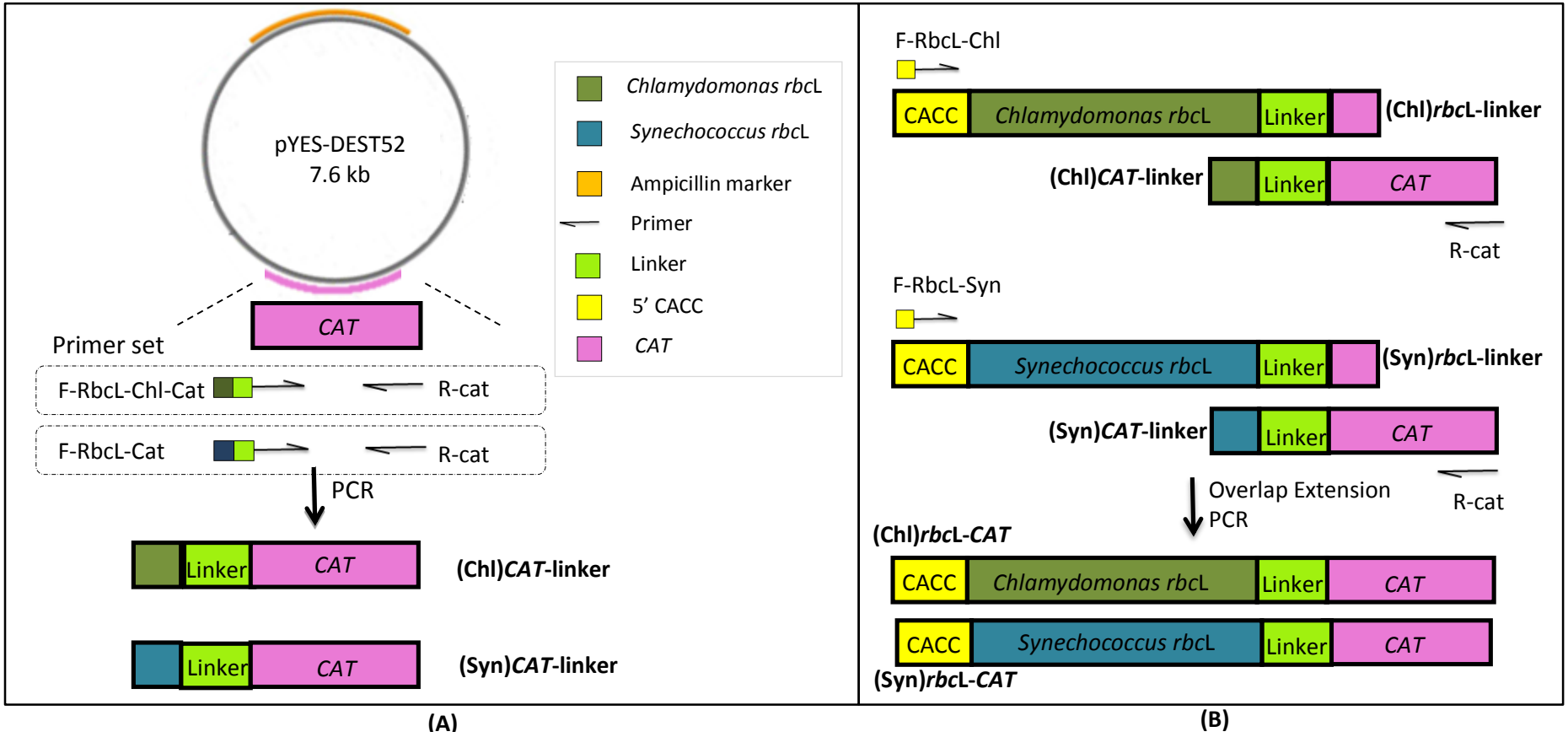


Figure 3.2 Schematic diagram showed amplification of *CAT* gene and fusion of amplified *rbc*L and *CAT* via overlap extension PCR. (A) CAT gene was amplified with two different set of primer consist of a linker and a *rbc*L segment derived from either *Chlamydomonas* or *Synechococcus*. Amplified *CAT* gene with the *Chlamydomonas rbc*L overlapping sequence is termed as (Chl)*CAT*-linker, while *CAT* gene with *Synechococcus rbc*L overlapping sequence is termed as (Syn)*CAT*-linker. (B) *CAT*-linker and *rbc*L-linker were fused at the overlapping sequences and amplified as a chimeric gene consists of both full length of *rbc*L and *CAT*(termed as *rbc*L-*CAT*) via overlap extension PCR.

3.3 Equipment and Materials Used

Table 3.2 List of equipment used and their respective manufacturers

Equipment	Manufacturer
ABJ Analytical Balance	KERN
Autoclave Sterilizer	HIRAYAMA HVE-50
Electrophoresis System	Electrophoresis system Mupid®-2plus
Electroporator 2510	Eppendorf
Imaging System	UVP
Incubator	Memmert
Laminar Flow	ESCO Laminar Flow Cabinet
Microcentrifuge	Thermo Electron Corporation
Micropipette	i. ThermoScientific ii. Eppendorf
Microwave	SHARP
Nanophotometer	Implen
PCR Thermal Cycler	Kyratec
Shaking Incubator	B-Biotek
Spectrophotometer	BIO-RAD SmartSpec™
Table Top Refrigerated Centrifuge	Dynamica
UV Transilluminator	UVP

Table 3.3 List of chemicals used and their respective manufacturers

General materials and reagents	Manufacturer
1 kb DNA ladder	i. Vivantis VC 1 kb DNA ladder ii. BioLineHyperLadder^{TM} 1 kb DNA ladder
100 bp DNA ladder	Genedirex 100 bp DNA ladder
Acetic acid	SYSTERM®
Agar powder	MERCK
Agarose powder	SeaKem® LE Agarose
Ampicillin sodium	Wako
Boric acid	QRëC^{TM}
D-Glucose	AMRESCO
DNA polymerase	i. Takara PrimeSTAR HS DNA Polymerase ii. Homemade *Taq* polymerase (Cheah 2012)
DNA polymerase buffer	i. Takara 5x PrimeSTAR Buffer-Mg^{2+} plus ii. Homemade DNA polymerase buffer (Cheah 2012)
dNTPs	Fermentas (2.5 mM)
Ethidium bromide	BIO BASIC INC.
Ethyknediaminetetraacetic acid (EDTA)	Merck Millipore OmniPur®
Glycerol	QRëC^{TM}
Hydrochloric acid	MERCK
Kanamycin sulfate	Wako
Loading dye	Fermentas 6X MassRuler DNA Loading Dye
Magnesium chloride	MERCK
Sodium chloride	MERCK
Streptomycin sulfate	United States Biological
Tris	BIO BASIC INC.
Tryptone	CONDA pronadisa
Yeast extract	Thermo Scientific Oxoid

3.4 Formulation of Solutions and Media Used

3.4.1 Culture Medium

Two types of culture medium were used in this project, Super Optimal broth with Catabolic repressor (S.O.C.) and 2x Yeast extract and Tryptone (2x YT) medium. 2x YT was used for culturing of bacteria while S.O.C. medium was only used in the final step of bacterial transformation. S.O.C. medium prolonged the logarithmic phase and enhanced the chances of harvesting bacterial culture at proper phase prior artificially induced competent state via physiochemical approach (Karcher 1995). The recipes for both media were tabulated in Table 3.4.

Table 3.4 Preparation of media.

Medium	Composition	Method of preparation
2x YT	16 g tryptone 10 g yeast extract 5 g NaCl *15 g Agar	All the ingredients were dissolved in distilled water and topped up to 1 L. pH was adjusted to 7.0 before autoclaving. *Agar was excluded for preparation of 2x YT broth
S.O.C.	20 g tryptone 5 g yeast extract 0.5 g NaCl 2.5 mL of 1 M KCl **5 mL of 2 M $MgCl_2$ **20 mL of 1 M glucose solution	All the ingredients were dissolved in distilled water and topped up to 1L. pH was adjusted to 7.0 before autoclaving. **$MgCl_2$ and glucose solution were sterilized by passing through 0.2 µm filter and added before use.

3.4.2 Antibiotic

Antibiotics, as tabulated in Table 3.5, were used in selecting transformants and propagation of bacteria cells.

Table 3.5 Antibiotics used and their respective application.

Antibiotic	Concentration (µg/mL)	Usage
Streptomycin sulfate	25	Propagation of Top10 for preparation of electrocompetent cells
Ampicillin sodium	100	Screening of transformants carrying recombinant pDEST^{TM}17
Kanamycin sulfate	50	Screening of transformants carrying recombinant pENTR^{TM}-D/TOPO

3.4.3 Solution

Two types of electrophoresis buffer, TAE and TBE buffer were used in this project. The latter was used in preparation of agarose gel for visualizing PCR products while TAE buffer is only used when agarose gel were subjected to gel purification of PCR product. The recipe of both buffers were tabulated in Table 3.6.

Table 3.6 Recipe of electrophoresis buffers

Solution	Ingredients
5x TBE buffer	54 g Tris base 27.5 g boric acid 20 mL of 0.5 M EDTA (pH 8.0)
10x TAE buffer	48.5g Tris base 11.4 mL glacial acetic acid 20 mL of 0.5 M EDTA (pH 8.0)

All the ingredients were dissolved in distilled water and topped up to 1 L.

3.5 Sterilization and Decontamination of Biological Materials

Culture media, solutions, glassware and materials such as pipette tips, toothpick, PCR tubes and microcentrifuge tube were all autoclaved at 121 ˚C, 15 psi for 15 minutes. Meanwhile, heat-sensitive solution including antibiotics and glucose solution were sterilized by passing through 0.2 µm syringe filter. On the other hand, waste materials and used glassware including petri dishes and tips were decontaminated by autoclaving at 121 ˚C, 15 psi for 30 minutes before washing or washing.

3.6 Preparation of Competent Cells

Isolated single colony of Top10 *E. coli* was picked and inoculated into 3 mL of culture medium for overnight culture with agitation of 220 rpm at 37 ˚C. Overnight culture was diluted until the OD_{600} was approximately 0.1. The diluted culture was then continue to agitate until the OD_{600} reading reaches 0.6 and rapidly being transferred to an ice water bath for 30 minutes with occasional swirling. The cells were spun down at 3,000 g, 4 ˚C for 15 minutes. Supernatant was removed and pellet was re-suspended with equal volume of ice cold 10% glycerol and spun down again with same parameter. The resulting cell pellet was re-suspended again with half and 1/25 volume of ice cold 10% glycerol. After the last centrifugation, the cell pellet was then re-suspended with adequate volume of 10% glycerol and aliquoted into ice cold cryo tubes with 50 µL per tube. The cells were frozen by dipping into liquid nitrogen and stored at -80 ˚C.

3.6.1 Cell Competency Test

Transformation efficiency of prepared competent cells was measured by transforming an aliquot with known amount of plasmid DNA. Both transformed and non-transformed aliquot (as negative control) were plated on selective plate and culture overnight in 37 ˚C incubator.

3.7 Polymerase Chain Reaction (PCR)

The overall PCR reaction including primer set and template used in each amplifications were shown in Figure 3.1. A negative control which consisting PCR reaction mixture without template was included in each amplification.

3.7.1 Amplification of *rbc*L

*rbc*L genes derived from *Chlamydomonas* and *Synechococcus* were used in this project. As shown in Figure 3.1, two sets of primer were used in amplification of each *rbc*L gene, in order to yield *rbc*L with and without *CAT* homologous sequence. The PCR parameter for amplification of *rbc*L was summarized in Table 3.7 and 3.8.

3.7.2 Amplification of *CAT* Gene

CAT gene was also amplified using two primer sets which comprised of different forward primer that consists of overlapping sequences of *Chlamydomonas* and

*Synechococcus rbc*L, respectively. PCR reaction mixture and thermocycling conditions for amplification of C*AT* gene were summarized in Table 3.7 and 3.9.

3.7.3 Overlap Extension PCR for Ligation of *rbc*L and *CAT* Gene

PCR product from amplification of *rbc*L-linker and *CAT*-linker were purified and adjusted to 1:1 molar concentration prior proceeding to fusion of genes. Overlap extension PCR comprised of 2 stages. At the initial stage, *rbc*L and *CAT* gene were fused at the complementary region and 3' overlap of each strand serves as primer for 3' extension of complimentary strand. The resulting fusion from first stage was used as the template for subsequent PCR and primers were added in order to amplify the fused gene. PCR conditions for both stages of overlap extension PCR was summarized in Table 3.10 and 3.11.

Table 3.7 PCR reaction mixture for amplification of *rbc*L and *CAT* gene.

Composition	Volume (µL)
DNA template	0.1
*2.5 mM dNTPs	0.6
*Forward primer (10 mM)	0.2
*Reverse primer (10 mM)	0.2
PrimeStar HS DNA Polymerase (2.5 units/µL)	0.1
5 x Thermo buffer	2.0
ddH_2O	10 in total

*For amplification of (Chl)*rbc*L-linker, 1.0 µL of 2.5 mM dNTPs and 0.4 µL of each primer were used.

Table 3.8 Thermocycling conditions for amplification of *rbc*L gene.

Temperature	Time	Number of cycles
96.0 ˚C	15 sec	1
96.0 ˚C	15 sec	30 cycles
*Varied	**15 sec	
72.0 ˚C	90 sec	
72.0 ˚C	90 sec	1
10.0 ˚C	∞	1

*The primer annealing temperature for amplification of (Chl)*rbc*L, (Syn)*rbc*L, (Chl)*rbc*L-linker and (Syn)*rbc*L-linker were 57.6 ˚C, 60.0 ˚C, 52.8 ˚C and 58.2 ˚C respectively.

**30 seconds of primer annealing time was used in amplification of (Chl)*rbc*L-linker.

Table 3.9 Thermocycling conditions for amplification of *CAT* gene.

Temperature	Time	Number of cycles
96.0 ˚C	15 sec	1
96.0 ˚C	15 sec	30 cycles
*Varied	5 sec	
72.0 ˚C	42 sec	
72.0 ˚C	42 sec	1
10.0 ˚C	∞	1

*Primer annealing temperature for amplification of (Chl)*CAT*-linker and (Syn)*CAT*-linker were 61.2 ˚C and 66.9 ˚C respectively.

Table 3.10 PCR mixtures for both stages of overlap extension PCR.

Stage 1

Composition	Volume (µL)
Amplified *CAT*-linker	0.1
Amplified *rbc*L-linker	0.1
2.5 mM dNTPs	0.6
PrimeSTAR HS DNA Polymerase (2.5 units/µL)	0.1
5 x Thermo buffer	2.0
ddH_2O	10 in total

Stage 2

Composition	Volume (µL)
PCR product from Stage 1	0.2
2.5 mM dNTPs	0.6
Forward primer (10 mM)	0.8
Reverse primer (10 mM)	0.8
PrimeSTAR HS DNA Polymerase (2.5 units/µL)	0.1
5 x Thermo buffer	2.0
DMSO	*Varied
ddH_2O	10 in total

*5% and 4% of DMSO were used in amplification of (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT*, respectively.

Table 3.11 Thermocycling conditions for overlap extension PCR.

Stage 1

Temperature	Time	Number of cycles
96.0 ˚C	60 sec	1
96.0 ˚C	30 sec	10 cycles
*Varied	30 sec	
72.0 ˚C	130 sec	
72.0 ˚C	180 sec	1
10.0 ˚C	∞	1

Stage 2

Temperature	Time	Number of cycles
96.0 ˚C	60 sec	1
96.0 ˚C	30 sec	30 cycles
**Varied	5 sec	
72.0 ˚C	130 sec	
72.0 ˚C	180 sec	1
10.0 ˚C	∞	1

*Annealing temperature used for fusion of both template was 63.7 ˚C for (Chl)*rbc*L-*CAT* and 65.6 ˚C for (Syn)*rbc*L-*CAT*.

**Primer annealing temperature for amplification of (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* was 58.2 ˚C and 60 ˚C, respectively.

3.8 Colony Screening

Colony PCR screens the presence of interest gene in recombinant plasmid by amplifying the particular desired gene using bacterial cells as template. First, PCR pre-mix was prepared and aliquoted into labeled PCR tubes. Single isolated colonies grown on overnight culture selective plate were picked by toothpick. The toothpick was gently touched on master plate and followed by mixing with

reaction mixture by stirring lightly. Thermocycling conditions of colony PCR resembles normal PCR, except ten minutes of initial denaturation at 96 ˚C was required for breaking of bacterial cell wall and releasing of plasmid. Primer annealing temperature of all colony PCR was set at 55 ˚C accompanied with 30 seconds of primer annealing time. The ingredients for PCR premix was tabulated in Table 3.12.

Table 3.12 PCR premix for colony PCR

Composition	Volume (µL)
ddH_2O	5.9
*Forward primer (10 mM)	1.0
*Reverse primer (10 mM)	1.0
10 x *Taq* buffer	1.0
2.5mM dNTPs	0.6
Homemade *Taq* DNA Polymerase (2 units/µL)	0.2
50mM $MgSO_4$	0.3
	Total : 10

*Primer set for colony PCR varied according to the plasmid screened. In this project, universal primer either F-M13 or R-M13 which binds to upstream and downstream of cloning site (refer Appendix D) was used in screening of recombinant pENTR^{TM}/D-TOPO. On the other hand, T7 terminator that binds to downstream of swapping region (refer Appendix E) was used for screening of recombinant pDEST^{TM}17.

3.9 Agarose Gel Electrophoresis

Agarose gel electrophoresis facilitates the visualization of PCR products. 1.0% of agarose gel was used in gel imaging and prepared by dissolving 1 g of agarose powder in 100 mL of 0.5x TBE (or TAE). All agarose gel electrophoresis was conducted with voltage of 80 V for 40 minutes. The resulting gel was stained

with ethidium bromide (in dark) for 10 minutes and visualized using UVP UV Transilluminator.

3.10 Purification of PCR Product

PCR product was purified in order to remove salt, dNTPs, primer and enzyme from reaction mixture which may affect downstream processes. Purification of PCR product was done using Thermo Scientific GENEJET PCR Purification Kit (Fermentas). It was carried out via direct purification or agarose gel purification and strictly followed the manufacturer's protocol.

3.11 Transformation

In this project, all transformation was conducted via electroporation. A volume of 2 µL (1 µL for competency test) of plasmid was mixed with 50 µL of electrocompetent cells and incubated on ice for 6 minutes. The mixture was then transferred to chilled electroporation tube and pulsed at 1.55 kV. A volume of 500 µL of pre-warmed S.O.C. broth was then quickly added into the electroporation tube and cells was re-suspended by pipetting the mixture up and down repeatedly. The mixture was transferred to microcentrifuge tube and agitated at 200 rpm, 37 ˚C for an hour and subsequently centrifuged at 300 g for 5 minutes before plated on selective agar. Centrifugation step was excluded when prepared electrocompetent cells were subjected for competency analysis.

3.12 Plasmid Extraction

All plasmid extraction was carried out using Geneaid High-Speed Mini Plasmid Kit by following the recommended protocol. The purity and concentration of isolated plasmids were measured with Nanophotometer.

3.13 Cloning of *rbc*L and *rbc*L-*CAT*

Gateway® technology (Invitrogen) enables insertion and moving of blunt-end DNA sequences into multiple vector systems depends on the requirement of particular study. There are two parts in Gateway® cloning: (1) insertion of *rbc*L or *rbc*L-*CAT* into entry vector; and (2) transferring of inserted gene from entry vector into destination vector for gene expression.

3.13.1 TOPO Cloning Reaction and LR Cloning

*rbc*L and *rbc*L-*CAT* was inserted into pENTR^{TM}-D/TOPO and/or pDEST^{TM}17 by following protocol listed in user manual of respective cloning kit. A slight modification was made in term of incubation time, where recommended 5 minutes incubation was extended to overnight incubation.

3.14 DNA Sequencing

Purified recombinant plasmid extracted from positive clones were sequenced by 1st Base Laboratories Sdn. Bhd. Recombinant pENTR^{TM}/D-TOPO carrying *rbc*L gene were sequenced with M13R and M13F, whereas recombinant

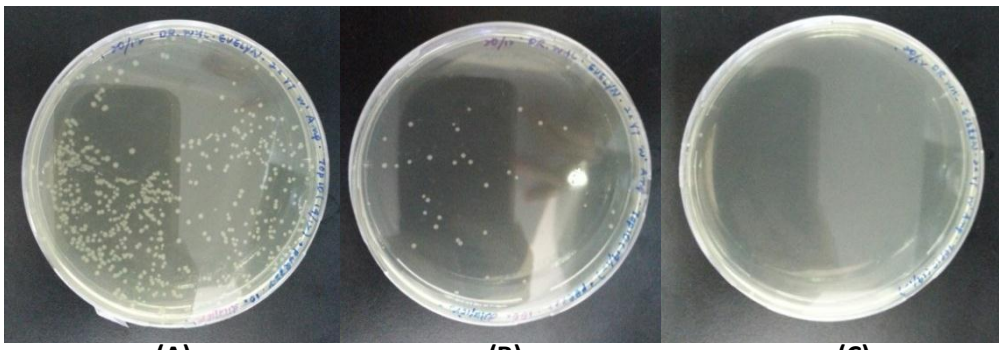
pDEST^{TM}17 were sequenced with T7 terminator and T7 promoter. On the other hand, considering the larger size of *rbc*L-*CAT*, recombinant vector carrying (Chl)*rbc*L-*CAT* was sequenced with two extra internal primers that target the insert. Primer F-RbcL-Chl-CAT and R-RbcL-Chl-CAT were used for screening of recombinant vector carrying (Chl)*rbc*L-*CAT*.

CHAPTER 4

RESULTS

4.1 Competency of Electrocompetent Cells

Prepared electrocompetent cells were transformed with pBR322 in order to facilitate the measuring of transformation efficiency prior proceeding to the downstream cloning steps. It was done by transforming 50 µL of freshly prepared electrocompetent Top10 with 1 µL of pBR322 (5 ng/µL) and topped up with 500 µL of S.O.C. medium. A serial dilution was carried out before plating of cells on selective agar containing 50 µg/mL Ampicillin (Figure 4.1).



(A) (B) (C)

Figure 4.1 Analysis of transformation efficiency of competent cells. Standard plate count (SPC) was used in enumeration of bacterial, where serial dilution was carried out in oder to obtain a countable plate. (A) A total of 490 colonies were found on plate with 10^{-1} dilution, while (B) A total of 37 colonies were observed on plate with 10^{-2} dilution. (C) No colony was observed after overnight incubation of negative control.

Transformation efficiency =

$$\frac{\text{Number of colonies}}{\text{Amount of DNA (µg)}} \text{ x } \frac{\text{Total transformation volume (µL)}}{\text{Amount plated (µL)}} \text{ x dilution factor}$$

Transformation efficiency was calculated from 10^{-2} dilution of cells (Figure 4.1; B) because its cell density is within the suitable colony counting range (25-250 CFU) recommended by U. S. Food and Drug Administration (2001). Transformation frequency was calculated as follows:

Number of colonies (on countable plate) = 37

Amount of DNA (pBR322) = 5 ng = 5 x 10^{-3} µg

Total transformation volumes (µL) = 551

Amount plated = 1 µL

Dilution factor = $\frac{551}{51}$

Hence, transformation efficiency = $\frac{37}{5 \text{ x } 10^{-3} \text{ µg}} \text{ x } \frac{551 \text{ µL}}{1 \text{ µL}} \text{ x } \frac{551}{51}$

= 4.41x 10^{7}/µg DNA

The competency of electrocompetent cell was deemed acceptable for molecular cloning.

4.2 Polymerase Chain Reaction (PCR)

4.2.1 Amplification of *rbc*L

(Chl)*rbc*L and (Syn)*rbc*L were amplified through PCR by using pLS-H and pTrcSynLS-MC as template, respectively. Figures 4.2 and 4.3 showed the PCR product at about 1.5 kb (lane 2 to 6) in amplification of *rbc*L, where five replicates of reaction mixtures were made. In Figure 4.2, there is also a faint

band lower than 200 bp in all the loaded lanes with highest intensity in the negative control lane. PCR products were then purified, the purity and concentration of purified PCR products are shown in Figure 4.4 and Table 4.1.

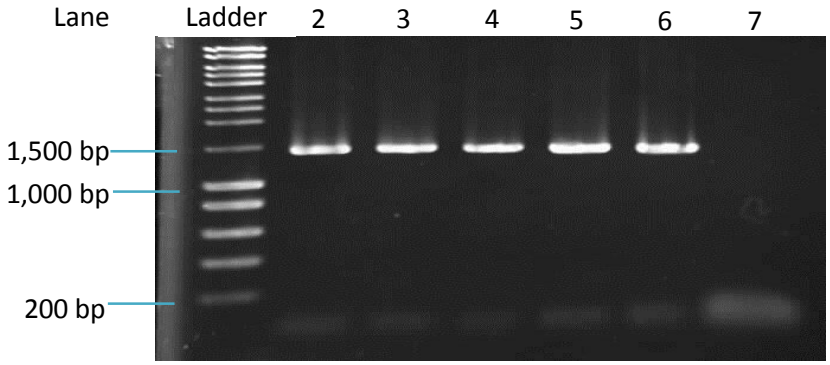


Figure 4.2 Amplification of (Chl)*rbc*L. Lane 1: BioLine HyperLadder™ 1 kb; Lane 2-6: PCR products of amplification of *Chlamydomonas rbc*L using F-RbcL-Chl and R-RbcL-Chl at annealing temperature of 57.6 ˚C; Lane 7: The negative control.

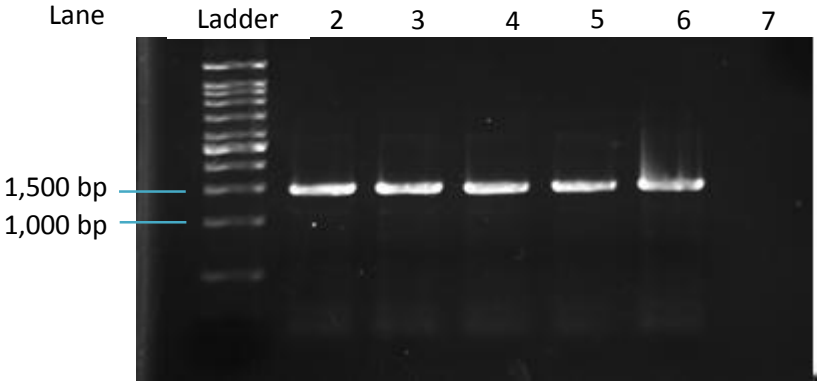


Figure 4.3 Amplification of (Syn)*rbc*L. Lane 1: VC 1 kb DNA ladder; Lane 2-6: PCR products of amplification of *Synechococcus* r*bc*L using F-RbcL-Syn and R-RbcL-Syn at annealing temperature of 60.0 ˚C; Lane 7: The negative control.

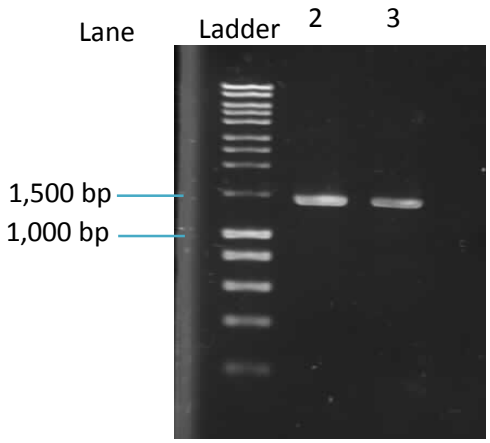


Figure 4.4 Purified (Chl)*rbc*L and (Syn)*rbc*L. Lane 1: BioLine HyperLadder™ 1 kb; Lane 2: Purified (Chl)*rbc*L; Lane 3: Purified (Syn)*rbc*L.

4.2.2 Amplification of *rbc*L-linker

(Chl)*rbc*L-linker and (Syn)*rbc*L-linker were also amplified from plasmid pLS-H and pTrcSynLS-MC, respectively, using different reverse primer from amplification of (Chl)*rbc*L and (Syn)*rbc*L. After PCR optimization, a distinct single band with desired size of about 1.5 kb was successfully achieved (Figures 4.6 and 4.7; Lanes 3, 4). Purity and concentration of resulting purified PCR products are showed in Table 4.1.

4.2.3 Amplification of *CAT*-linker

The *CAT* gene was amplified by using plasmid vector pYES-DEST52 as template with two different sets of primer which consists of homologous sequence of (Chl)*rbc*L and (Syn)*rbc*L, respectively. The gel image of resulting PCR products, (Chl)*CAT*-linker and (Syn)*CAT*-linker showed single desired band of about 700 bp (Figure 4.6 and 4.7; Lane 5, 6), while the purity and concentration of purified PCR products are tabulated in Table 4.1.

4.2.4 Fusion of *rbc*L-linker with *CAT*-linker and Amplification of the Resulting *rbc*L-*CAT* Chimeric Gene

Fusion and amplification of *rbc*L-*CAT* gene was carried out using overlap extension PCR. Template for (Chl)*rbc*L-linker is (Chl)*rbc*L-linker and (Chl)*CAT*-linker. On the other hand, (Syn)*rbc*L-*CAT* was constructed by fusing (Syn)*rbc*L-linker with (Syn)*CAT*-linker. Optimization of PCR was done via varying the annealing temperature of second PCR, as shown in Figure 4.5. Further optimization for strict PCR condition such as adding in DMSO to 5% and reducing primer annealing time to 5 sec had results in desired single band at about 2.1 kb in amplification of (Chl)*rbc*L-*CAT* (Figure 4.6, Lane 7, 8), but not (Syn)*rbc*L-*CAT* (Figure 4.7 Lane 7, 8). PCR product from amplification of (Syn)*rbc*L-*CAT* was purified using agarose gel purification, to isolate the fragment with desired size (Figure 4.8; Lane 3) Concentration and purity of both purified chimeric genes are noted in Table 4.1.

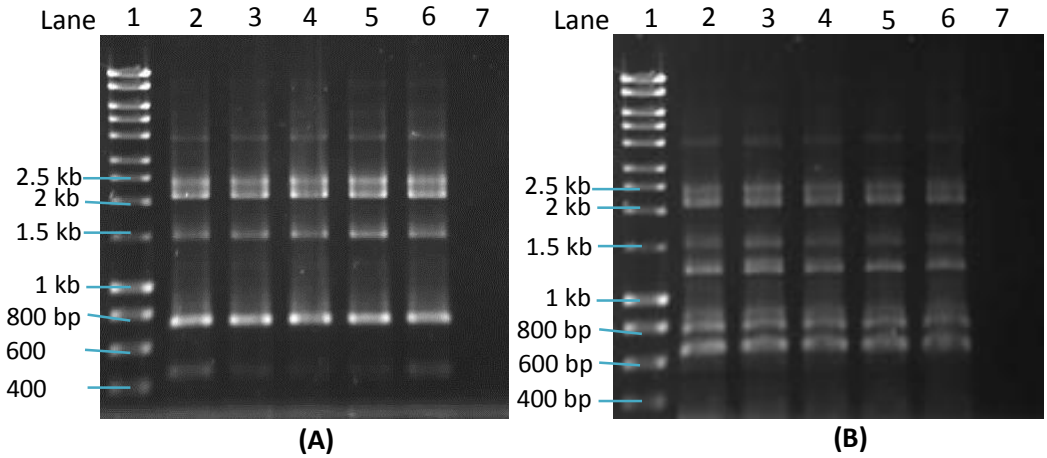


Figure 4.5 PCR products of overlap extension PCR of chimeric gene. Gradient PCR for amplification of (A) (Chl)*rbc*L-*CAT* and (B) (Syn)*rbc*L-*CAT* were carried out by setting T_a ranging from 55 ˚C to 60 ˚C using overlap extension PCR. Lane 1: BioLine HyperLadder^{TM} 1 kb DNA ladder; Lane 2: T_a = 55.0 ˚C; Lane 3: T_a = 55.8 ˚C; Lane 4: T_a = 57.9 ˚C, Lane 5: T_a = 59.1 ˚C; Lane 6: T_a = 60.0 ˚C; Lane 7: The negative control.

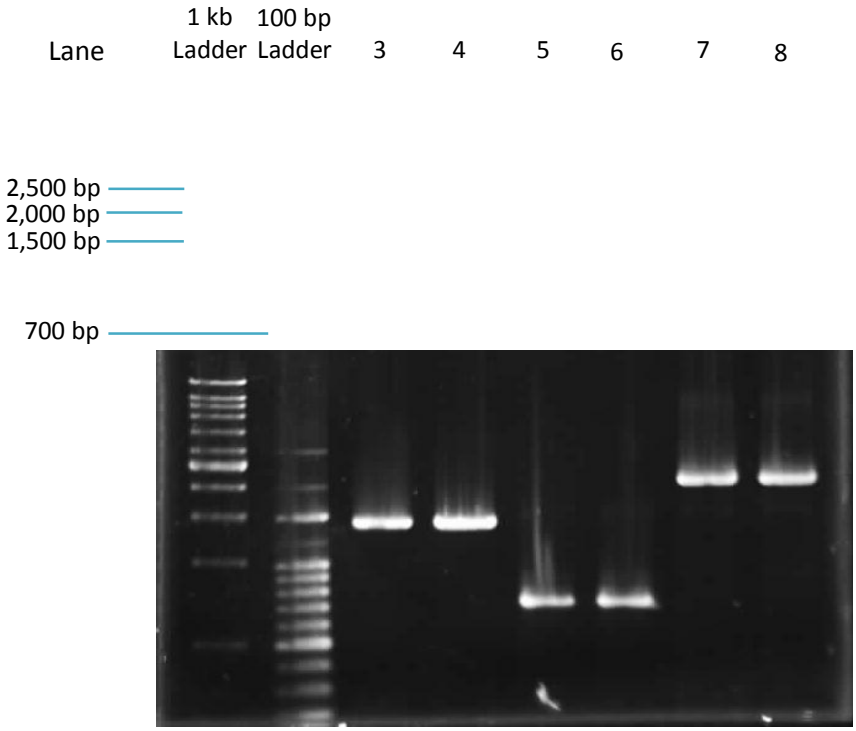


Figure 4.6 Construction of (Chl)*rbc*L-*CAT*. Lane 1: VC 1 kb DNA ladder; Lane 2: Genedirex 100 bp DNA ladder; Lane 3,4: Amplified (Chl)*rbc*L-linker; Lane 5, 6: Amplified (Chl)*CAT*-linker; Lane 7,8: Amplified (Chl)*rbc*L-*CAT*.

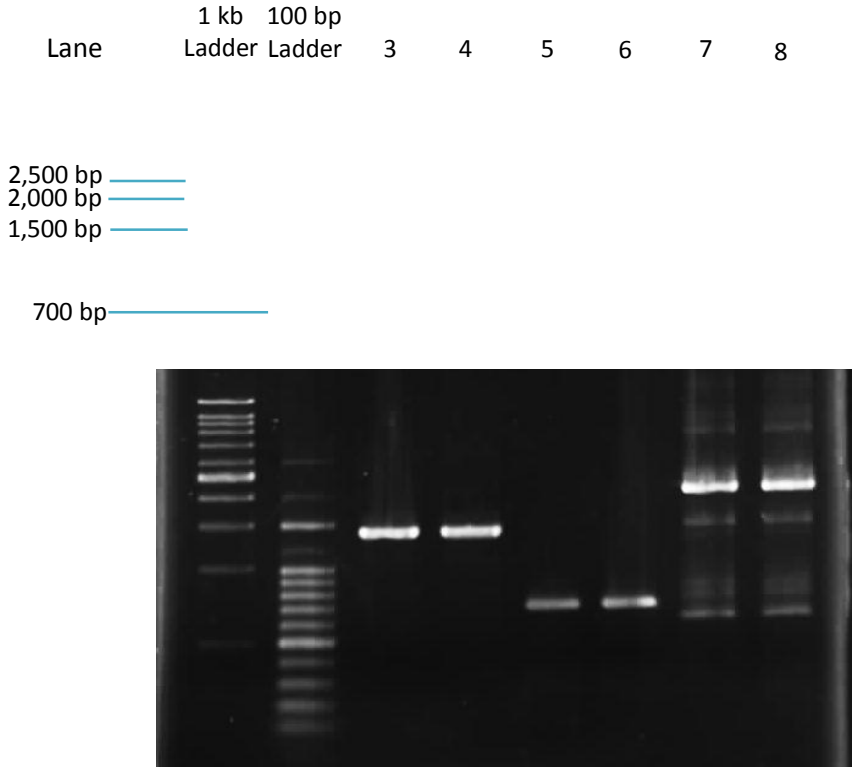


Figure 4.7 Construction of (Syn)*rbc*L-*CAT***.** Lane 1: VC 1 kb DNA ladder; Lane 2: Genedirex 100 bp DNA ladder; Lane 3,4: Amplified (Syn)*rbc*L-linker; Lane 5, 6: Amplified (Syn)*CAT*-linker; Lane 7,8: Amplified (Syn)*rbc*L-*CAT*.

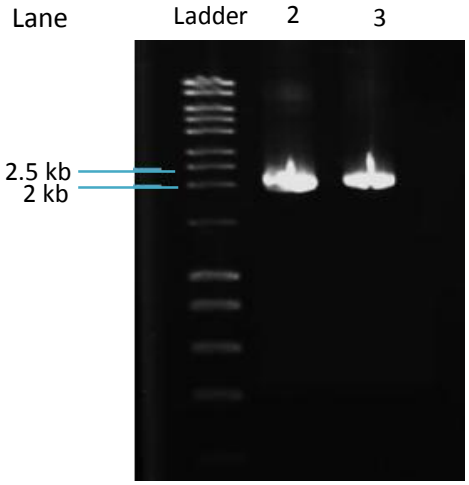


Figure 4.8 Purified *rbc*L-*CAT*. Lane 1: BioLine HyperLadder™ 1 kb DNA ladder**;** Lane 2: Purified (Chl)*rbc*L-*CAT*; Lane 3: Purified (Syn)*rbc*L-*CAT*.

Table 4.1 Concentration and purity of all purified PCR product.

Purified PCR product	Concentration (ng/µL)	A_{260}/A_{280}
(Chl)*rbc*L	174.7	1.83
(Syn)*rbc*L	90.2	1.83
(Chl)*rbc*L-linker	87.8	1.85
(Syn)*rbc*L-linker	107.6	1.86
(Chl)*CAT*-linker	112.6	1.82
(Syn)*CAT*-linker	56.3	1.80
(Chl)*rbc*L-*CAT*	170.8	1.86
(Syn)*rbc*L-*CAT*	106.3	1.86

4.3 Colony PCR of Recombinant pENTR™/D-TOPO

4.3.1 *rbc*L Gene

Both purified (Chl)*rbc*L and (Syn)*rbc*L were cloned into pENTR™/D-TOPO and transformed into electrocompetent *E. coli* Top10 through electroporation. Approximately 40 colonies were found on both selective plate and 23 colonies were selected randomly to proceed with colony PCR. In order to screen for positive transformants which carry the recombinant vector in accurate

orientation, M13F was used as forward primer. As shown in Figure 4.9, three positive clones which were designated as C8, C12 and C18 showed positive result for colony PCR that screened for cloning vector carrying (Chl)*rbc*L gene. On the other hand, one colony (designated as S20) showed positive result for screening of recombinant pENTR bearing (Syn)*rbc*L (Figure 4.10).

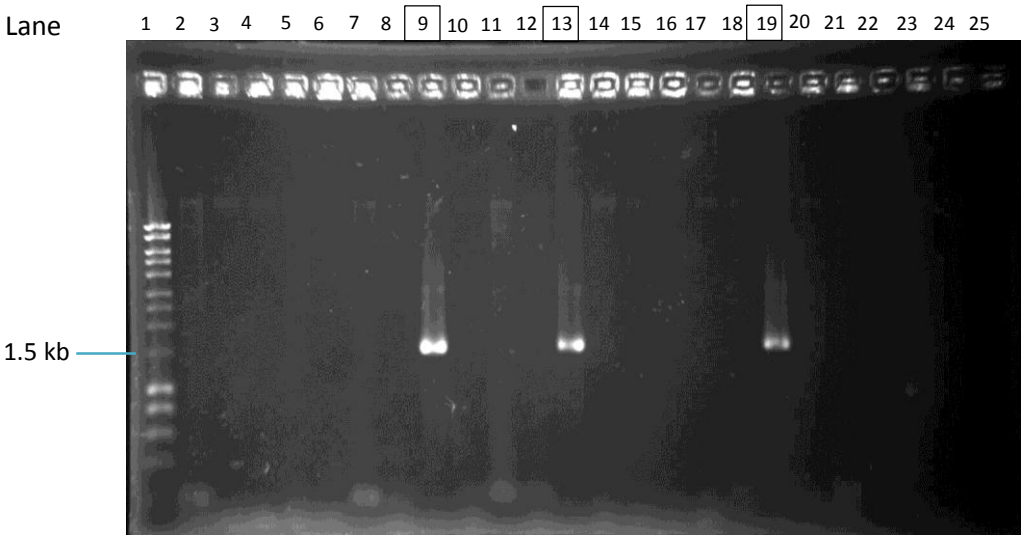


Figure 4.9 Colony PCR to screen for positive clones which consists of recombinant pENTR^{TM}/D-TOPO bearing (Chl)*rbc*L by using M13F and R-RbcL-Chl. Lane 1: BioLine HyperLadder^{TM} 1 kb DNA ladder; Lane 2-24: Products of colony PCR with randomly selected colonies C1-C23; Lane 25: The negative control. Boxed lanes are those with positive result.

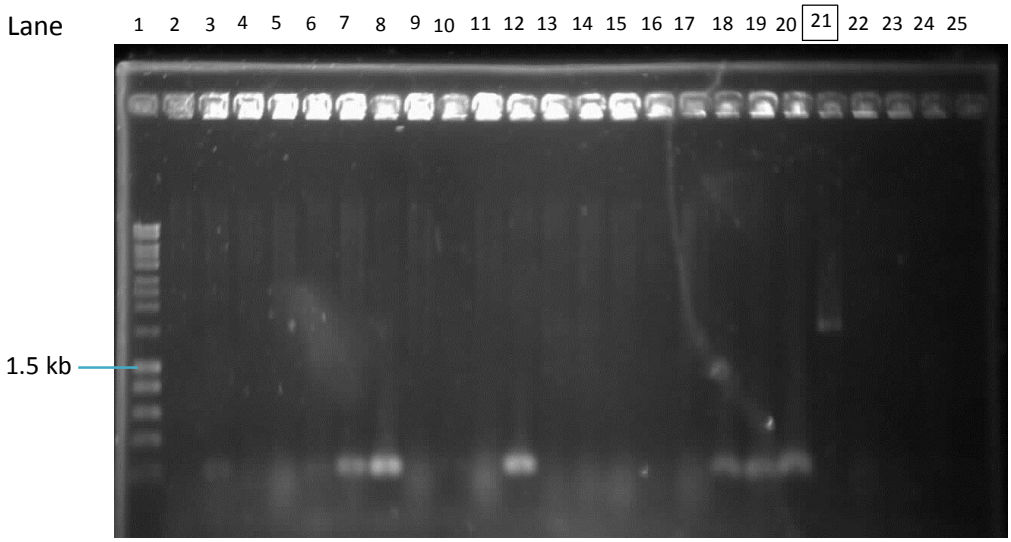


Figure 4.10 Colony PCR to screen for positive clones which consists of recombinant pENTR^{TM}/D-TOPO bearing (Syn)*rbc*L by using M13F and R-RbcL-Syn. Lane 1: BioLine HyperLadder^{TM} 1 kb DNA ladder; Lane 2-24: Products of colony PCR with randomly selected colonies S1-S23; Lane 25: The negative control. Boxed lane is that with positive result.

4.3.2 Chimeric Gene

Both purified chimeric gene were also cloned into pENTR^{TM}/D-TOPO and transformed into electrocompetent Top10 *E. coli*. Figure 4.11 showed the gel visualization of colony PCR for screening of positive transformants consists of recombinant plasmid (Chl)*rbc*L-*CAT*. A band with desired size at approximately 2.5 kb was observed on lane 3, 8 and 9 where the respective colonies were designated as CC1, CC6 and CC7. On the other hand, positive transformants which bearing (Syn)*rbc*L-*CAT* was not found.

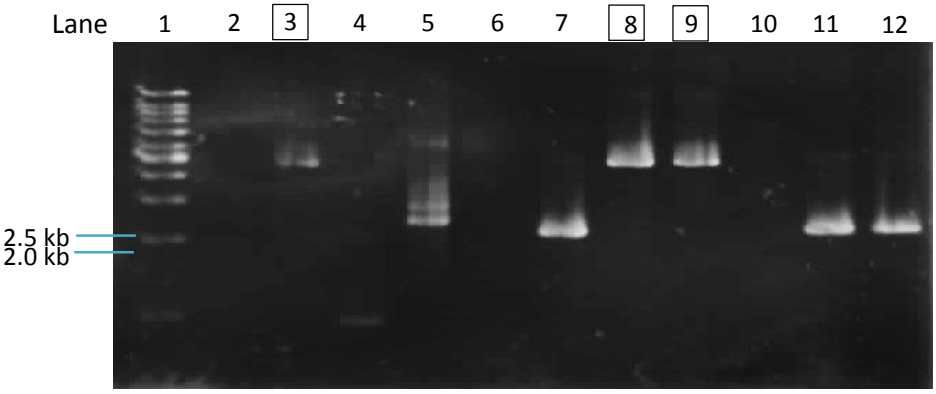


Figure 4.11 Colony PCR to screen for positive clones which consists of recombinant pENTR™/D-TOPO bearing (Chl)*rbc*L-*CAT* by using M13F and R-Cat. Lane 1: BioLine HyperLadder™ 1 kb DNA ladder; Lane 2: The negative control; Lane 3-12: Products of colony PCR with randomly selected colonies CC1-CC10.

4.4 Plasmid Extraction of Recombinant pENTR™/D-TOPO

4.4.1 pENTR-*rbc*L

Overnight culture of four colonies which showed positive result in colony PCR for screening of recombinant vector bearing (Chl)*rbc*L (C8, C12, C18) and (Syn)*rbc*L (S20) were proceeded with plasmid extraction. Purified plasmids were then used as PCR template for amplification of gene of interest. Two sets of primer were used in this verification steps: (1) M13-F paired with reverse primer of respective *rbc*L gene; (2) Forward primer of respective *rbc*L gene paired with M13-R. Gel image which visualizes the respective PCR products were showed in Figure 4.12 and Figure 4.13, band size at approximately 1.5 kb were observed with amplification of plasmid C8 and S20. Beyond expectation, extracted plasmid from colony C12 and C18 showed negative results in this verification step (Figure 4.12). The concentration and purity of extracted plasmids were tabulated in Table 4.2.

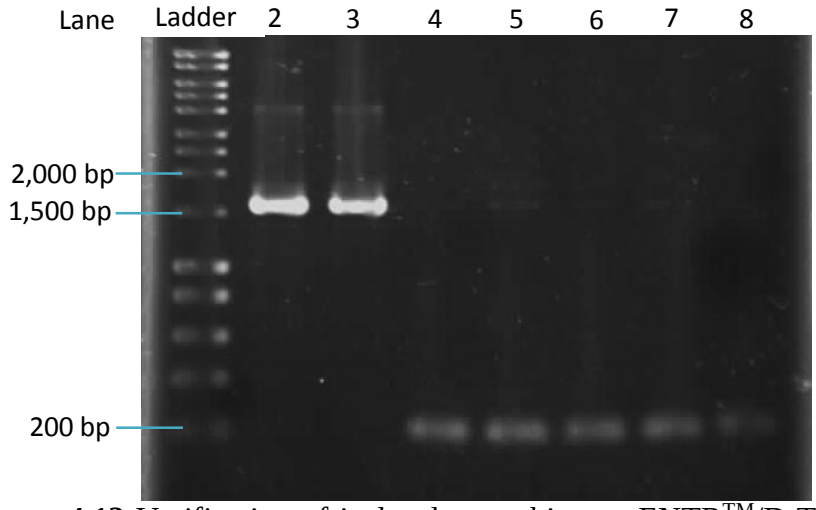


Figure 4.12 Verification of isolated recombinant pENTR™/D-TOPO bearing (Chl)*rbc*L. Lane 1: BioLine HyperLadder™ 1 kb DNA ladder; Lane 2, 4, 6: PCR product of plasmid extracted from colonies C8, C12, C18 using M13-F and R-RbcL-Chl as primer; Lane 3, 5, 7: Plasmid extracted from colonies C8, C12, C18 were used as template for PCR using F-RbcL-Chl and M13-R as primer; Lane 8: The negative control

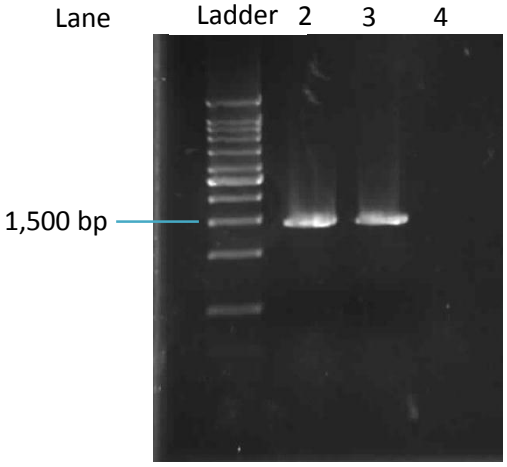


Figure 4.13 Verification of isolated recombinant pENTR™/D-TOPO bearing (Syn)*rbc*L. Lane 1: VC 1 kb DNA ladder; Lane 2: PCR product of plasmid extracted from colony S20 using M13-F and R-RbcL-Syn; Lane 3: Plasmid extracted from colony S20 were used as template for PCR using F-RbcL-Syn and M13-R as primer; Lane 4: The negative control.

4.4.2 *rbc*L-*CAT* Chimeric Gene

Plasmid of positive transformants that consists of pENTR^{TM}/D-TOPO bearing (Chl)*rbc*L-*CAT* (CC1, CC6, CC7) were extracted and used as template for amplification of insert. PCR products of extracted plasmids with band size at approximately 2.5 kb were observed in Figure 4.14. Purity and concentration of extracted plasmids were tabulated in Table 4.2.

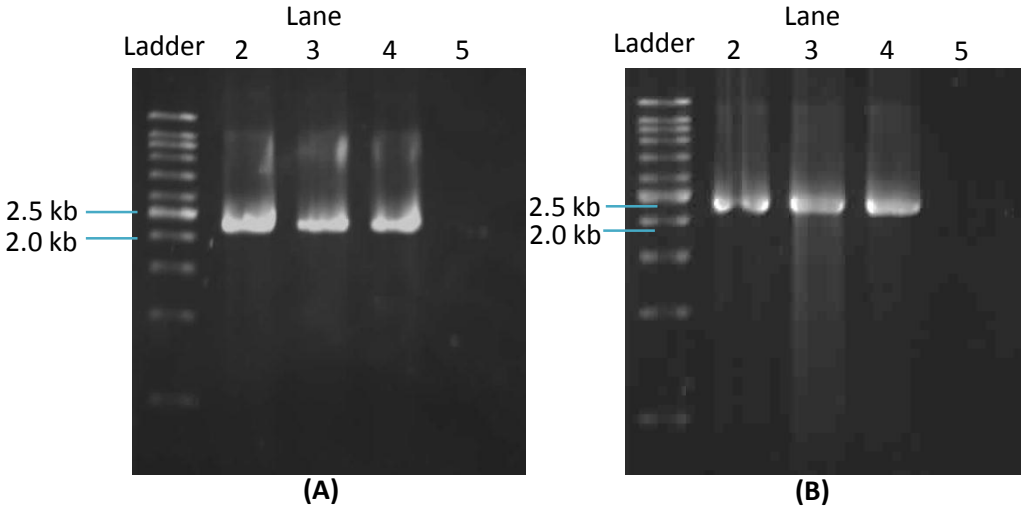


Figure 4.14 Verification of isolated recombinant pENTR^{TM}/D-TOPO bearing (Chl)*rbc*L-*CAT*. Extracted plasmid of all three positive clones were undergone PCR using (A) M13-F with R-Cat and (B) F-RbcL-Chl with M13-R. Lane 1: VC 1 kb DNA ladder; Lane 2-4: PCR product of plasmid extracted from colony CC1, CC6 and CC7, respectively; Lane 4: The negative control.

Table 4.2 Concentration and purity of all purified recombinant pENTR.

Purified pENTR^{TM}/D-TOPO Recombinant Plasmid	Concentration (ng/µL)	A_{260}/A_{280}
C8	232.1	2.06
C12	253.5	2.08
C18	261.4	2.12
S20	144.7	1.88
CC1	736.8	1.88
CC6	752.5	1.88
CC7	726.4	1.87

4.5 DNA Sequencing of pENTR^{TM}/D-TOPO Recombinant

Purified recombinant pENTR^{TM}/D-TOPO that possessed positive result in plasmid identification (C8, S20, CC1, CC6, CC7) were sent for sequencing. The sequencing results were aligned with known sequences of respective *rbc*L gene by using A plasmid Editor (ApE) software. DNA sequencing results and alignment shown clones C8 and S20 had correct sequences and the cloned genes were inserted in correct orientation. On the other hand, a mismatch at position 2031 (C·T mismatch) which causes a silent mutation was found on CC1, CC6 and CC7.

4.6 Colony PCR of pDEST^{TM}17 bearing *rbc*L

Recombinant pENTR^{TM}-D/TOPO C8 and S20 which carrying (Chl)*rbc*L and (Syn)*rbc*L, respectively were further transformed into pDEST^{TM}17 via LR recombination reaction. Four and five colonies were found on selective plates screening for pDEST-*rbc*L derived from *Chlamydomonas* and *Synechococcus*, respectively. As shown in Figure 4.15 and 4.16, three colonies showed positive result for presence of (Chl)*rbc*L, while four positive clones of (Syn)*rbc*L were also identified.

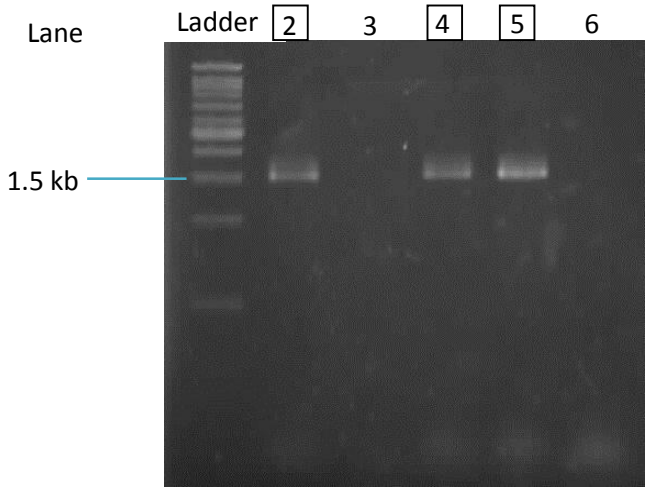


Figure 4.15 Colony PCR of cells transformed with pDEST^{TM}17 carrying (Chl)*rbc*L using T7 terminator and F-RbcL-Chl. Lane 1: VC 1 kb DNA ladder; Lane 2-5: PCR product with colonies DC1-DC4; Lane 6: The negative control. Boxed lanes are those with positive results.

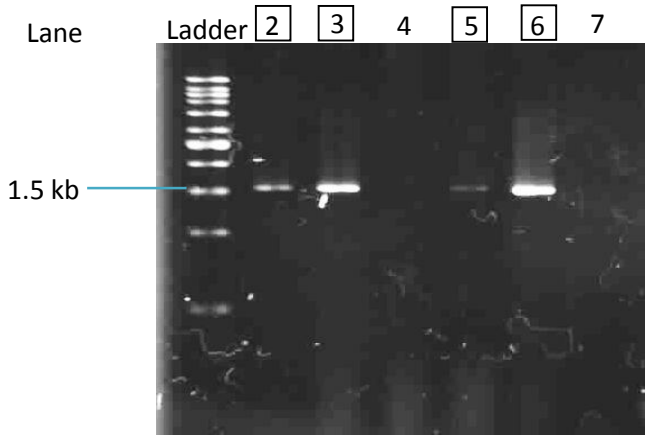


Figure 4.16 Colony PCR of cells transformed with pDEST^{TM}17 carrying (Syn)*rbc*L using T7 terminator and F-RbcL-Syn. Lane 1: VC 1 kb DNA ladder; Lane 2-6: PCR product with colonies DS1-DS5; Lane 7: The negative control. Boxed lanes are those with positive results.

4.7 Plasmid Extraction of Recombinant pDEST^{TM}17

Recombinant pDEST^{TM}17 from overnight culture of colonies DC1, DC3, DC4, DS1, DS2, DS4 and DS5 were isolated and used as PCR template for amplification of respective *rbc*L. All extracted plasmids shown positive result

for presence of respective *rbc*L in the correct orientation. Purity and concentration of purified PCR products are tabulated in Table 4.3.

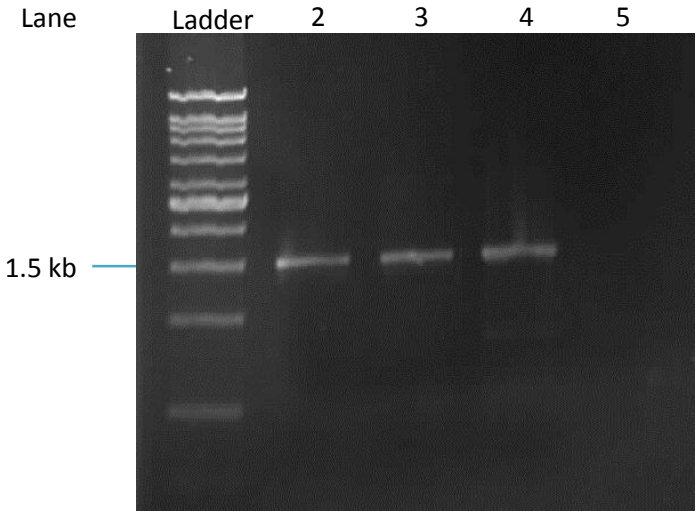


Figure 4.17 Verification of isolated recombinant pDEST^{TM}17 carrying (Chl)*rbc*L. Extracted plasmids of all positive clones were undergone PCR using T7 terminator and F-RbcL-Chl. Lane 1: VC 1 kb DNA ladder; Lane 2-4: PCR product of plasmid extracted from colony DC1, DC3, and DC4, respectively; Lane 5: The negative control.

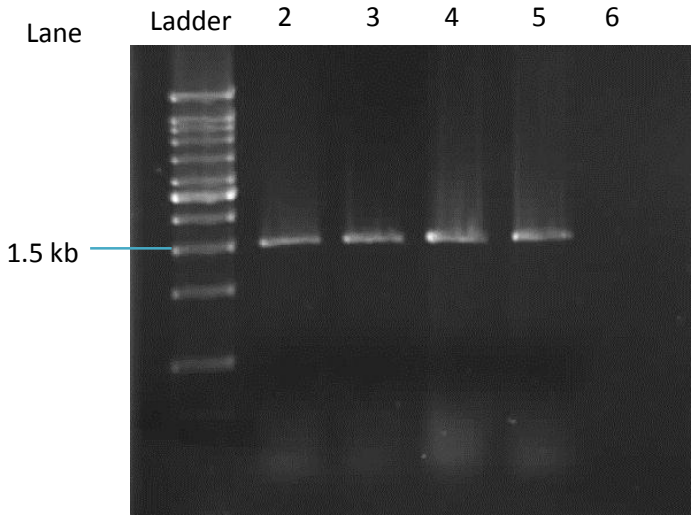


Figure 4.18 Verification of isolated recombinant pDEST^{TM}17 carrying (Syn)*rbc*L. Extracted plasmids of all positive clones were undergone PCR using T7 terminator and F-RbcL-Syn. Lane 1: VC 1 kb DNA ladder; Lane 2-5: PCR product of plasmid extracted from colony DS1, DS2, DS4 and DS5 respectively; Lane 6: The negative control.

Table 4.3 Concentration and purity of all purified recombinant pDEST^{TM}17.

Purified pENTR^{TM}/D-TOPO Recombinant Plasmid	Concentration (ng/µL)	A_{260}/A_{280}
DC1	91.9	1.83
DC3	121.4	1.89
DC4	100.4	1.85
DS1	159.4	1.96
DS2	191.4	1.91
DS4	112.2	1.91
DS5	137.0	1.89

4.8 DNA Sequencing of Recombinant pDEST^{TM}17

Purified recombinant pDEST^{TM}17 (DC1, DC3, DC4, DS1, DS2, DS4 and DS5) that possessed positive result in plasmid identification were sent for sequencing. Based on alignment of sequencing result with respective *rbc*L gene sequences, all plasmids shown correctly oriented incorporation of desired gene into pDEST^{TM}17 with 100% correct sequences.

CHAPTER 5

DISCUSSION

5.1 Transformation by electroporation

Transformation is the uptake of naked DNA and incorporation of foreign DNA into genome by an organism (Karcher 1995). Extracellular DNA are abundant in natural environment and DNA uptake by bacteria plays important roles in genetic transformation, DNA repair, and as source of nutrients during starvation (Palchevskiy and Finkel 2006). It was named by Frederick Griffth, who demonstrated first 'heritable alteration' in *Streptococcus pneumoniae* prior discovery of DNA as genetic materials (Tu 2013). It has been observed that at least 40 bacterial species under certain genera of bacteria such as *Streptococcus*, *Acinetobacter* and *Neisseria* possess natural transformation system (Chen and Dubnau 2004).

Transporting exogenous DNA across outer membrane barrier into cytosolic compartment is a complex task, and it was until early 1970 that Mandel and Higa demonstrated induced transformation in *E. coli* by altering the permeability of cell membrane through addition of calcium chloride (Nicholl 2002). Besides, the bacterial bilayer lipid membrane that acts as permeation barrier is readily modified by imposing a transmembrane electric potential, which brought up another fascinating transformation approach-electroporation (Casali and Preston

2003). It is reported that electroporation is probably the most efficient and reliable method for transformation of *E. coli*.

In this project, introduction of plasmid by transformation was done via electroporation using Top 10 *E. coli* as host cell. Eppendorf electroporation cuvettes with 1 mm gap width were used together with 1.55 kV, which corresponds to 15.5 kV/cm. The voltage applied is within the optimum field strength for *E. coli* recommended by Nickoloff (1995), which is 12.5-16.7 kV/cm. Although Zhang et al (2007) demonstrated that the optimum voltage for electroporation varied with different size of inserts, yet all the plasmids used in this project (pDEST^{TM}17, pENTR^{TM}/D-TOPO, pBR322) do not varied much in size (2.6 to 6.3 kb), thus all the transformation were carried out with constant voltage. Another influencing factor of electroporation is time constant. Eppendorf electroporator 2510 used in this project featured with exponential decay pulse with fixed time constant (5 ms), which is ideal for *E. coli* transformation (Eppendorf 2005).

E. coli is not known to be naturally transformable and hence *E. coli* in competent state are artificially induced by physiochemical method. The ability of an organism to uptake DNA is measured as competency. The calculation showed in Section 4.1 indicated that the competency of prepared electrocompetent cells was 4.41×10^7/µg DNA. Commercially prepared competent *E. coli* generally possess 10^9 competency and according to Casali and Preston (2003), transformation efficiencies of 10^6 to 10^7 are typical. The commercially

prepared high competency competent cells are needed mostly for making a complete library of complex genome (Karcher 1995), thus the prepared electrocompetent cells in this research are sufficient for cloning experiment. Furthermore, there was no sign of cells growing on negative control (Figure 4.1; C) and hence indicated that prepared electrocompetent cells were not contaminated throughout the preparing process.

5.2 Polymerase Chain Reaction

5.2.1 Amplification of *rbc*L and *CAT* gene

*rbc*L genes derived from both *Synechococcus* and *Chlamydomonas* were amplified with two sets of primers which comprised of different reverse primers in order to yield PCR products with and without *CAT* overlapping sequence at downstream of *rbc*L (Figure 3.1). A 5' CACC was placed at the upstream of *rbc*L gene by incorporating as part of forward primer. The incorporated 5' CACC sequence facilitates cloning by subsequently binds to the 3' GTGG overhang of pENTR/D-TOPO. On the other hand, *CAT* gene was amplified using two set of primers which comprise of different forward primers that encodes for overlapping sequence of *Chlamydomonas* and *Synechococcus rbc*L gene, respectively.

Optimization of PCR involves manipulating of several variables, and one of the critical factors is annealing temperature. Annealing temperature (T_a) plays crucial role upon the molecular interaction between primer and template, and

hence affecting the specificity of PCR product (McPherson and Møller 2006). In general, lower T_a shows higher tolerance towards mismatches between primer and template, and hence leading to non-specific binding and amplification of secondary products. On the other hand, high T_a ensures high specificity binding of primer and yet exceeding optimum T_a may cause low product yield due to poor annealing of primers (Rychlik, Spencer and Rhoads 1990). DNA melting temperature (T_m), temperature at which 50% of a given DNA sequence are hybridize as a double strand, while the remaining are denature and present as single strands, gives the best clue of optimum annealing temperature. Primers purchased from 1st Base were given T_m value which attached as product info. However, most of the primers used in this project consist of additional sequences, including *CAT* and *rbc*L overlapping sequence, and 5'CACC which is absence in template, indicates that primer will not anneal to the template 100%. Thus, it is not surprising that optimum T_a for amplifications were slightly lower than T_m.

As shown in Figure 4.2, 4.3, 4.6 (Lane 3-6) and 4.7 (Lane 3-6), all lanes displayed a distinct band in gel image visualizing the PCR product for amplification of *rbc*L and *CAT* gene. The size of (Chl)*rbc*L, (Syn)*rbc*L, (Chl)*rbc*L-linker, (Syn)*rbc*L-linker, (Chl)*CAT*-linker and (Syn)*CAT*-linker were 1465 bp, 1416 bp, 1510 bp, 1461 bp, 704 bp and 697 bp respectively. All the PCR products were in expected size, which were approximately 1.5 kb for *rbc*L and 700 bp for *CAT* gene. Thus, suggested that *rbc*L gene derived from *Synechococcus* and *Chlamydomonas* together with *CAT* gene had been successfully amplified from template at optimum parameter. Furthermore, it was also noted that a faint band with size smaller than 200 bp was found in

amplification of (Chl)*rbc*L, and the intensity of band was much higher in negative control (Figure 4.2). The non-specific product was known as primer dimer, product of duplex formation among two primers (Das, Mohapatra and Hsu 1999). Based on primer analysis done via OligoAnalyzer 3.1 (Integrated DNA Technologies), it is found that there is a higher extent of complementarity between primer set used in amplification of (Chl)*rbc*L, compared to other amplification reaction. Accumulation of primer-template during PCR greatly reduces formation of primer-dimer, and hence the primer-dimers were observed at higher intensity in the negative control which does not contain of template.

By comparing the PCR reaction mixture and thermocycling conditions among the six PCR-amplification reactions, the amplification of (Chl)*rbc*L-linker possesses the most sophisticated requirements. It is mainly due to non-specific binding of primer set at lower T_a that results in amplification of secondary product which has size of approximately 4 kb (result not shown). However, while increasing T_a, a lower yield of desired product was observed. Thus, in order to compensate for high T_a which gives high specificity, a less stringent PCR condition was set up by varying other PCR parameters, including increasing the volume of dNTPs to 1.0 µL, and doubling the primer annealing time and amount of primer added.

5.2.2 Overlap Extension PCR

PCR-amplified *rbc*L and *CAT* gene which both comprised of overlapping fragments were fused via overlap extension PCR. Prior to ligation, the amplified fragments were purified in order to remove PCR reaction mixture, particularly magnesium ions. According to McPherson and Møller (2006), concentration of Mg^{2+} is critical in PCR as it formed complex with dNTP which subsequently interact with sugar-phosphate backbone of nucleic acid and substantially influenced the activity of DNA polymerase. Without Mg^{2+} removal prior processing to subsequent PCR amplification may lead to excessive Mg^{2+} that reduces the fidelity of DNA polymerase and results in amplification of nonspecific products. Purification also removes interfering proteins, and the products were pure, with A_{260}/A_{280} of 1.8 to 2.0 (Table 4.1).

The purified fragments with a 1:1 molar ratio were then subjected to fusion via overlap extension PCR. Optimization of overlap extension PCR was carried out with varying primer T_a from 55 ˚C to 60 ˚C of second PCR. The expected size of fusion product (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* were 2103 bp and 2094 bp respectively. As shown in Figure 4.5, PCR products with the desired size were observed with high background of non-specific products in all the lanes and intensity of PCR product (both desired and undesired secondary product) does not seem to be affected by varying T_a.

Chakrabarti and Schutt (2001) had reported the relationship of low yield and specificity in amplification of high GC content target sequence, which resulted by increased hydrogen bond strength and formation of intermolecular secondary

structure that caused difficulty in primer annealing (Hardjasa et al 2010). Thus, it is suggested that the high background of non-specific products in amplification of *rbc*L-*CAT* was mainly due to high GC content of template used, particularly (Syn)*rbc*L-*CAT* (53.8% GC). Addition of DMSO is particularly useful in this case as DMSO with high polarity and high dielectric constant, forms hydrogen bond with major and minor grooves of template, destabilizes double helix structure and eliminate secondary structure (Weiner 2008).

Adding in dimethyl sulfoxide (DMSO) to 5% and reducing the primer annealing time to 5 sec at subsequent optimization step had resulted in the desired single band in the amplification of (Chl)*rbc*L-*CAT* (Figure 4.6; Lane 7,8), but not (Syn)*rbc*L-*CAT* (Figure 4.7; Lane 7,8). The non-specific binding in amplification of (Syn)*rbc*L-*CAT* (2.1 kb) resulted in amplification of secondary products which were approximately 650 bp and 1.6 kb. Several attempts including reducing template in both PCR respectively and further increasing T_a still unable to yield the desirable result. In general, smaller inserts are cloned more efficiently than larger inserts. Since undesired PCR products were smaller in size compared to the desired chimeric gene, gel purification was carried out in order to isolate and purify desired DNA fragments based on size. The purified PCR product from amplification of (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* were determined to be pure and consist of only desired overlap extension PCR product, based on agarose gel electrophoresis (Figure 4.8).

5.3 Colony Screening of Recombinant pENTR™/D-TOPO

All purified PCR products, including (Chl)*rbc*L, (Syn)*rbc*L, (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* were subjected to cloning into entry vector, pENTR^{TM}/D-TOPO followed by transformed into Top 10 *E. coli* for expression. Colony screening was carried out in order to screen *E. coli* that bearing recombinant plasmid with properly incorporated insert. Since primer that targeting both vector and insert were used in colony PCR, additional sequence was included in the amplification, and the expected size of PCR products is approximately 150 bp longer compared to original insert, making the expected size of PCR product were about 1.6 kb and 2.2 kb for screening of recombinant entry vector bearing *rbc*L and *rbc*L-*CAT*, respectively.

Comparing the cloning of *rbc*L and chimeric *rbc*L-*CAT*, the former was easier to achieve and suggested smaller inserts will be cloned more efficiently compared to the larger one. Positive clones carrying recombinant plasmids of (Chl)*rbc*L, (Syn)*rbc*L and (Chl)*rbc*L-*CAT* were obtained (Figure 4.9, 4.10 and 4.11), but not (Syn)*rbc*L-*CAT*. Colonies were observed on selective plate screening for transformants carrying recombinant plasmid with (Syn)*rbc*L-*CAT*, yet colony PCR of colonies obtained did not yield PCR products with the desired size. On the contrary, distinct bands with sizes smaller than 1.5 kb (result not shown) were observed in every attempts of colony PCR. Perhaps the absence of PCR product in desired size from colony screening of (Syn)*rbc*L-*CAT* may due to failure in incorporation of (Syn)*rbc*L-*CAT* gene into the entry vector, or , (Syn)*rbc*L-*CAT* was being incorporated yet unable to be amplified fully via colony PCR.

Despite cloning of *rbc*L-*CAT* derived from *Synechococcus* was failed, the PCR-amplified product of respective chimeric gene were sent for DNA sequencing to identify its identity. Sequencing using two internal primers targeting the insert unable to yield full sequence of PCR product. However, based on alignment of sequencing result with known sequence, it showed the presence of both desired genes in the PCR-amplified product, suggesting that chimeric *rbc*L-*CAT* derived from *Synechoccocus* was successfully constructed (refer Appendix F).

5.3.1 Troubleshooting Cloning of Chimeric *Synechococcus rbc*L-*CAT*

PCR-amplification of (Syn)*rbc*L-*CAT* resulted in multiple non-specific products which was subsequently removed via agarose gel purification using Thermo Scientific GeneJET PCR purification kit (Fermentas). It is suggested that the unsuccessful downstream PCR reaction may due to presence of reaction mixture residual that interferes with PCR reaction enzymes. Ethanol aids precipitation of nucleic acid during elution, yet insufficient centrifugation may results in eluate with residual ethanol. Residual ethanol not only leads to samples floating out of agarose wells when attempting to load followed by loss of purified DNA product, it also acts as PCR inhibitor (Mackay 2007). In addition, excessive salt resulting from ineffective washing is also known to affect PCR activity (Kennedy and Oswald 2011). Guanidinium isothiocyanate, one of the reagents of GeneJET DNA purification kit can also inhibit some enzymatic activities when present in trace amounts (Gerstein 2001). Apart from residual reaction mixture, it is

reported that residual traces of agarose may also inhibits subsequent manipulation (Tabak and Flavell 1978).

Taq polymerase used in colony PCR lack proofreading activity and had higher error rate which is approximately 2 x 10^{-4} per nucleotide per cycle, corresponds to ~1 substitution per 1000 bp fragment (Zylstra et al 998). It is reported that proofreading activity is substantially important in PCR-amplification of large sequence (Fromenty 2000). This is mainly because misincorporated nucleotides at 3' terminal may subsequently lead to prematurely termination of strand synthesis (Cheng et al 1994) and yielded PCR products with smaller size. Furthermore, the processivity of *Taq* polymerase reduces while amplifying long sequence and further increase the rate of mismatches (Su et al 1996), where processivity is the number of nucleotides being incorporated by polymerase during elongation before it dissociates. The inherent limitation of *Taq* polymerase is even more significant in crude sample PCR application (including colony PCR) where polymerase inhibitors, including bacterial cell debris and components of cellular media are generally presence (Costa and Weiner 2006).

According to Lonza Group (n.d.), the choice of agarose is also regarded as one of the important factors in recovering DNA. SeaKem LE Agarose used in this project possess Molecular Biology grade and are generally appropriate for analytical separation of DNA, yet it is not the best choice for recovery of DNA. SeaKem LE Agarose required high temperature for melting and potentially

causes denaturation of double-stranded DNA which substantially interfered with the downstream cloning reaction.

Besides, repeated thermocyling and prolonged high temperature elongation time in amplification of large sequence may also lead to template damage, such as DNA depurination which may also block progress of DNA polymerase. In this project, the PCR parameter of colony PCR was generalized including fixed 55 ˚C of T_a and reaction mixture regardless features of target sequence and primer set used. Thus, the PCR conditions used may not be suitable for screening of recombinant pENTR^{TM}/D-TOPO bearing (Syn)*rbc*L-*CAT*.

In order to resolve the potential problems that affect the cloning of chimeric (Syn)*rbc*L-*CAT*, several solutions were proposed. First and foremost, it is suggested that some precaution steps such as longer centrifugation time, efficient membrane wash and longer incubation time of purification column in Wash Buffer should be taken in order to remove residual from agarose gel purification, including ethanol and excessive salt that may interfere with downstream PCR reaction (Thermo Fisher Scientific Inc. 2013). In addition, contaminant such as guanidinium isothiocyanate can also be efficiently eliminated with efficient ethanol wash (Gerstein 2001).

On the other hand, Cambrex Corporation (n.d.) offered Low Melting SeaPlaque® Genetic Technology Grade (GTG) and Standard Melting SeaKem

GTG Agarose, modified agarose which characterized with lower gelling and melting point. Since conventional agarose required high melting temperature and potentially causes denaturation of DNA template followed by unsuccessful downstream cloning reaction. It is suggested that modified agarose gel with low melting point gives better alternatives for purification of DNA.

Besides, failure in detecting positive chimeric *Synechecoccus rbc*L-*CAT* clones through colony PCR may also cause by inability to amplify desired gene in colony PCR. The failure of amplifying desired gene may categorize into two main factors, non-favorable amplification conditions or limitation of reaction enzymes involved. It is suggested that more optimization steps particularly primer annealing temperature should be carried out in order to screen the desired gene under optimum condition. On the other hand, considering limitation of *Taq* polymerase in amplifying long fragments, an internal primer targeting desired gene should be used. Thus a smaller fragment is being amplified and increases the efficiency of colony PCR.

According to research carried out by Su et al (1996), they suggested that processivity of *Taq* polymerase is greatly reduced when amplifying AT rich sequences at 72 ˚C and a reduced extension temperature of 60 ˚C is advisable. Thus, it is postulated that a reduction of elongation temperature may increases the efficiency of colony PCR. Furthermore, decreasing temperature of elongation further reduces the chances of template damage. Besides, instead of using colony PCR solely to screen transformants carrying recombinant plasmid

bearing desired insert, it is postulated that additional confirmation methods can be carried out. Restriction digestion of extracted plasmids and screening of desired plasmids by evaluation of its digested size may plays as a classic yet reliable additional confirmation assay.

5.4 Verification of Recombinant pENTR^{TM}/D-TOPO

Purified plasmids of positive clones carrying recombinant pENTR^{TM}/D-TOPO with (Chl)*rbc*L, (Syn)*rbc*L and (Chl)*rbc*L-*CAT*, respectively, were used as template for another round of PCR. Figure 4.12, 4.13 and 4.14 shown the gel image visualizing the PCR products and band with expected size were found for each construct (C8, S20, CC1, CC6, CC7), thus suggesting that desired inserts were successfully being incorporated into the entry vector at appropriate orientation. On the other hand, out of three positive clones, which shown positive results in screening of transformants carrying recombinant pENTR^{TM}/D-TOPO with (Chl)*rbc*L via colony PCR, of which two shown negative results (designated as C12 and C18) with another round of PCR using extracted plasmid. Both clones C12 and C18 were re-streaked from the master plate and subjected to colony PCR again. However, negative results for both clones were observed.

Most purified plasmids shown desired A_{260}/A_{280} value, which is between 1.8 to 2.0 that indicates relatively pure nucleic acid samples (refer Table 4.2). Purified plasmids (C8, C12, C18) that shown slightly higher than expected of A_{260}/A_{280} purity ratio, suggested that it may had contaminated with RNA. It is also suggested that slight deviant from desired purity ratio would also be due to

improper handling such as measurement of blank on dirty pedestal as all samples showing higher purity ratio were measured on same particular day (Thermo Fischer Scientific Inc. 2011). However, since RNA contamination may lead to serious consequences including overestimation of plasmid concentration and interferes with sequencing reaction, it is suggested that precaution step should be implemented. Apart from measuring A_{260}/A_{280} purity ratio, agarose gel electrophoresis is a simple method that reviews the yield and purity of purified plasmid where a faint and smeary band below plasmid DNA band would be observed if RNA contamination is present.

5.5 DNA Sequencing of Recombinant pENTR^{TM}/D-TOPO

Purified plasmids from positive clones bearing recombinant pENTR^{TM}/D-TOPO with incorporated *rbc*L (C8, S20) and *rbc*L-*CAT* (CC1, CC6, CC7) were sent for DNA sequencing. Aligning sequencing result with expected sequence using ApE software shown that plasmids C8 and S20 consists of recombinant pENTR^{TM}/D-TOPO with incorporated *rbc*L derived from *Chlamydomona*s and *Synechococcus*, respectively, with 100% identities and desired orientation. According to sequencing result of CC1, CC6 and CC7, *rbc*L-*CAT* was also successfully incorporated into entry vector at appropriate orientation, yet a C·T mismatch in the *CAT* gene at position 2031 ('clean' peaks were observed from two reads which covered the particular nucleotide) was found on all three clones (refer Appendix G). However, due to the redundancy of genetic code, nucleotide mismatch at position 2031, which also third position of triplet codon causes only a silent mutation, thus it does not affect the function of the protein. The sequence

chromatogram of CC1 and CC6 had a region with low signal intensity and high background noise. Since the respective gap was covered only by single read, the identity of CC1 and CC6 was doubtful. Hence, plasmid CC7 was subjected for subsequent cloning reactions.

In addition, the sequencing results are unreadable for clones C12 and C18, which shown false positive results in colony PCR. Their results were shown in Appendix H and I. The chromatogram data of both clones were messy and multiple peaks of differing heights were observed with low number of readable peaks. Moreover, both sequencing results showed early signal drop with 'blank' sequence found. It is suggested that the 'dirty' sequence was mainly due to RNA contamination as A_{260}/A_{280} of both plasmids exceeded 2.0. From the alignment, perhaps clones C12 and C18 had truncated insert that resulted in failure of amplifying the respective insert.

5.6 LR Recombination of Recombinant pENTR-*rbc*L with pDEST™17

Incorporated *rbc*L derived from *Chlamydomonas* and *Synechococcus* in recombinant pENTR™/D-TOPO C8 and S20, respectively, were further subcloned into expression vector, pDEST™17 via LR recombination. In colony PCR screening of positive clones bearing *rbc*L, *rbc*L gene was being amplified with additional sequence (sequence between TOPO cloning site 2 and *att*L2, and sequence between pDEST™17 *att*R2 and T7 terminator complementary site), thus the expected size of PCR products were approximately 161 bp longer compared to gene size of *rbc*L solely. The expected size for colony PCR of

pDEST-*rbc*L derived from *Chlamydomonas* and *Synechococcus* were 1589 bp and 1577 bp, respectively. As shown in Figure 4.15 and 4.16, bands with desired size of approximately 1.5 kb were found on gel image visualizing PCR products of colony PCR. Positive clones comprising pDEST-*rbc*L derived from *Chlamydomonas* were designated as DC1, DC3 and DC4. On the other hand, 4 colonies shown positive result for screening of recombinant pDEST^{TM}17 bearing (Syn)*rbc*L were designated as DS1, DS2, DS4 and DS5. Plasmids isolated from the positive clones were subjected to further identification by proceeding to another round of PCR. All purified plasmids shown desired A_{260}/A_{280} value, which is between 1.8 to 2.0 that indicates relatively pure nucleic acid samples (refer Table 4.3). Based on Figure 4.17 and 4.18, PCR of extracted plasmids shown a band with the expected size of about 1.5 kb, indicating that recombination between entry clone and destination clone was obtained. All the isolated plasmid were also sent for DNA sequencing and alignment of sequencing result with sequence of respective *rbc*L shown that insert from TOPO vector was subcloned into destination vector while retaining its correct sequence and desired orientation.

CHAPTER 6

CONCLUSION

This project focused on construction and cloning of *rbc*L-*CAT* derived from *Chlamydomonas* and *Synechococcus*. *rbc*L gene derived from *Chlamydomonas* and *Synechococcus* were amplified by polymerase chain reaction (PCR) from pLS-H (Du and Spreitzer 2000) and pTrcSynLS-MC (Mueller-Cajar and Whitney 2008a), respectively. Two set of primers were used in PCR-amplification of each *rbc*L gene in order to yield both PCR product with and without *CAT* gene overlapping sequence at downstream of *rbc*L gene. On the other hand, *CAT* gene was also being amplified successfully by PCR from pYES-DEST52 (Invitrogen) using two set of primers that comprising overlapping sequence of *Chlamydomonas* and *Synechococcus rbc*L gene, respectively. Fusion of PCR-amplified *rbc*L gene with *CAT* gene at the overlapping sequences via overlap extension PCR was also successfully performed. Amplified fragments were cloned into entry vector, pENTR^{TM}-D/TOPO, then subcloned into expression vector, pDEST^{TM}17. Each cloning step was followed by transformation into Top 10 *E. coli* for expression and screening of transformants bearing recombinant vector via colony PCR. Extracted plasmids from positive clones were further identified through PCR of insert and DNA sequencing. Alignment of sequencing results with known sequences was done via A plasmid Editor (ApE) software.

Recombinant plasmids pENTR-*rbc*L derived from both *Chlamydomonas* and *Synechococcus* were successfully obtained and sub-cloned into pDEST^{TM}17. Besides that, recombinant plasmid pENTR-*rbc*L-*CAT* derived from *Chlamydomonas* was also obtained. Although chimeric *rbc*L-*CAT* derived from *Synechococcus* was constructed, cloning of *rbc*L-CAT into pENTR^{TM}/D-TOPO was unsuccessful and the actual cause was unable to be identified. However, some potential causes were discussed, including limitation of *Taq* polymerase in amplifying long fragment, residual reaction mixture and agarose that interferes with downstream reaction, inappropriate agarose gel used in gel purification of PCR product and non-favorable amplification conditions. Thus, it is proposed that efficient removal of residual, using modified agarose gel that characterized with lower melting point in gel purification of PCR product, further optimization of PCR parameter, reducing PCR elongation temperature, targeting a smaller fragment during colony PCR and restriction digestion of suspicious plasmids should be considered for cloning of (Syn)*rbc*L-*CAT*.

From the course of this project, four out of five objectives were successfully obtained. For future works, (Syn)*rbc*L-*CAT* should be re-amplified in order to obtain fresh PCR product prior proceeds to subsequent cloning reaction. After successfully cloning of (Syn)*rbc*L-*CAT* into pENTR^{TM}-D/TOPO, both pENTR-*rbc*L-*CAT* derived from *Chlamydomonas* and *Synechococcus* will be subcloned into expression vector, pDEST^{TM}17. It is postulated that *Chlamydomonas* (eukaryotic) rbcL unable to be folded properly by bacterial chaperone will possess distorted structure that subsequently disrupts conformation, thus function of CAT downstream to it. Furthermore, denatured protein with altered

conformation is also more sensitive for protease and increases its susceptibility to protease. Therefore, *E. coli* carrying the non-functional rbcL-CAT is expected to have limited resistance or susceptible to chloramphenicol. On the other hand, *Synechococcus* (prokaryotic) rbcL is able to be folded properly by bacterial chaperone and *E. coli* expressing a functional (Syn)rbcL-CAT would confer chloramphenicol resistance. By plating BL21(DE3) *E. coli* expressing (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* on culture medium consisting different concentration of chloramphenicol (5 to 34 μg/mL), the susceptibility towards chloramphenicol among transformants expressing (Chl)*rbc*L-*CAT* and (Syn)*rbc*L-*CAT* can be compared and it is postulated that a suitable concentration of chloramphenicol that able to distinguish between functional and non-functional rbcL is able to be determined. When this functional selection system is established, sequence swapping between DNA libraries of *Chlamydomonas* and *Synechococcus rbc*L genes can be performed and selection of functional rbcL can be performed based on chloramphenicol resistance.

REFERENCES

Ainsworth, E. A., Rogers, A. and Leakey, D. B., 2008. Targets for crop biotechnology in a future high-CO_2 and high O_3 world. *Plant Physiol*, vol. 147, no. 1, pp. 13-19.

Alberts, B., 2002. *Molecular biology of the cell*. [e-book] New York: Garland Science. Available through: National Center for Biotechnology Information website http://www.ncbi.nlm.nih.gov/ [Accessed 16 April 2014].

Andersson, I., 2008. Catalysis and regulation in Rubisco. *J Exp Bot*, vol. 59, no. 7, pp. 1555-68.

Andersson, I. and Backlund, A., 2008. Structure and function of Rubisco, *Plant Physiol. Biochem*, vol. 46, no. 3, pp. 275-291.

Andrews, T. J., 1988. Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits. *J Biol Chem*, vol. 263, no. 25, pp. 12213-12219.

Bloom, A. J. and Davis, U. C., 2009. As carbon dioxide rises, food quality will decline without careful nitrogen management. *Calif Agr*, vol. 63, no. 72, pp. 67-72.

Braakman, R. and Smith, E. 2012. The emergence and early evolution of biological carbon-fixation. *PLoS Comput Biol*, vol. 8, no. 4, e1002455.

Bond, S. R. and Naus, C. C., 2012. RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acid Res*, vol. 40, pp. W209-W213.

Bryksin, V. A. and Matsumura, I., 2010. Overlap extensions PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques*, vol. 48, no. 6, pp.463-465.

Bryksin, A. and Matsumura, I., 2013. Overlap extension PCR cloning. *Methods Mol Biol*, vol. 1073, pp. 31-42.

Cambrex Corporation, n.d. *A handbook for gel electrophoresis*. [Online]. Available at: http://lifeserv.bgu.ac.il/wb/zarivach/media/Documentation/Manual%20for%20Gel%20electrophoeresis.pdf [Accessed 7 April 2014].

Casali, N. and Preston, A., 2003. *E. coli plasmid vectors*. [e-book] Totowa, N. J.: Humana Press. Available at: Google Books <books.google.com> [Accessed 28 March 2014].

Chakrabarti, R. and Schutt, C. E., 2001. The enhancement of PCR amplification by low molecular-weight sulfones. *Gene*, vol. 274, no. 1-2, pp. 293-8.

Charpentier, B. and Branlant, C., 1994. The *Escherichia coli* gapA gene is transcribed by the vegetative RNA polymerase holoenzyme E sigma 70 and by the heat shock RNA polymerase E sigma 32. *J Bactriol*, vol. 176, no. 3, pp. 830-839.

Cheah, P. Y., 2012. *Expression and purification of Taq DNA polymerase from Escherichia coli and molecular cloning of rice metallothionein.* Degree Thesis, Universiti Tunku Abdul Rahman. Malaysia.

Chen, I. and Dubnau, D., 2004. DNA uptake during bacterial transformation. *Nat Rev*, vol. 2, no. 3, pp. 241-249.

Cheng, S. et al, 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Natl Acad Sci USA*, vol. 91, no. 12, pp. 5695-5699.

Cloney, L. P., Bekkaoui, D. R. and Hemmingsen, S. M., 1993. Co-expression of plastid chaperonin genes and a synthetic plant Rubisco operon in *Escherichia coli*. *Plant Mol Biol*, vol. 23, no. 6, pp. 1285-1290.

Costa, G. L. and Weiner, M. P., 2006. Colony PCR. *CSH Protoc*, vol. 1, pp. pdb.prot4141.

Curtis, M. D. and Grossniklaus, U., 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol*, vol. 133, no. 2, pp. 462-469.

Das, S., Mohapatra, S. C. and Hsu, J. T., 1999. Studies on primer-dimer formation in polymerase chain reaction (PCR). *Biotechnol Tech*, vol. 13, pp. 643-646.

Dortay, H. et al, 2011. High-throughput protein expression using a combination of ligation-independent cloning (LIC) and infrared fluorescent protein (IFP) detection. *PLoS One*, vol. 6, no. 4, pp. e18900.

Du, Y. C. and Spreitzer, R. J., 2000. Suppressor mutations in the chloroplast-encoded large subunit improve the thermal stability of wild-type ribulose-1,5-bisphosphate carboxylase/oxygenase. *J Biol Chem*, vol. 275, no. 26, pp. 19844-19847.

Ellis, R. J., 1979. The most abundant protein in the world. *Trends Biochem Sci*, vol 4, pp. 241-244.

Eppendorf, 2005. *Electroporation made easy*. [Online]. Available at: http://www.eppendorf.com/int/img/na/lit/pdf/8310-C106D-07.pdf [Accessed 16 March 2014].

Esposito, D., Garvey, L. A. and Chakiath, C. S., 2009. Gateway cloning for protein expression. *Methods Mol Biol*, vol. 498, pp. 31-54.

Esquível, M. G., Anwaruzzaman, M. and Spreitzer, R. J., 2002. Deletion of nine carboxy-terminal residues of the Rubisco small subunit decreases thermal stability but does not eliminate function. *FEBS Lett*, vol. 520, no. 1-3, pp. 73-76.

Evans, J. R., 2013. Improving photosynthesis. *Plant Physiol*, vol. 162, no. 4, pp. 1780-1793.

Fromenty, B. et al, 2000. *Escherichia coli* exonuclease III enhances long PCR amplification of damaged DNA templates. *Nucl Acids Res*, vol. 28, no. 11, pp. e50.

Gatenby, A. A., 1988. Synthesis and assembly of bacterial and higher plant Rubisco subunits in *Escherichia coli*. *Photosynth Res*, vol. 17, no. 1-2, pp. 145-157.

Gerstein, A. S., 2001. *Molecular biology problem solver*. [e-book] New York: Wiley. Available through: Wiley Online Library website http://onlinelibrary.wiley.com/ [Accessed 8 April 2014].

Goldberg, A. L., 1972. Degradation of abnormal proteins in *Escherichia coli*. *Proc Nat Acad Sci USA*, vol. 69, no. 2, pp. 422-426.

Gottesman, S. and Maurizi, M. R., 1992. Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol Rev*, vol. 56, no. 4, pp. 592-621.

Greene, D. N., Whitney, S. M., Matsumura, I., 2007. Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochem J*, vol. 44, no. Pt3, pp. 517-524.

Gutteridge, S. et al, 1984. A site specific mutation within the active site of ribulose-1,5-bisphosphate carboxylase of *Rhodospirillum rubrum*. *EMBO J*, vol. 3, no. 12, pp. 2737-2743.

Hallenbeck, P. L. and Kaplan, S., 1987. Cloning of the gene for phosphoribulokinase activity from *Rhodobacter sphaeroides* and its expression in *Escherichia coli*. *J Bacteriol*, vol. 169, no. 8, pp. 3669-3678.

Hardjasa, A. et al, 2010. Investigating the effects of DMSO on PCR fidelity using a restriction digest-based method. *JEMI*, vol. 14, pp. 161-164.

Invitrogen Corporation, 2012. *E. coli expression system with Gateway® Technology*. [Online]. Available at:

http://tools.lifetechnologies.com/content/sfs/manuals/ecoli_gateway_man.pdf. [Accessed 16 March 2014].

Invitrogen Corporation, 2006. *pRNTR™ directional TOPO® cloning kits*. [Online]. Available at: http://tools.lifetechnologies.com/content/sfs/manuals/pentr_dtopo_man.pdf [Accessed 16 March 2014].

Karcher, S. J., 1995. *Molecular biology*. [e-book]. San Diego: Academic Press. Available at: Google Books <books.google.com> [Accessed 28 March 2014].

Kennedy, S. and Oswald, N., 2011. *PCR troubleshooting and optimization*. [e-book] Norfolk, UK: Caister Academic Press. Available at: Google Books <books.google.com> [Accessed 5 April 2014].

Lee, B., Read, B. A. and Tabite, F. R., 1991. Catalytic properties of recombinant octameric, hexadecameric, and heterologous cyanobacterial/bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch Biochem Biophys*, vol. 291, no. 2, pp. 263-269.

Liu, C. et al 2010. Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature*, vol. 463, no. 7278, pp. 197-202.

Lonza Group, n.d. *Recovery of DNA from agarose gels*. [Online]. Available at: http://bio.lonza.com/uploads/tx_mwaxmarketingmaterial/Lonza_BenchGuides_SourceBook_Section_VI_-_Recovery_of_DNA_from_Agarose_Gels.pdf [Accessed 7 April 2014].

Mackay, I. M., 2007. *Real-time PCR in microbiology*. [e-book] Norfolk: Caister Academic Press. Available at: Google Books <books.google.com> [Accessed 5 April 2014].

McPherson, M. J. and Møller, S. G., 2006. PCR. [e-book] New York: Taylor & Francis. Available at: Google Books <books.google.com> [Accessed 29 March 2014].

Moroney, J. V. and Ynalvez, R. A., 2007. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*, vol. 6, no. 8, pp. 1251-1259.

Mueller-Cajar, O., Morell, M. and Whitney, S. M., 2007. Directed evolution of Rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the Form II Enzyme. *Biochemistry*, vol. 46, no. 49, pp.14067-14074.

Mueller-Cajar, O. and Whitney, S. M., 2008a. Directing of evolution of Rubisco and Rubisco activase: first impression of a new tool for photosynthesis research. *Photosynth Res.*, vol. 98, no. 1-3, pp. 667-675.

Mueller-Cajar, O. and Whiteny, S. M., 2008b. Evolving improved Synechococcus Rubisco functional expression in *Escherichia coli*. *Biochem J*, vol. 414, no. 2, pp. 205-214.

Muto, M., Henry, R. E. and Mayfield, S. P., 2009. Accumulation and processing of a recombinant protein designed as a cleavable fusion to the endogenous Rubisco LSU protein in *Chlamydomonas* chloroplast. *BMC Biotechno*l, vol. 9, no. 26,

Nath, K. and Koch, A. L., 1971. Protein degradation *in Escherichia coli*: II, strains differences in the degradation of protein and nucleic acid resulting from starvation. *J Biol Chem*, vol. 246, pp. 6956-6967.

Nelson, M. D. and Fitch, D. H. A., 2011. Overlap extension PCR: an efficient method for transgene construction. *Methods Mol Biol*, vol. 772, pp. 459-470.

Nicholl, D. S. T., 2002. *An introduction to genetic engineering*. [e-book] New York: Cambridge University Press. Available at: Google Books <books.google.com> [Accessed 27 March 2014].

Nickoloff, J. A., 1995. *Electroporation protocols for microorganisms*. [e-book]. Totowa, N. J.:Humana Press. Available at: Google Books <books.google.com> [Accessed 29 March 2014].

Palchevskiy, V. and Finkel, S. E., 2006. *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. *J Bacteriol*, vol. 188, no. 11, pp. 3902-3910.

Paoli, G. C. and Tabita, F. R., 1998. Aerobic chemolithoautotrophic growth and RubisCO function in *Rhodobacter capsulatus* and a spontaneous gain of function mutant of *Rhodobacter sphaeroides. Arch Microbiol*, vol. 170, no. 1, pp. 8-17.

Paoli, G. C., Vichivanives, P. and Tabita, F. R., 1998. Physiological control and regulation of the *Rhodobacter capsulatus* cbb operons. *J Bacteriol*, vol. 180, no. 16, pp. 4258-4269.

Parikh, M. R. et al, 2006. Directed evolution of RuBisCO hypermorphs through genetic selection in engineered *E. coli*. *Protein Eng Des Sel*, vol, 19, no. 3, pp. 113-9.

Parry et al, 2003. Manipulation of Rubisco: the amount, activity, function and regulation. *J Exp Bot*, vol. 54, no. 386, pp. 1321-33.

Raghavendra, A. S. and Sage, R. F., 2010. *C4 photosynthesis and related CO2 concentrating mechanisms*. [e-book] Springer. Available at: Google Books <books.google.com> [Accessed 6 March 2014].

Reyes-Duarte, D. et al, 2012. Functional-based screening methods for lipases, esterses, and phospholipases in metagenomic libraries. *Methods Mol Biol*, 861, pp. 101-113.

Rychlik, W., Spencer, W. J. and Rhoads, R. E., 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Res*, vol. 18, no. 21, pp. 6409-6412.

Salvador, M. L., Klein, U. and Bogorad, L., 1993. 5' sequences are important positive and negative determinants of the longevity of *Chlamydomonas* chloroplast gene transcripts. *PNAS*, vol. 90, no. 4, pp. 1556-1560.

Saschenbrecker, S. et al, 2007. Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. *Cell*, vol. 129, no. 6, pp, 1189-20.

Smith, S. A. and Tabita, F. R., 2003. Positive and negative selection of mutant forms of prokaryotic (cyanobacterial) ribulose-1,5-bisphosphate carboxylase/oxygenase. *J Mol Biol*, vol. 331, pp. 557-569.

Somerville, C. R. and Somerville, S. C., 1984. Cloning and expression of the *Rhodospirillum rubrum* ribulose bisphosphate carboxylase gene in *E. coli*. *Mol Gen Genet*, vol. 193, no. 2, pp. 214-219.

Spreitzer, R. J., 2003. Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch Biochem Biophys*, vol. 414, no. 2, pp. 141-9.

Strauch, K. L. and Backwith, J., 1988. An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. *Proc Natl Acad Sci USA*, vol. 85, pp.1576-1580.

Su, X. et al 1996. Reduced extension temperature required for PCR amplification of extremely A+T rich DNA. *Nucl Acids Res*, vol. 24, no. 8, pp. 1574-1575.

Sugawara, H. et al, 1999. Crystal structure of carboxylase/oxygenase from a thermophilic red alga, *Galdieria partita*. *J Biol Chem*, vol. 274, no. 22, pp. 15655-15661.

Tabak, H. F. and Flavell, R. A., 1978. A method for recovery of DNA from agarose gels. *Nucleic Acid Res*, vol. 5, no. 7, pp. 2321-2332.

Tabita, F. R., 1999. Microbial ribulose-1,5-bisphosphate carboxylase/oxygenase: A different perspective, *Photosynth. Res*, vol. 60, pp. 1-28.

Thermo Fischer Scientific Inc., 2013. Thermo Scientific GeneJET PCR purification kit #K0701, #K0702. [Online]. Available at:

http://www.thermoscientificbio.com/uploadedFiles/Resources/k070-product-information.pdf [Accessed 16 March 2014].

Thermo Fischer Scientific Inc., 2011. *Assessment of nucleic acid purity*. [Online]. Available at: http://www.nanodrop.com/Library/T042-NanoDrop-Spectrophotometers-Nucleic-Acid-Purity-Ratios.pdf [Accessed 6 April 2014].

Tu, A. H. T., 2013. *Transformation of Escherichia coli made competent by calcium chloride protocol*. [Online]. Available at: http://www.microbelibrary.org/component/resource/laboratory-test/3152-transformation-of-escherichia-coli-made-competent-by-calcium-chloride-protocol [Accessed 28 March 2014].

U. S. Food and Drug Administration, 2001. *BAM: aerobic plate count*. [Online]. Available at: http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm063346.htm [Accessed 16 April 2014].

Unger, T et al, 2010. Applications of the restriction free (RF) cloning procedure for molecular manipulations and protein expression. *J Struct Biol*, vol. 172, no. 1, pp. 34-44.

Varshavsky, A., 2003. The N-end rule pathway of protein degradation. *Genes Cells*, vol. 2, no. 1, pp. 13-28.

Weiner, L. P., 2008. *Neural stem cells*. [e-book] Totowa, NJ: Humana Press. Available at: Google Books <books.google.com> [Accessed 30 March 2014]

Whiteny, S. M. et al, 2001. Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J*, vol. 26, no. 5, pp. 535-547.

Zhang, Y. et al, 2007. The study of optimal conditions of electroporation in *Escherichia coli* strain. *Sheng Wu Gong Cheng Xue Bao*, vol. 23, no. 2, pp.347-51.

Zhu, X. G., Long, S. P. and Ort, D. R., 2010. Improving photosynthetic efficiency for greater yield, *Annu Rev Plant Biol*, vol. 61, pp. 235-61.

Zylstra, P. et al, 1998. PCR amplification of murine immunoglobulin germline V genes: strategies for minimization of recombination artefacts. *Immunol Cell Biol*, vol. 76, no. 5, pp. 395-405.

APPENDIX A

Gene Sequence of *rbc*L Derived from *Chlamydomonas*

ATGGTTCCACAAACAGAAACTAAAGCAGGTGCTGGATTCAAAGCCG
GTGTAAAAGACTACCGTTTAACATACTACACACCTGATTACGTAGT
AAGAGATACTGATATTTTAGCTGCATTCCGTATGACTCCACAACCA
GGTGTTCCACCTGAAGAATGTGGTGCTGCTGTAGCTGCTGAATCTTC
AACAGGTACATGGACTACAGTATGGACTGACGGTTTAACAAGTCTT
GACCGTTACAAAGGTCGTTGTTACGATATCGAACCAGTTCCGGGTG
AAGACAACCAATACATTGCTTACGTAGCTTACCCAATCGACTTATTC
GAAGAAGGTTCAGTAACTAACATGTTCACTTCTATTGTAGGTAACGT
ATTCGGTTTCAAAGCTTTACGTGCTCTACGTCTTGAAGACCTTCGTA
TTCCACCTGCTTACGTTAAAACATTCGTAGGTCCTCCACACGGTATT
CAGGTAGAACGTGACAAATTAAACAAATATGGTCGTGGTCTTTTAG
GTTGTACAATCAAACCTAAATTAGGTCTTTCAGCTAAAAACTACGGT
CGTGCAGTTTATGAATGTTTACGTGGTGGTCTTGACTTTACTAAAGA
CGACGAAAACGTAAACTCACAACCATTCATGCGTTGGCGTGACCGT
TTCCTTTTCGTTGCTGAAGCTATTTACAAAGCTCAAGCAGAAACAGG
TGAAGTTAAAGGTCACTACTTAAACGCTACTGCTGGTACTTGTGAA
GAAATGATGAAACGTGCAGTATGTGCTAAAGAATTAGGTGTACCTA
TTATTATGCACGACTACTTAACAGGTGGTTTCACAGCTAACACTTCA
TTAGCTATCTACTGTCGTGACAACGGTCTTCTTCTACACATCCACCG
TGCTATGCACGCGGTTATTGACCGTCAACGTAACCACGGTATTCACT
TCCGTGTTCTTGCTAAAGCTCTTCGTATGTCTGGTGGTGACCACCTT
CACTCTGGTACTGTTGTAGGTAAACTAGAAGGTGAACGTGAAGTTA
CTCTAGGTTTCGTAGACTTAATGCGTGATGACTACGTTGAAAAAGA
CCGTAGCCGTGGTATTTACTTCACTCAAGACTGGTGTTCAATGCCAG
GTGTTATGCCAGTTGCTTCAGGTGGTATTCACGTATGGCACATGCCA
GCTTTAGTTGAAATCTTCGGTGATGACGCATGTCTTCAGTTCGGTGG
TGGTACTCTAGGTCACCCTTGGGGTAACGCTCCAGGTGCTGCAGCT
AACCGTGTAGCTCTTGAAGCTTGTACTCAAGCTCGTAACGAAGGTC
GTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAAA
TGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTA
AATTCGAATTTGATACTATTGACAAACTTTAA

APPENDIX B

Gene Sequence of *rbc*L Derived from *Synechococcus*

ATGCCCAAGACGCAATCTGCCGCAGGCTATAAGGCCGGGGTGAAG
GACTACAAACTCACCTATTACACCCCCGATTACACCCCCAAAGACA
CTGACCTGCTGGCGGCTTTCCGCTTCAGCCCTCAGCCGGGTGTCCCT
GCTGACGAAGCTGGTGCGGCGATCGCGGCTGAATCTTCGACCGGTA
CCTGGACCACCGTGTGGACCGACTTGCTGACCGACATGGATCGGTA
CAAAGGCAAGTGCTACCACATCGAGCCGGTGCAAGGCGAAGAGAA
CTCCTACTTTGCGTTCATCGCTTACCCGCTCGACCTGTTTGAAGAAG
GGTCGGTCACCAACATCCTGACCTCGATCGTCGGTAACGTGTTTGGC
TTCAAAGCTATCCGTTCGCTGCGTCTGGAAGACATCCGCTTCCCCGT
CGCCTTGGTCAAAACCTTCCAAGGTCCTCCCCACGGTATCCAAGTCG
AGCGCGACCTGCTGAACAAGTACGGCCGTCCGATGCTGGGTTGCAC
GATCAAACCAAAACTCGGTCTGTCGGCGAAAAACTACGGTCGTGCC
GTCTACGAATGTCTGCGCGGCGGTCTGGACTTCACCAAAGACGACG
AAAACATCAACTCGCAGCCGTTCCAACGCTGGCGCGATCGCTTCCT
GTTTGTGGCTGATGCAATCCACAAATCGCAAGCAGAAACCGGTGAA
ATCAAAGGTCACTACCTGAACGTGACCGCGCCGACCTGCGAAGAAA
TGATGAAACGGGCTGAGTTCGCTAAAGAACTCGGCATGCCGATCAT
CATGCATGACTTCTTGACGGCTGGTTTCACCGCCAACACCACCTTGG
CAAAATGGTGCCGCGACAACGGCGTCCTGCTGCACATCCACCGTGC
AATGCACGCGGTGATCGACCGTCAGCGTAACCACGGGATTCACTTC
CGTGTCTTGGCCAAGTGTTTGCGTCTGTCCGGTGGTGACCACCTCCA
CTCCGGCACCGTCGTCGGCAAACTGGAAGGCGACAAAGCTTCGACC
TTGGGCTTTGTTGACTTGATGCGCGAAGACCACATCGAAGCTGACC
GCAGCCGTGGGGTCTTCTTCACCCAAGATTGGGCGTCGATGCCGGG
CGTGCTGCCGGTTGCTTCCGGTGGTATCCACGTGTGGCACATGCCCG
CACTGGTGGAAATCTTCGGTGATGACTCCGTTCTCCAGTTCGGTGGC
GGCACCTTGGGTCACCCCTGGGGTAATGCTCCTGGTGCAACCGCGA
ACCGTGTTGCCTTGGAAGCTTGCGTCCAAGCTCGGAACGAAGGTCG
CGACCTCTACCGTGAAGGCGGCGACATCCTTCGTGAAGCTGGCAAG
TGGTCGCCTGAACTGGCTGCTGCCCTCGACCTCTGGAAAGAGATCA
AGTTCGAATTCGAAACGATGGACAAGCTC

APPENDIX C

Nucleotide Sequences of Linkers and *CAT* Gene

Nucleotide sequence of linker for (Chl)*rbc*L-*CAT*

GGTGGAGGGTCTGGCGGAGGT

Nucleotide sequence of linker for (Syn)*rbc*L-*CAT*

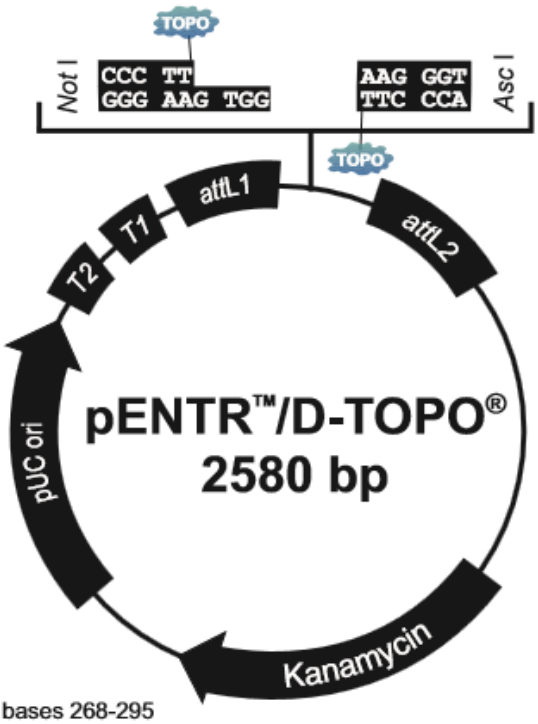
GGTGGATCTGGCGGAGGT

Nucleotide sequence of *CAT* gene

GAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCG
TAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCA
GACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATA
AGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATG
CTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGG
GATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTT
TCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATA
TATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAA
GGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTC
ACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTT
TTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCT
GGCGATTCAGGTTCATCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAAT
GCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAA

APPENDIX D

Map of pENTR™/D-TOPO



Comments for pENTR™/D-TOPO®
2580 nucleotides

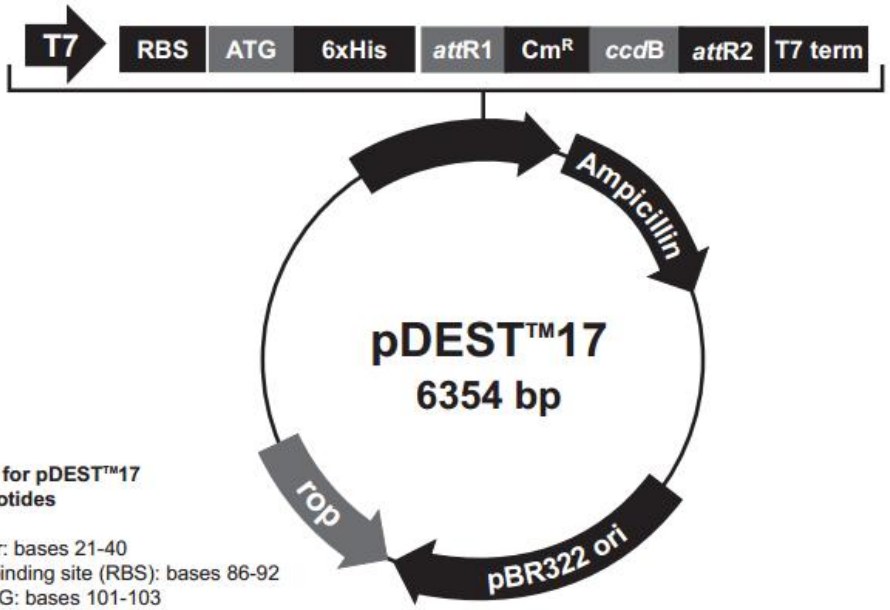
*rrn*B T2 transcription termination sequence: bases 268-295
*rrn*B T1 transcription termination sequence: bases 427-470
M13 forward (-20) priming site: bases 537-552
*attL*1: bases 569-668 (c)
TOPO® recognition site 1: bases 680-684
Overhang: bases 685-688
TOPO® recognition site 2: bases 689-693
*attL*2: bases 705-804
T7 Promoter/priming site: bases 821-840 (c)
M13 reverse priming site: bases 845-861
Kanamycin resistance gene: bases 974-1783
pUC origin: bases 1904-2577

(c) = complementary sequence

(Invitrogen Corporation 2006)

APPENDIX E

Map of pDEST-17



Comments for pDEST™17
6354 nucleotides

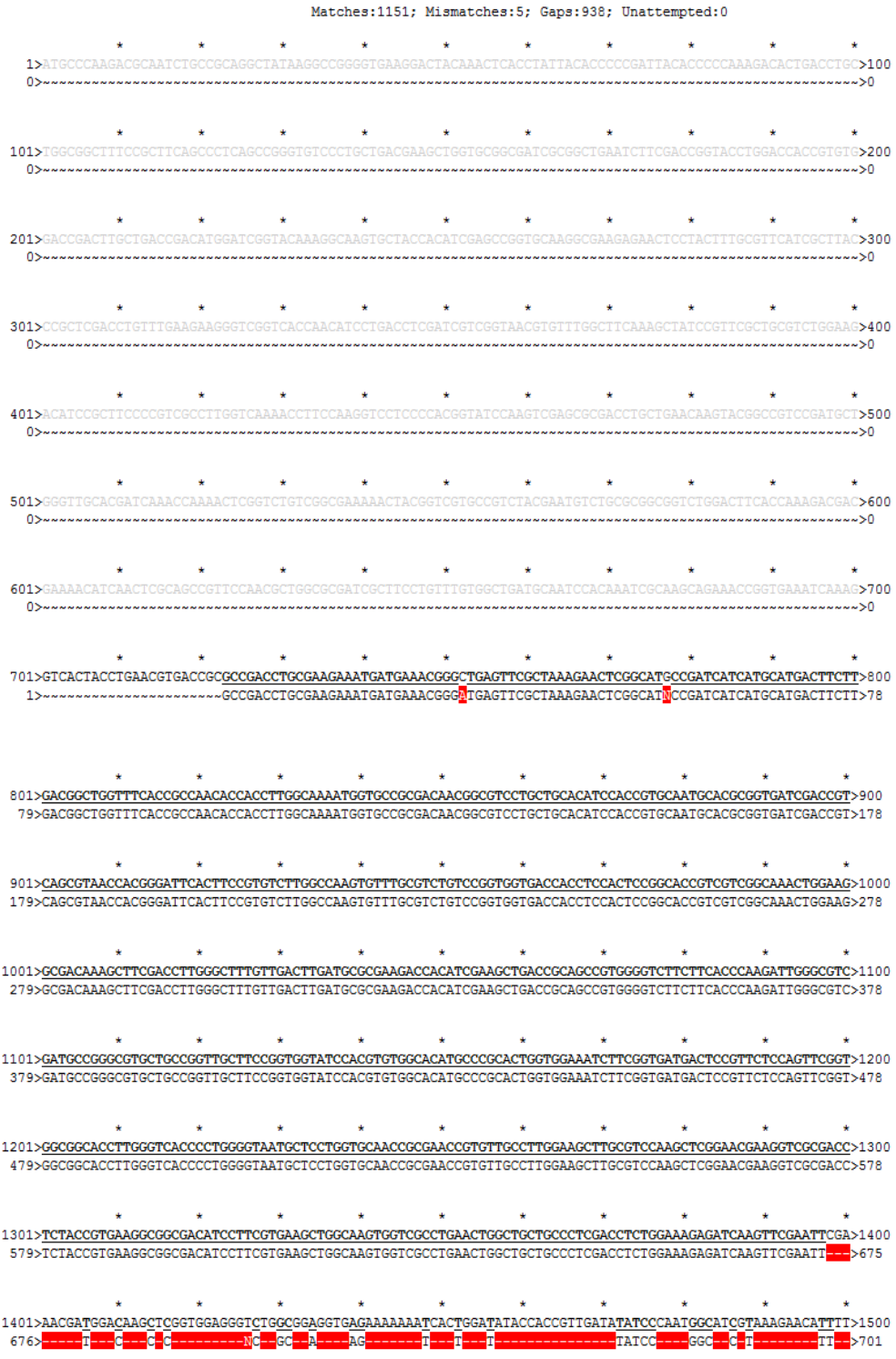
T7 promoter: bases 21-40
Ribosome binding site (RBS): bases 86-92
Initiation ATG: bases 101-103
6xHis tag: bases 113-130
*att*R1: bases 140-264
Chloramphenicol resistance gene (Cm^R): bases 373-1032
*ccd*B gene: bases 1374-1679
*att*R2: bases 1720-1844
T7 transcription termination region: bases 1855-1983
bla promoter: bases 2471-2569
Ampicillin (*bla*) resistance gene: bases 2570-3430
pBR322 origin: bases 3575-4248
ROP ORF: bases 4619-4810 (C)

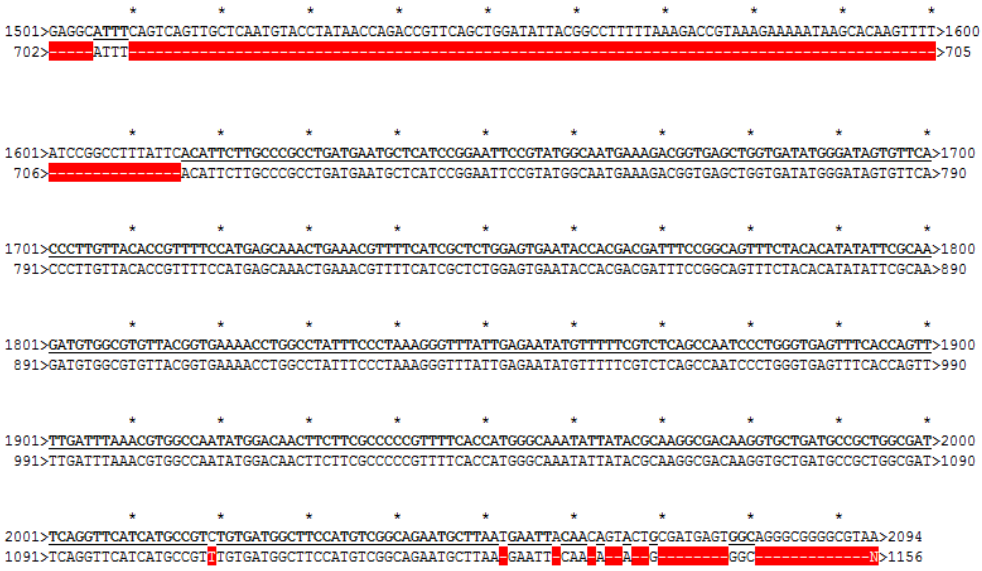
C=complementary strand

(Invitrogen Corporation 2012)

APPENDIX F

Sequence Alignment of Expected (Syn)*rbc*L-*CAT* Sequence and PCR Product of Amplification of (Syn)*rbc*L-*CAT*




1501>GAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTT>1600
702>-----ATTT--->705
1601>ATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCA>1700
706>---------------ACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCA>790
1701>CCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAA>1800
791>CCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAA>890
1801>GATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTT>1900
891>GATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTT>990
1901>TTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGAT>2000
991>TTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGAT>1090
2001>TCAGGTTCATCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAA>2094
1091>TCAGGTTCATCATGCCGT-TGTGATGGCTTCCATGTCGGCAGAATGCTTAA-GAATT-CAA-A--A--G--------GGC----------N>1156

APPENDIX G

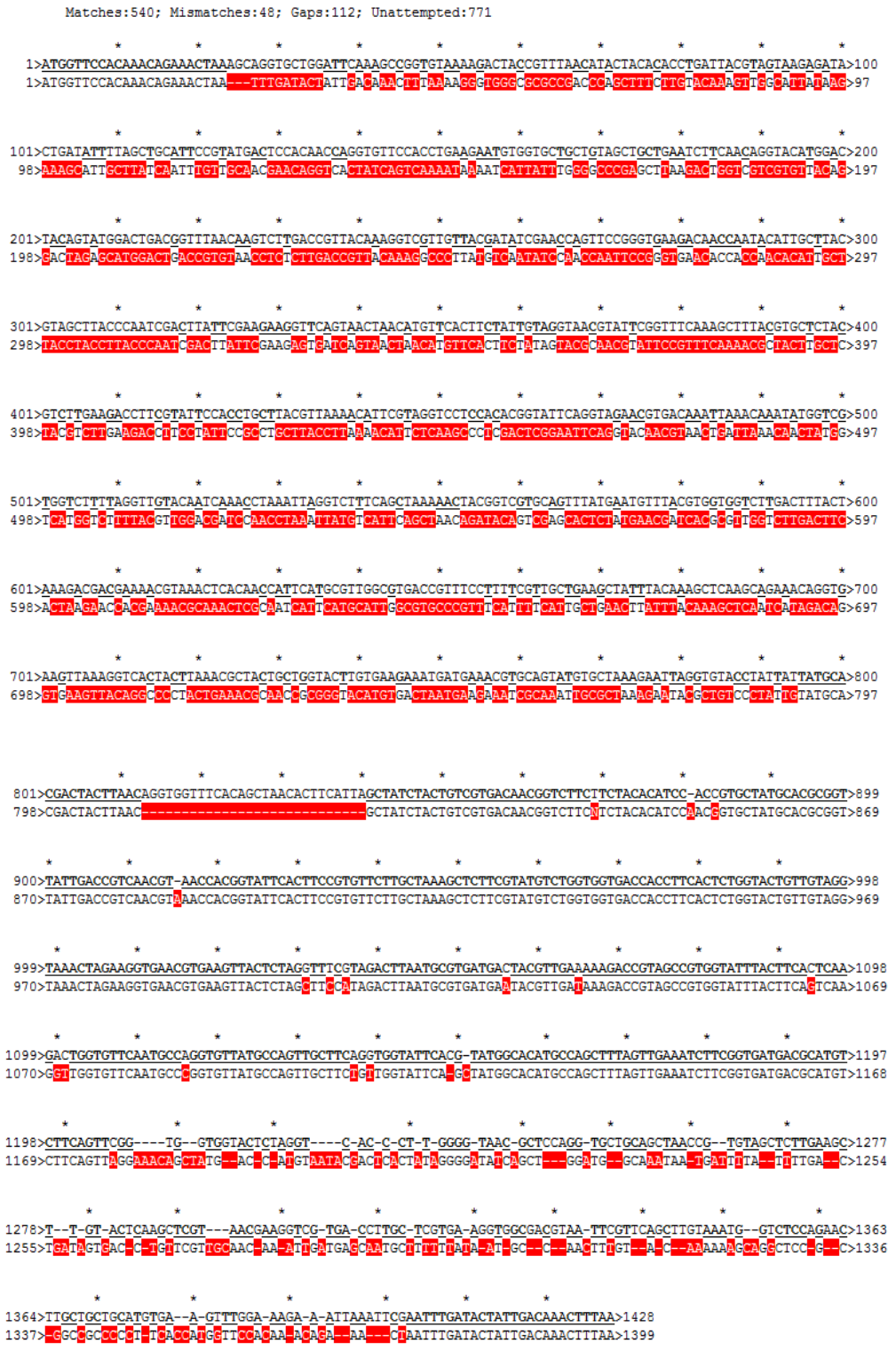
Sequence Alignment of Expected (Chl)*rbc*L-*CAT* Sequence with Recombinant pENTR/D-TOPO Extracted from Clone CC1

```
Matches:2105; Mismatches:1; Gaps:0; Unattempted:0

          *         *         *         *         *         *         *         *         *         *
  1>ATGGTTCCACAAACAGAAACTAAAGCAGGTGCTGGATTCAAAGCCGGTGTAAAAGACTACCGTTTAACATACTACACACCTGATTACGTAGTAAGAGATA>100
  1>ATGGTTCCACAAACAGAAACTAAAGCAGGTGCTGGATTCAAAGCCGGTGTAAAAGACTACCGTTTAACATACTACACACCTGATTACGTAGTAAGAGATA>100

          *         *         *         *         *         *         *         *         *         *
101>CTGATATTTTAGCTGCATTCCGTATGACTCCACAACCAGGTGTTCCACCTGAAGAATGTGGTGCTGCTGTAGCTGCTGAATCTTCAACAGGTACATGGAC>200
101>CTGATATTTTAGCTGCATTCCGTATGACTCCACAACCAGGTGTTCCACCTGAAGAATGTGGTGCTGCTGTAGCTGCTGAATCTTCAACAGGTACATGGAC>200

          *         *         *         *         *         *         *         *         *         *
201>TACAGTATGGACTGACGGTTTAACAAGTCTTGACCGTTACAAAGGTCGTTGTTACGATATCGAACCAGTTCCGGGTGAAGACAACCAATACATTGCTTAC>300
201>TACAGTATGGACTGACGGTTTAACAAGTCTTGACCGTTACAAAGGTCGTTGTTACGATATCGAACCAGTTCCGGGTGAAGACAACCAATACATTGCTTAC>300

          *         *         *         *         *         *         *         *         *         *
301>GTAGCTTACCCAATCGACTTATTCGAAGAAGGTTCAGTAACTAACATGTTCACTTCTATTGTAGGTAACGTATTCGGTTTCAAAGCTTTACGTGCTCTAC>400
301>GTAGCTTACCCAATCGACTTATTCGAAGAAGGTTCAGTAACTAACATGTTCACTTCTATTGTAGGTAACGTATTCGGTTTCAAAGCTTTACGTGCTCTAC>400

          *         *         *         *         *         *         *         *         *         *
401>GTCTTGAAGACCTTCGTATTCCACCTGCTTACGTTAAAACATTCGTAGGTCCTCCACACGGTATTCAGGTAGAACGTGACAAATTAAACAAATATGGTCG>500
401>GTCTTGAAGACCTTCGTATTCCACCTGCTTACGTTAAAACATTCGTAGGTCCTCCACACGGTATTCAGGTAGAACGTGACAAATTAAACAAATATGGTCG>500

          *         *         *         *         *         *         *         *         *         *
501>TGGTCTTTTAGGTTGTACAATCAAACCTAAATTAGGTCTTTCAGCTAAAAACTACGGTCGTGCAGTTTATGAATGTTTACGTGGTGGTCTTGACTTTACT>600
501>TGGTCTTTTAGGTTGTACAATCAAACCTAAATTAGGTCTTTCAGCTAAAAACTACGGTCGTGCAGTTTATGAATGTTTACGTGGTGGTCTTGACTTTACT>600

          *         *         *         *         *         *         *         *         *         *
601>AAAGACGACGAAAACGTAAACTCACAACCATTCATGCGTTGGCGTGACCGTTTCCTTTTCGTTGCTGAAGCTATTTACAAAGCTCAAGCAGAAACAGGTG>700
601>AAAGACGACGAAAACGTAAACTCACAACCATTCATGCGTTGGCGTGACCGTTTCCTTTTCGTTGCTGAAGCTATTTACAAAGCTCAAGCAGAAACAGGTG>700

          *         *         *         *         *         *         *         *         *         *
701>AAGTTAAAGGTCACTACTTAAACGCTACTGCTGGTACTTGTGAAGAAATGATGAAACGTGCAGTATGTGCTAAAGAATTAGGTGTACCTATTATTATGCA>800
701>AAGTTAAAGGTCACTACTTAAACGCTACTGCTGGTACTTGTGAAGAAATGATGAAACGTGCAGTATGTGCTAAAGAATTAGGTGTACCTATTATTATGCA>800



          *         *         *         *         *         *         *         *         *         *
 901>ATTGACCGTCAACGTAACCACGGTATTCACTTCCGTGTTCTTGCTAAAGCTCTTCGTATGTCTGGTGGTGACCACCTTCACTCTGGTACTGTTGTAGGTA>1000
 901>ATTGACCGTCAACGTAACCACGGTATTCACTTCCGTGTTCTTGCTAAAGCTCTTCGTATGTCTGGTGGTGACCACCTTCACTCTGGTACTGTTGTAGGTA>1000

          *         *         *         *         *         *         *         *         *         *
1001>AACTAGAAGGTGAACGTGAAGTTACTCTAGGTTTCGTAGACTTAATGCGTGATGACTACGTTGAAAAAGACCGTAGCCGTGGTATTTACTTCACTCAAGA>1100
1001>AACTAGAAGGTGAACGTGAAGTTACTCTAGGTTTCGTAGACTTAATGCGTGATGACTACGTTGAAAAAGACCGTAGCCGTGGTATTTACTTCACTCAAGA>1100

          *         *         *         *         *         *         *         *         *         *
1101>CTGGTGTTCAATGCCAGGTGTTATGCCAGTTGCTTCAGGTGGTATTCACGTATGGCACATGCCAGCTTTAGTTGAAATCTTCGGTGATGACGCATGTCTT>1200
1101>CTGGTGTTCAATGCCAGGTGTTATGCCAGTTGCTTCAGGTGGTATTCACGTATGGCACATGCCAGCTTTAGTTGAAATCTTCGGTGATGACGCATGTCTT>1200

          *         *         *         *         *         *         *         *         *         *
1201>CAGTTCGGTGGTGGTACTCTAGGTCACCCTTGGGGTAACGCTCCAGGTGCTGCAGCTAACCGTGTAGCTCTTGAAGCTTGTACTCAAGCTCGTAACGAAG>1300
1201>CAGTTCGGTGGTGGTACTCTAGGTCACCCTTGGGGTAACGCTCCAGGTGCTGCAGCTAACCGTGTAGCTCTTGAAGCTTGTACTCAAGCTCGTAACGAAG>1300

          *         *         *         *         *         *         *         *         *         *
1301>GTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAAATGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATT>1400
1301>GTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAAATGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATT>1400

          *         *         *         *         *         *         *         *         *         *
1401>CGAATTTGATACTATTGACAAACTTTAAGGTGGAGGGTCTGGCGGAGGTGAGAAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGT>1500
1401>CGAATTTGATACTATTGACAAACTTTAAGGTGGAGGGTCTGGCGGAGGTGAGAAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGT>1500
```

```
           *       *       *       *       *       *       *       *       *       *
1501>AAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATA>1600
1501>AAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATA>1600


           *       *       *       *       *       *       *       *       *       *
1701>GGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACAC>1800
1701>GGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACAC>1800

           *       *       *       *       *       *       *       *       *       *
1801>ATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGA>1900
1801>ATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGA>1900

           *       *       *       *       *       *       *       *       *       *
1901>GTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGAT>2000
1901>GTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGAT>2000

           *       *       *       *       *       *       *       *       *       *
2001>GCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGG>2100
2001>GCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGG>2100


2101>GCGTAA>2106
2101>GCGTAA>2106
```

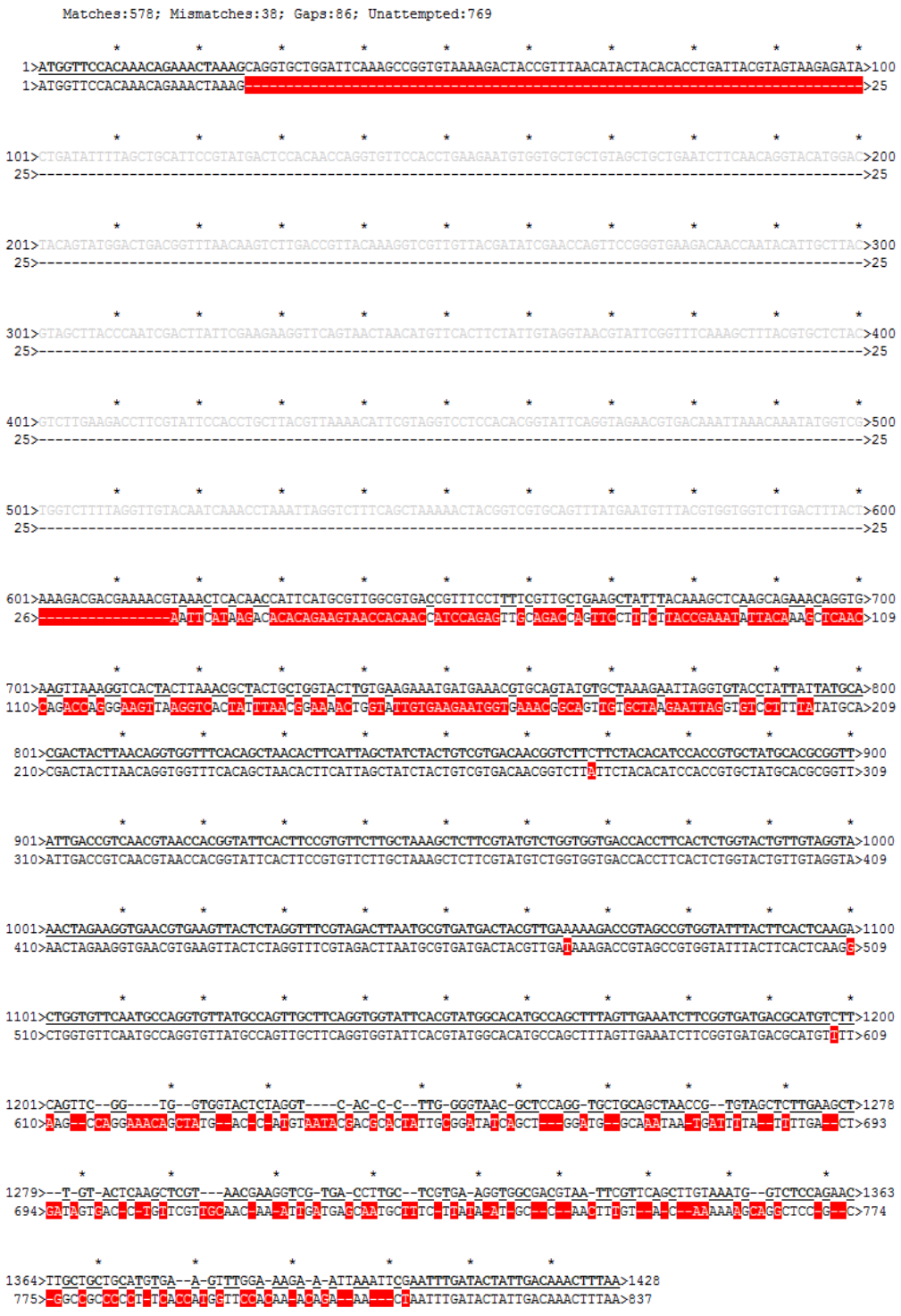
APPENDIX H

Sequence Alignment of (Chl)*rbc*L Sequence with Recombinant pENTR/D-TOPO Extracted from Clone C12



APPENDIX I

Sequence Alignment of (Chl)*rbc*L Sequence with Recombinant pENTR/D-TOPO Extracted from Clone C18


Matches:578; Mismatches:38; Gaps:86; Unattempted:769