Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

Eirik Solbakken

Engineering of Orthogonal Aminoacyl tRNA Synthetase/tRNA Pairs for Codon Reassignment in the Chloroplast of *Chlamydomonas reinhardtii*

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Martin F. Hohmann-Marriott
July 2020

NTNU
Norwegian University of
Science and Technology

Eirik Solbakken

Engineering of Orthogonal Aminoacyl tRNA Synthetase/tRNA Pairs for Codon Reassignment in the Chloroplast of *Chlamydomonas reinhardtii*

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Martin F. Hohmann-Marriott
July 2020

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

Preface

This master thesis concludes my degree in Chemical Engineering and Biotechnology (M.Sc.) at the Norwegian University of Science and Technology (NTNU) in Trondheim. The project was carried out at the Department of Biotechnology and Food Science (IBT).

I want to thank my supervisor, Professor Martin F. Hohmann-Marriott, for letting me take part in this project and all his guidance along the way. A special thanks go to PhD candidate Maxime Fages-Lartaud. His help and guidance have been invaluable and greatly appreciated. Furthermore, I would like to thank the other members of the PhotoSynLab for good company and assistance.

With the challenging situation caused by the ongoing COVID-19 pandemic, gratitude is expressed towards the way NTNU and especially IBT have handled it. Their focus on organizing an as safe as possible working environment, allowing us master's students to finish up most of our lab work is much appreciated.

Eirik Solbakken Trondheim, July 1, 2020

Summary

Microalgae are emerging as an attractive production platform for the synthesis of recombinant proteins and other biobased products. The chloroplast maintains its own protein synthesis machinery and a compact chromosome, which can be easily manipulated, establishing it as an attractive site within the field of synthetic biology. Features like these make the chloroplast an ideal target for codon reassignment to alter or expand the genetic code through the utilization of additional amino acids. Incorporating such non-canonical amino acids (ncAA) during protein synthesis could improve existing characteristics or facilitate the acquisition of new functionalities. Aminoacyl tRNA synthetases (aaRS)/tRNA pairs have been previously utilized for codon reassignment in several microorganisms. Yet, so far, the chloroplast of microalgae has not been targeted.

The work presented in this thesis establishes the molecular tools to identify and utilize orthogonal translation systems for the incorporation of ncAA during protein synthesis in the chloroplast of the microalgae *Chlamydomonas reinhardtii*.

Vectors containing genetic constructs for reassigning opal stop codon were constructed, a requirement for the incorporation of ncAA. Genes encoding pyrrolysyl-tRNA synthetase from *Methanomethylophilus alvus*, *Methanosarcina barkeri* and *Methanosarcina mazei* that have been previously optimized for enhanced affinity towards different ncAA, were genetically combined with genes for cognate opal-suppressor tRNA. All vectors were assembled with *aadA* gene conferring spectinomycin resistance and flanking sequences for homologous recombination within the *psbH* chloroplast locus. The gene for fluorescent protein mVenus, engineered to contain an additional opal codon, was included as the reporter gene. Functionally scarless vectors were constructed utilizing the Start-Stop Assembly, where start and stop codons served as fusion sites between promoter/5'untranslated region (UTR), coding sequence and 3'UTR of each gene. These vectors provide an efficient toolbox for establishing, analyzing and refining the incorporation of ncAA in proteins synthesized in the chloroplast of *Chlamydomonas reinhardtii*.

Sammendrag

Mikroalger etablerer seg som en attraktiv produksjonsplatform for rekombinante proteiner og andre biobaserte produkter. Kloroplasten opprettholder sitt eget maskineri for proteinsyntese og et kompakt kromosom, som på enkelt vis kan manipuleres og dermed etablerer organellen som et spennende område innenfor syntetisk biologi. Slike egenskaper gjør kloroplasten til et ideelt mål for kodon omdisponering for å endre eller utvide den genetiske koden via bruk av flere aminosyrer. Inkorporering av slike ikke-kanoniske aminosyrer under proteinsyntese kan både forbedre eksisterende egenskaper eller sørge for tilegning av nye egenskaper. Aminoasyl tRNA syntetase/tRNA parr har tidligere vært benyttet for kodon omdisponering i flere ulike mikroorganismer, men til nå har ikke mikroalgers kloroplast vært et mål.

Arbeidet som er presentert i denne avhandlingen etablerer molekylære verktøy for å kunne identifisere og utnytte ortogonale translasjonsystemer for inkorporering av ncAA under proteinsyntese i kloroplasten til mikroalgen *Chlamydomonas reinhardtii*.

Vektorer som inneholder genetisk konstrukt for omdisponering av opal stoppkodon ble konstruert, som er en betingelse for å kunne inkorporere ikke-kanoniske aminosyrer. Gener som koder for pyrrolysyl-tRNA syntetaser fra *Methanomethylophilus alvus*, *Methanosarcina barkeri* and *Methanosarcina mazei* ble genetisk kombinert med gener for opal-supressor tRNA. Disse pyrrolysyl-tRNA syntetasene har tidligere blitt optimalisert for affinitet til ikke-kanoniske aminosyrer. Alle vektorene ble konstruert med et *aadA* gen som sørger for resistans mot spektinomysin. Flankerende sekvenser ble også benyttet for å tilrettelegge for homolog rekombinasjon med kloroplast lokuset *psbH*. I tillegg ble et gen som koder for fluoriserende mVenus, med ett ekstra opal kodon innsatt, benyttet som rapportørgen.'Start-Stop assembly' ble benyttet for å konstruere funksjonelt arrløse gener. Start og stopp kodoner fungerte som fusjonsområder mellom promoter/5'utranslatert region, kodende sekvens og 3'utranslatert region. Disse vektorene danner en effektiv verktøykasse for å etablere og analysere inkorporering av ikke-kanoniske aminosyrer i proteiner syntetisert i kloroplasten til *Chlamydomonas reinhardtii*.

Abbreviations

AA	Amino acid
aaRS	Aminoacyl-tRNA syntethase
AlocLys	Nϵ-alloloxycarbonyl-L-lysine
ATP	Adenosine triphosphate
BB	BioBrick
BocK	Nϵ-Boc-L-lysine
bp	Base pair
C. reinhardtii	*Chlamydomonas reinhardtii*
CDS	Coding sequence
CbzK	Nϵ-benzyloxycarbonyl-K-lysin
Cyc-K	Nϵ-cyclopentyloxycarbonyl-L-lysin
CypK	Nϵ(((2-methylcycloprop-2-en-1-yl)methoxy) carbonyl)-L-lysine
D-prolyl-lysin	Nϵ-D-prolyl-L-lysin
E. coli	*Escherichia coli*
LB	Lysogeni Broth
M. barkeri	*Methanosarcina barkeri*
*Ma*PylRS	*Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase
*Mb*PylRS	*Methanosarcina barkeri* pyrrolysyl-tRNA synthetase
*Mj*TyrRS	*Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase
*Mm*PylRS	*Methanosarcina mazei* pyrrolysyl-tRNA synthetase
Me-His	3-methyl-L-histidine
mRNA	Messenger ribonucleic acid
mVenus	Monomeric Venus
N-Ac-lys	N$^{\epsilon}$-acetyllysin
ncAA	Non-canonical amino acid
NitroTyr	3-nitrotyrosine
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
PylRS	Pyrrolysyl-tRNA syntethase
TAE	Tris-Acetate-EDTA
TAP	Tris-Acetate-Phosphate
TyrRS	Tyrolysyl-tRNA syntethase
RBS	Ribosome binding site
RF	Release factor
UTR	Untranslated region
Wt	Wild type
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
YGE	Crude yeast genomic DNA extraction

Contents

1 Introduction

In recent years there has been an increasing interest in utilizing microalgae as a production platform for various products [1, 2]. Microalgae are rich in carbon containing-compounds. They can, therefore, serve as a sustainable and renewable source for products such as pharmaceuticals, biofuels or food ingredients [3]. Genetically engineered organisms have become a popular tool for the production of proteins and other substances. To date, most such systems utilize prokaryotes such as *Escherichia coli* (*E. coli*) or yeast such as *Saccharomyces cerevisiae*. However, as microalgae are emerging as a new production platform, research has also focused on genetically engineering these organisms to express different recombinant proteins [4]. Production based on proteins encoded in the nucleus has been reported. However, the chloroplast has received special interest due to its potential as a site for the synthesis of different bio-based products. These include products such as therapeutic proteins, vaccine compounds and enzymes with an industrial application. The advantages of using the chloroplast for heterologous protein expression involve high expression rate, cost-efficacy and swift scalability. Furthermore, the chloroplasts come with the relative ease of genetic manipulation compared to the nucleus' genome. The absence of certain toxins is also a desirable trait of the chloroplast [5]. The chloroplast is capable of performing several post-translational modifications necessary for a wide range of mammalian proteins, a property lacking in many prokaryotic organisms [6, 7]. As the global demand for recombinant proteins rise, the need for developing and improving effective systems increases. To date, there exist established techniques to synthesize recombinant proteins using bacteria, mammalian cells, yeast, plants and transgenic animals. Each of these systems come with their own set of advantages and disadvantages [8, 9, 10].

Green microalgae are a group of photosynthetic eukaryotes. Unlike bacteria, these organisms contain metabolic pathways leading to post-transcriptional and post-translational modifications necessary for the production of different proteins. Furthermore, the use of microalgae brings along advantages such as low growth cost, high growth rate and relatively simple cultivation. Microalgae are photoautotrophs, capable of utilizing light and carbon dioxide to generate energy through photosynthesis. Solar energy can thus be harnessed to generate energy to fuel the production of different biomolecules such as fatty acids, vitamins, antibiotics and other natural and recombinant proteins [5].

Protein synthesis usually utilizes 20 different amino acids (AA), the proteinogenic AA. With few exceptions, all organisms contain a genetic code encoding the same AAs. In total, 61 different codons, triplets of nucleotides, direct the incorporation of these 20 AA into the growing polypeptide. The genetic code is shared between all three domains (procaryotes, archaea and eukaryotes). Hence, which codons correspond to which AA is considered universal. The three remaining codons are stop codons, initiating termination of translation. To date, several techniques exist for the incorporation of AA, which are not part of these 20 proteinogenic AA [11]. Non-canonical amino acids (ncAA), are amino acids that most organisms do not utilize during protein synthesis. Although, there exist some organisms like methanogenic archaea that do. These organisms contain aminoacyl tRNA synthetase (aaRS)/tRNA pairs specific for the incorporation of such ncAA. Often, what is generally recognized as stop codons are the codons that initiate the incorporation of such ncAA. These naturally existing systems can, and have already been used to introduce ncAA in proteins synthesized in other organisms[12, 13].

1.1 Orthogonal Translation Systems for ncAA Incorporation

To expand the genetic code for inclusion of ncAA, research has focused on the development and use of orthogonal translation systems. In this sence, orthogonality referes to pathways for incorporation of AA which do not cross react with a hosts endogenous translation system [14]. The significant elements of such a system are: genetically engineered aaRS, their cognate tRNA and ncAA recognized both by the tRNA and aaRS [15]. To achieve an orthogonal system, it is imperative to introduce aaRS and tRNA which do not interact with the organisms own aaRS and tRNA, as not to disturb their natural function. The orthogonal aaRS cannot recognize any of the endogenous tRNA, which could lead to the incorporation of the wrong AA at undesired locations. Neither can the orthogonal tRNA be recognized by any endogenous aaRS, for the same reasons. Often, such orthogonal aaRS/tRNA pairs are found in organisms that are phylogenetically distant [15]. Recognition of tRNA by corresponding aaRS relies on short nucleotide sequences within the tRNA, called identity elements and anti-identity elements. Identity elements help activate aminoacylation by cognate aaRS, while anti-identity elements prevent the same reaction by non-cognate aaRS. Such sequences are usually found within the acceptor stem or anticodon region of the tRNA. Two tRNA with different identity elements could, therefore, potenitially be orthogonal towards each other, as they are likely recognized by different aaRS. The entire aaRS/tRNA pairs would then not cross react with each other and, therefore, be orthogonal [16].

An orthogonal system which offers site-specific incorporation of ncAA could improve the ability to manipulate both protein structure and function. The ncAA could alter the chemical and physical properties of the protein. Furthermore, fluorescent ncAA could be utilized as biophysical probes, allowing detection and study of protein structure and function *in vitro* as well as *in vivo* [17, 18].

1.2 Project Aim

This thesis is part of a larger project where the aim is to identify and utilize orthogonal translation system for codon reassignment to incorporation of ncAA in chloroplast proteins of the microalgae *Chlamydomonas Reinhardtii* (*C. reinhardtii*). To ahieve this goal, genetic vectors for gene delivery were constructed, providing the required genes for the incorporation of ncAA. As a response to the increased focus on microalgae, a system for incorporating ncAA within proteins produced in the chloroplast could potentially contribute to several opportunities within heterologous protein expression.

These vectors contain aaRS and tRNA genes that are to be introduced into the chloroplast DNA (cpDNA) of *C. reinhardtii*. This could enable the algae to synthesize the proteins needed for the expanded translational machinery. Plasmids are used as gene vectors, introducing the necessary genes into the algae. These plasmids contain homologous regions, flanking the sequence of the cpDNA where the genes are introduced, allowing homologous recombination between the vector and target site, thereby introducing the genes of the aaRS/tRNA pairs. An illustration of such a vector can be seen in Figure 1.1. In addition to the homologous flanking regions, the vector contains four gene cassettes. Within these cassettes are genes encoding the aaRS, tRNA and reporter gene used to verify the successful utilization of the system. The vectors also include a selectable marker, providing the possibility for effective screening of organisms containing the correct inserts.

aadA is used as the selectable marker. The *aadA* gene encodes an adenylyltransferase, conferring resistance towards the antibiotics streptomycin and spectinomycin [19]. Screening for correct phenotype could be carried out by supplying the growth medium with one of these antibiotics, which are toxic to wild-type (wt) *C. reinhartii*. Additionally, the antibiotic resistance will help drive the evolution of a homoplasmic cell line, containing only copies of recombinant cpDNA with the insert of interest. This is a result that could be achieved through gradually increasing antibiotic concentration. With increasing antibiotic concentration, the algal cells with multiple copies of the *aadA* gene will have a higher survival rate, which can drive the population towards homoplasmy. [20, 21].

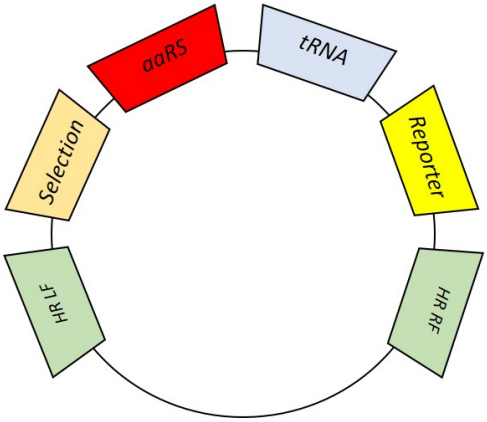


Figure 1.1: Vector containing four gene cassettes: reporter gene, aminoacyl tRNA synthethase (aaRS), tRNA and selection gene, in addition to left flank (LF) and right flank (RF) for homologouse recombination (HR).

To observe if the system is functional within *C. reinhartii*, a reporter gene encoding monomeric Venus (mVenus) is also included in the vector. mVenus is an improved variant of a yellow fluorescent protein, which fold faster, more efficiently and have increased brightness. Additionally, dimerization tendency of the yellow fluorescent protein have been eliminated in this variant [22]. The gene encoding mVenus is hereby referred to as *mVenus*. As the tRNA/aaRS pairs incorporate ncAA in response to stop codons, such codons are introduced into the mVenus gene. If the system functions properly, mVenus will be translated with additional ncAA. Previouse research indicates that mVenus most likely keeps its fluorescent properties, even with an additional ncAA at specific sites [23]. If the orthogonal system does not function, mVenus will not be fully synthesized as translation will terminate upon reaching the introduced stop codon. Hence, the detection of fluorescence will indicate the successful integration of the altered translation system. If fluorescence within *C. reinhardtii* is not detected; the orthogonal system could be proven unsuccessful. In total, three different mVenus mutants are generated. Two mVenus contain the amber or opal stop codon, respectively, while the third version contains both stop codons.

Stop codons initiate the termination of messenger ribonucleic acid (mRNA) translation. There exist three known stop codons, universal for all organisms. These have been given the names ochre (UAA) amber (UAG) and opal (UGA) stop codon. Upon interaction between the ribosome and the stop codons, certain release factors are recruited. Release factors activate the large ribosomal subunit, the peptidyl transferase center, which is responsible for catalyzing the cleavage of tRNA from the peptide chain. Through hydrolysis

of ester bonds, the growing peptide chain is released from the ribosome and tRNA, ending the translation process [24]. Due to sequencing of *C. reinhartii's* entire plastome, it is known that the opal codon does not occur at all within the cpDNA. Therefore, this spare codon can be introduced into the DNA and re-assigned for the incorporation of ncAA without disturbing translation termination elsewhere in the plastome. The alga has a strong preference for the ochre stop codon. A total of 65 out of the 69 protein encoding genes within the cpDNA utilize the ochre stop codon. The remaining four genes use the amber stop codon [5].

1.3 Aminoacyl-tRNA Syntethases Role in Protein Translation

During translation, aaRS are responsible for attaching AA to their corresponding tRNA. aaRS are enzymes, catalyzing the formation of covalent bonds between AA and tRNA. Every aaRS is specific for certain AA. Linking the AA to its cognate tRNA is achieved through two reactions. Initially, the amino acid is adenylated using adenosine triphosphate (ATP), as can be seen in Equation (1.1) [25].

$$AA + ATP = AA - AMP + PP_i \tag{1.1}$$

Afterward, the adenylated AA (AA-AMP) is covalently linked to a hydroxyl group at the isoacceptor of tRNA with the release of adenosine monophosphate (AMP), as shown in Equation (1.2) [25].

$$AA - AMP + tRNA^{AA} \longrightarrow AA - tRNA^{AA} + AMP \tag{1.2}$$

Consensus sequences are DNA sequences with similar structure and function across different species. Based on such consensus sequences, properties and construction of catalytic domain aaRS have been categorized into two groups. Class I aaRS is defined by what is known as a Rossmann ATP binding fold, where the aminoacylation occurs. The enzymes' active site contains two consensus sequences Lysine-Methionine-Serine-Lysie-Serine (KMSKS) and Histidine-Isoleucine-Glycine-Histidine (HIGH). During activation of AA, KMSKS is involved in the stabilization of ATP, while the histidine residue within HIGH binds phosphate in ATP causing stabilization. Another common feature among class I aaRS is the tendency to be monomeric, while class II aaRS exists as oligomers. The second class of aaRS is defined by three different consensus sequences; motif 1, 2 and 3. It is believed that motif 1 is involved in signaling activation of the aaRS subunits, while motif 2 and 3 are important for the active site for aminoacylation [25]. Table 1.1 lists examples of aaRS belonging to each class.

Table 1.1: Classification of different aminoacyl tRNA syntethases based on consensus sequence, catalytic domain and other properties [25].

Class I	Class II
ArgRS, CysRS, GlnRS, GluRS, IleRS, LeuRS, LysRS-I, MetRS, TrpRS, TyrRS & ValRS	AlaRS, AsnRS, AspRS, GlyRS, HisRS, LysRS-II, PheRS, ProRS, PylRS, SepRS, SerRS & ThrRS

1.3.1 Pyrrolysyl-tRNA Synthetase

Pyrrolysine is a ncAA, derived from the canonical amino acid lysine. The ncAA was first discovered in methyltransferases belonging to the archaea *Methanosarcina barkeri* (*M. barkeri*). Incorporation during protein synthesis is directed by the amber codon [26]. As can be seen from Table 1.1, PylRS has been categorized as a class II aaRS. During translation, pyrrolysyl-tRNA synthetase (PylRS) catalyzes the esterification of pyrrolysine to its cognate tRNA^{Pyl}. The wt-tRNA contains the anticodon CUA, complementary to the amber stop codon, thereby inserting the AA in response to said codon [12]. Variants of PylRS have been isolated from several different archaea, which have since been re-programmed to modify the substrate specificity [13, 27, 28]. An overview of some of these PylRS variants are listed in Table 1.2.

Methanosarcina barkeri pyrrolysyl-tRNA synthetase

M. barkeri is known to incorporate the amino acid pyrrolysine in response to the amber codon. *M. barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS) together with its cognate amber suppressor $Mb\text{tRNA}^{Pyl}_{CUA}$ assures the incorporation of the ncAA in response to the amber codon. Previous research indicate that *E. coli* can be modified to express functionally *Mb*PylRS/$Mb\text{tRNA}^{pyl}_{CUA}$. Additionally, it has been confirmed that the *Mb*PylRS/$Mb\text{tRNA}^{pyl}_{CUA}$ pair do not interact with endogenous aaRS and tRNA, indicating orthogonality towards *E. coli's* already existing translation machinery [13]. tRNA^{Pyl}_{CUA} contains several characteristics in both its primary and secondary structure, which is believed to be the reason it does not interact with endogenous aaRS of *E. coli*. These deviations include the CUA anticodon, a larger acceptor stem and smaller D-loop than tRNA endogenous to *E. coli* [12]. It has further been confirmed that *Mb*PylRS is able to esterify several derivatives of lysine and pyrrolysine. Among these are Nϵ-D-prolyl-L-lysin (D-prolyl-lysin) and Nϵ-cyclopentyloxycarbonyl-L-lysin (CycK) [29].

*Mb*PylRS-Ack1 is a mutant of *Mb*PylRS designed for optimized attachment of N$^{\epsilon}$-acetyllysine (N-Ac-lys) to $Mb\text{tRNA}^{Pyl}_{CUA}$. The mutant aaRS contain six mutations: D76G, L226V, L270I, Y271F, L274A and C313F. The five last mutations are found within 6 Å of where the aaRS bind pyrrolysine's pyrroline ring. It is believed that these mutations cause a change in the aaRSs' hydrophobic cavity, a rearrangement that promotes binding to the acetyl group of N$^{\epsilon}$-acetyllysine, rather than the pyrroline ring of pyrrolysine. The mutations cause an enlarged aaRS volume. It is has been proposed that enlargement compensate for the volume difference between pyrrolysine and N-Ac-lys [13].

Methanomethylophilus alvus pyrrolysyl-tRNA synthetase

A second PylRS/tRNA^{Pyl} pair have been isolated from the methanogene *Methanomethylophilus alvus*. The *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase (*Ma*PylRS) has been identified as orthogonal to the endogenouse aaRS/tRNA of *E. coli*. *Ma*PylRS do not contain the N-terminal domain which first was believed to be necessary for *in vivo* activity of PylRS/tRNA^{Pyl}. In fact, studies show that *Ma*PylRS/tRNA^{Pyl}_{CUA} incorporate the ncAA Nϵ-Boc-L-lysine (BocK) more efficiently than *Mb*PylRS/tRNA^{Pyl}_{CUA} [28].

Several different variants of *Ma*PylRS have been synthesized in attempts to develop or increase affinity towards different AA. Two of these are *Ma*PylRS-RS1 (Y126M, M129G and V168T) and *Ma*PylRS-3MeHis (L121M, L125I, Y126F, M129A, V168F). *Ma*PylRS-RS1/tRNA$^{Pyl}_{CUA}$ is optimised to direct incorporation of the ncAA Nϵ-benzyloxycarbonyl-K-lysin (CbzK) [27]. *Ma*PylRS-3MeHis is a mutant of *Ma*PylRS, where the active site has been differentiated to re-program the substrate specificity. The mutant is known to incorporate 3-methyl-L-histidine (Me-His). However, it lacks the ability of wt-*Ma*PylRS to incorporate BocK [28].

Methanosarcina mazei pyrrolysyl-tRNA synthetase

A PylRS/tRNA*Pyl* pair has also been identified in *Methanosarcina mazei*. Research show that *Methanosarcina mazei* pyrrolysyl-tRNA synthetase (*Mm*PylRS) is capable of esterifying tRNA^{Pyl} with ncAA such as BocK, Nϵ-alloloxycarbonyl-L-lysine (AlocLys) and Nϵ(((2-methylcycloprop-2-en-1-yl)methoxy) carbonyl)-L-lysine (CypK) [12, 27]. As for the previously mentioned PylRS, *Mm*PylRS have been successfully used for the incorporation of ncAA in *E. coli*. The *Mm*PylRS/tRNA$^{Pyl}_{CUA}$ pair is reported orthogonal towards both *E. coli* and eukaryotic cells [27].

Wt-*Mm*PylRS/^{Pyl} can attach certain AA to *Ma*tRNA$^{Pyl}_{CUA}$. Therefor, *Mm*PylRS/tRNA$^{Pyl}_{CUA}$ and *Ma*PylRS/tRNA$^{Pyl}_{CUA}$ are not mutually orthogonal. However, variants of MatRNA^{Pyl} which to not interact with *Mm*PylRS have been developed. The N-terminal domain of *Mm*PylRS interacts with the variable loop of tRNA. *Ma*PylRS tRNA lack the N-termianl domain and interacts differently with the tRNA. Hence, by introducing mutations to the variable loop of *Ma*tRNA^{Pyl} to avoid assocciation with *Mm*PylRS/^{Pyl}, mutually orthogonal *Mm*PylRS/tRNA$^{Pyl}_{CUA}$ and *Ma*PylRS/tRNA$^{Pyl}_{CUA}$ have been developed [27].

A mutant version of *Mm*PylRS, *Mm*PylRS-AF (Y306A, Y384) is known to display higher *in vitro* aminoacylation and *in vivo* amber suppression with pyrrolysine, BocK, AlocLys than the wt-*Mm*PylRS. The mutation Y306A causes an expanse of the amino acid binding pockets, which allow binding of N$^{\epsilon}$-(o-acidobenzyloxycarbonyl)-L-lysin (AzZLys) and esterification to tRNA^{Pyl} [27, 12].

Wt-tRNA^{Pyl} for all the mentioned PylRS initially contains a CUA anticodon, responding to the amber codon (UAG), making it a natural occurring amber suppressor. It suppresses the terminal effect of the UAG stop codon by competing with release factors for the binding to the ribosome. Several PylRS is known to tolerate mutations within the anticodon of tRNA^{Pyl}, as it is often not considered identity element. This gave rise to the idea of changing the anticodon from CUA to UCA. This mutation alters the tRNAs specificity from recognizing the amber stop codon, to recognizing the opal stop codon [30]. For this project, the mutated tRNA$^{Pyl}_{UCA}$ is used together with all PylRS to incorporate ncAA in response to the opal codon.

1.3.2 Tyrosyl-tRNA Synthetase

Another aaRS is the tyrosyl-tRNA synthetase (TyrRS), which has been given its name due to its role in catalyzing attachment of tyrosine to the aaRS's cognate tRNA. Unlike PylRS, TyrRS is categorized as a class I aaRS. A TyrRS/tRNA isolated from the archaea *Methanocaldococcus jannaschii* (*Mj*TyrRS) have been used for codon reassignment in *E.*

coli. The archaea's and *E.coli's* tRNA^{tyr} contain different identity elements. Hence, *Mj*TyrRS/tRNA^{Tyr} have been identified as orthogonal to the endogenous aaRS/tRNA of *E. coli* [31].

Furthermore, this aaRS does not contain any form of editing domain, making it more suitable for ncAA incorporation. As a result of lacking an editing domain, it is believed that *Mj*TyrRS do not proofread if it is a ncAA ligated to the cognate tRNA [31]. Recognition and positioning of substrates at their binding site is rather directed by conserved active sites within TyrRS [32].

Advances have been made to change the anticodon of *Mj*tRNA to CUA, only slightly reducing its affinity for *Mj*TyrRS. With a CUA anticodon, the tRNA incorporates AA in response to the amber codon (UAG) [33]. In this project, such a tRNA with an anticodon for the amber codon is utilized. TyrRS/tRNA^{Tyr} originally ensures incorporation of tyrosine, one of the 20 standard AA. However, mutant variants *Mj*TyrRS-NitroTyr capable of efficiently esterifying the ncAA 3-nitrotyrosine (NitroTyr) have been developed [34].

Additionally, *Mj*TyrRS/tRNA$^{Tyr}_{CUA}$ is orthogonal towards both *Mm*PylRS/tRNA$^{Pyl}_{CUA}$ and *Mb*PylRS/tRNA$^{Pyl}_{CUA}$. Hence, using a combination of these aaRS/tRNA pairs, it is possible to facilitate the co-translational incorporation of two different ncAA simultaneously [27].

Table 1.2: A list of Pyrrolysyl-tRNA synthetases (PylRS) and a tyrosyl-tRNA synthetas (TyrRS) and cognate tRNA used in the project. These enzymes have been isolated with in *Methanomethylophilus alvus* (*Ma*), *Methanosarcina barkeri* (Mb), *Methanocaldococcus jannaschii* (Mj) and *Methanosarcina mazei* (Mm) and variants have been constructed for optimized ncAA affinity. Each pair is listed with examples of ncAA, which they have a confirmed affinity for.

aaRS	tRNA	ncAA
*Ma*PylRS-3MeHis	*Ma*tRNA$^{Pyl}_{CUA}$	Me-His
*Ma*PylRS-RS1	*Ma*tRNA$^{Pyl}_{CUA}$	CbzK
*Mb*PylRS	*Mb*tRNA$^{Pyl}_{CUA}$	Pyrrolysin, BocK, D-prolyl-lysin, Cyc-K
*Mb*PylRS-Ack1	*Mb*tRNA$^{Pyl}_{CUA}$	N-Ac-lys
*Mm*PylRS-AF	*Mm*tRNA$^{Pyl}_{CUA}$	CypK, pyrrolysin, Bock, AzZLys
*Mj*TyrRS-NitroTyr	*Mj*tRNA$^{Tyr}_{CUA}$	NitroTyr

1.4 Choice of Selection Cassettes for Chloroplast Engineering

To date, there are several different choices for selectable markers in algae. These are broadly categorized into antibiotic resistance, herbicide resistance or essential genes for photosynthesis. The rest are categorized as either positive selectable markers or other negative selectable markers. Although herbicide selectable markers have been developed for other algae, per 2018 no selective markers offering resistance in *C. reinhardtii* have been identified [21].

Among the selectable markers, the antibiotic gene *aadA* is the most widely used in chloroplast engineering. The coding sequence (CDS) of *aadA* is often fused with chloroplast promoters and untranslated regions (UTR) to achieve increased expression rate [21].

aadA encodes an aminoglycoside adenyl transferase, which, as earlier described, confer resistance towards spectinomycin and streptomycin. Although this is a bacterial gene, a stable transgenic expression conferring antibiotic resistance in *C. reinhardtii* has been confirmed [35].

Another choice for a selectable marker in *C. reinhardtii* is the use of the *aphA-6* gene isolated from the bacteria *Acinetobacter baumannin*. The gene encodes an aminoglycoside phosphatasetransferase, conferring resistance to the antibiotics kanamycin and amikacin [36]. Several factors have to be considered when developing a functioning antibiotic selective marker for the green algae. It has to be established that *C. reinhardtii* in fact is sensitive to the antibiotic within a suitable concentration range. Additionally, since the wt-algae depend on a light source for photosynthesis, a light-sensitive antibiotic would be undesirable. Thirdly, the selectable marker has to confer enough resistance for survival of heteroplasmic cells, where the genes' copy number could be low [21].

C. reinhardtii is capable of growing heterotrophically. Mutants carrying alterations in essential photosynthetic genes have been isolated and described [37]. This gave rise to the potential of selection based on the restoration of photosynthesis. In such cases, the selectable marker would be a fully functional wild-type version of the photosynthetic gene. Through homologous recombination, the mutated version can be replaced by the wild-type restoring photosynthetic functionality. A general scheme of how selection based on restoration of photosynthesis can be achieved is shown in Figure 1.2. A wild-type version is used in the gene vector together with other genes of interest and inserted at the site of the mutated gene. Restoration of the wt-gene would then allow for wt-phenotype selection, by growing the algae with a light source and medium without acetate. Furthermore, the use of such a marker eliminates the risk of transferring unwanted antibiotic resistance [21].

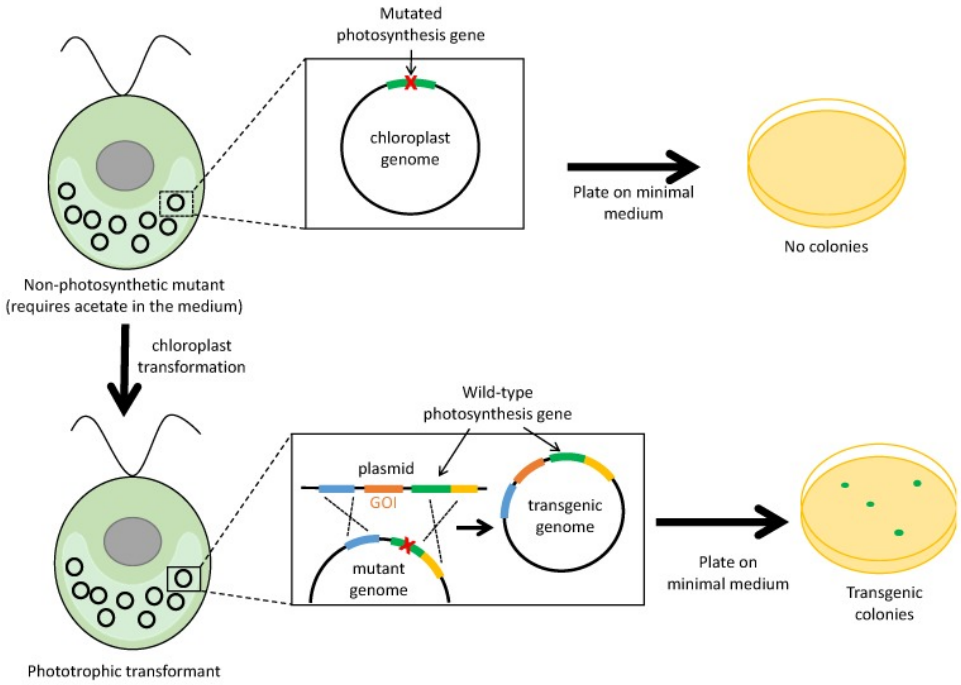


Figure 1.2: The use of a selectable marker containing wild-type gene involved in photosyntesis, allowing for selection based on restoration of photosynthesis in mutants with defective photosynthesis. Transgenic colonies are then able to grow on medium without acetate in the presence of a light source. The figure is adapted from Esland, L. et al [21]

1.5 Chloroplast Promoters and UTR for Enhanced Expression Rate and mRNA Stability

In order for the algae to successfully utilize the transgenic coding regions which are to be inserted, they have to be combined with appropriate promoters and untranslated regions (UTR). Gene promoters are nucleotide sequences where transcription is initiated. These sequences contain sites where RNA polymerase can bind and initiate the synthesis of mRNA [38]. 5'UTR help regulate mRNA stability and initiation of translation, while 3'UTR take part in mRNA stability and transcription termination [6].

Endogenouse regulatory elements are often used to efficiently express transgenes in the algal chloroplast. Among these are the promoters for genes encoding the large subunit of ribulose bisphophate carboxylase (*rbcl)* and a subunit of the ATP synthase (*atpA*). The promoters for the D1 protein (psbA) and D2 protein (psbD), which are part of photosystem II, have also been utilized in transgenic expression [39]. In a study performed by D. Barnes et al. the effect of 5'UTRs containing the promoters for *psbD*, *psbA*, *rbcl* and *atpA* were studied using a green fluorescent (GFP) reporter. It was found that the promoters for *psbD* and *atpA* gave the highest expression rate of chimeric mRNA and GFP. Varying the 3'UTR, however, did not seem to alter the expression rate much. As long as a 3'UTR was present, the expression rate did not significantly change [40]. The 5'UTR and promoter of *psbA* have promoted a high protein expression in *C. reinhardtii*. However, this was only achieved in D1-deficient strains. Another promising promoter is *C. reinhardtii's* own 16S rRNA promoter. Fused to 5'UTR regulatory elements such as those belonging to *psbA* and *atpA*, this promoter stimulates a high transgenic expression rate [6].

1.6 *Chlamydomonas reinhardtii* as a Model Organism

C. reinhardtii serve as the model organism for implementing and testing of the orthogonal translation system. Microalgae such as *C. Reinhardtii* could potentially become an important platform for the production of different recombinant proteins. Previously, *C. reinhardtii* has been engineered for the production of different proteins such as hormones, antibodies, reporter proteins and antigens to be included in vaccines [1]. *C. reinhardtii* is widely used as a model organism for unicellular algae, especially among studies focusing on photosynthesis, nutrition or cilia function and structure. In more recent years, the algae have also been used as a platform for the production of different bio-products and biofuel production [41].

The single-cell green algae grow as a haploid cell. As it only has one set of chromosomes, this allows for the immediate expression of mutant phenotype. This feature makes it an excellent laboratory model species. When supplied with acetate-containing medium, the normally photosynthetic organism, have been proven able to grow and still retain a functional photosynthetic apparatus. Hence, the growth of mutants with light-sensitive photosynthesis can be sustained [42]. The nuclear, chloroplast and mitochondrial genomes have all been sequenced. Having a sequenced genome is an important aspect of providing useful information in many situations. Additionally, procedures for transformation of exogenous DNA into all these three genomes have been established, making the organism well suited for genetic studies [21].

Chloroplasts are organelles found in both plants and green algae, containing their own genetic system. Furthermore, the chloroplast is an essential site for photosynthesis, but also biosynthesis of fatty acid, amino acid and secondary metabolites. Other vital processes in the organelle include nitrate and sulfate assimilation. However, a limited number of proteins involved in these processes are encoded by the cpDNA itself, as most are encoded within the nuclear DNA. The chloroplast genome is referred to as a plastome, and contain on average as few as 120-130 genes. Most of these genes encode the proteins involved in the chloroplasts' own gene expression or photosynthesis machinery. The chromosomes within the chloroplast are organized as nucleoids, consisting of approximately 20 copies of the plastome, a variety of proteins and RNA. It has been suggested that both mRNA processing, editing and ribosome assembly all occur co-transcriptionally and in association with the nucleoid, displaying a feature commonly found within prokaryotic organisms [43, 44].

The circular plastome of *C. reinhardtii* is approximately 206 kb and found within the single large chloroplast. Each cell can contain 50-80 copies of the plastome. The chloroplast genome has become an attractive site for genome editing and the production of recombinant proteins, especially due to its genome's simplicity compared to the nuclear genome. Other favorable traits include high levels of gene expression as well as the possibility to introduce whole operons, allowing multiple genes to be expressed and controlled simultaneously. The chloroplast can also function as a site for storage, having proven to accumulate high levels of different substances. Furthermore, *C. reinhardtii* has been recognized as GRAS (Generally Recognized As Safe), opening several possibilities such as the use of whole cells in vaccines or animal feed [45]. Due to evolutionary origin from cyanobacteria, the chloroplast has several traits similar to that of prokaryotes. Among these traits is the translation apparatus, allowing for the production and accumulation of substances which otherwise would have killed the cell due to toxicity [46].

Change in expression yield is an important aspect to consider when introducing transgenes to subcellular compartments. Expression yield is affected by both protein assembly, folding and post-translational modification. As an example, the generation of disulfide bridges and proper folding is essential to generate functional proteins. The chloroplast contains high concentrations of proteins called chaperons, which are required for protein folding. Additionally, the chloroplast contains an oxidizing environment, promoting the development of disulfide bridges. Conditions such as these elevate the chloroplast as an attractive site for the synthesis of recombinant proteins [47]. The chloroplast could also provide a more sheltered environment for the production of recombinant proteins. Algal chloroplast usually contains fewer proteases than the cytoplasm, reducing the proteolysis of proteins. On the other hand, downsides to using the chloroplast involve lack of performing complex post-translational modifications, such as glycosylation. These limitations should be taken into account when choosing which proteins to be synthesized. It is crucial to consider what effect the lack of these modifications has for the protein of interest [48].

1.7 Techniques for DNA Amplification and Assembly

Within the field of molecular biology, methods for assembling genetic parts are crucial. It is possible to chemically synthesize DNA strands *de novo* up to a certain length [49]. However, the development of larger parts still relies on enzymatic assembly methods. A more complicated DNA construct requires higher efficacy and specificity than what can be accomplished by the standard restriction-ligation reactions. Additionally, as DNA constructs often contain a large amount of DNA, selection of appropriate restriction site and enzymes have become increasingly difficult. It could be possible to identify suitable restriction sites for one construct. However, there is no guarantee they would fit other constructs, thereby limiting the modularity of such assembly methods. Hence, the development and use of both efficient and broadly applicable techniques to combine DNA are sought [50].

DNA assembly are methods to develop DNA constructs relatively quick and reliable, compared to standard restriction and ligation cloning. Often, these methods are separated into two categories: modular and bespoke methods. The modular assembly techniques are often more suitable for combinatorial assembly, combining different DNA parts (promoters, CDS, terminators, etc.) in a defined order. The high efficacy of modular methods stems from a system of highly specific steps. However, these methods often leave scarring at the links between the different parts of the construct. Additionally, modular methods are often more resource requiring and time-consuming [51].

Bespoke methods, also referred to as non-modular methods, are less suitable for combinatorial assembly because of low efficacy. On the other hand, these techniques have the advantage of creating scarless constructs. Additionally, DNA to be combined through bespoke methods require less preparatory work than in modular methods. However, bespoke methods rely on sequencing overlap between parts. This introduce the need for specific primers to get the correct sequences at the ends. Sequencing of the DNA is also required due to possible mistakes during amplification by polymerase chain reaction (PCR)

1.7.1 Type II Restriction Endonucleases for Plasmid Construction and DNA Degradation

Restriction endonucleases play a central role in most DNA assembly techniques. These are enzymes able to cut DNA strands at specific sequences. These enzymes are primarily found in bacteria, used as a defense mechanism to degrade foreign DNA. However, they have also been identified within archaea, viruses and eukaryotes. The enzymes are harvested and widely used within the field of biotechnology, as useful tools for DNA assembly and cloning. Different types of restriction endonucleases cut DNA differently and at different sequences. Some cut DNA unsymmetrical resulting in a sticky end, a single-stranded DNA overhang. Others create blunt ends as a result of cutting both DNA strands at the same position. Restriction endonucleases can be classified into four different types (I-IV), each with additional subclasses. Classification is based on restriction site, cleavage site, structure, activators and cofactors[52, 53]. All type II restriction enzymes share several similarities. These endonucleases are normally homodimers, which in the presence of Mg_2^+ cleave double-stranded DNA within or close to its recognition site. Additionally, these recognition sites are often palindromic. However, not all type II endonucleases fit the exact same definition. Therefore, different subclasses exist [54].

Figure 1.3 illustrates the mechanism of how type II restriction endonuclease bind DNA and catalyzes bond cleavage. Initially, the enzyme binds non-specifically to the DNA. Following attachment, the enzyme moves along the DNA in what is defined as a random diffusional walk until it identifies its corresponding recognition site. Upon binding to the recognition site, conformational changes lead to activation of the catalytic site. The enzyme catalyzes cleavage of phosphodiester bonds in both DNA strands. Finally, the enzyme either dissociate from the DNA completely or is transferred to another non-specific binding site on the DNA molecule [54].

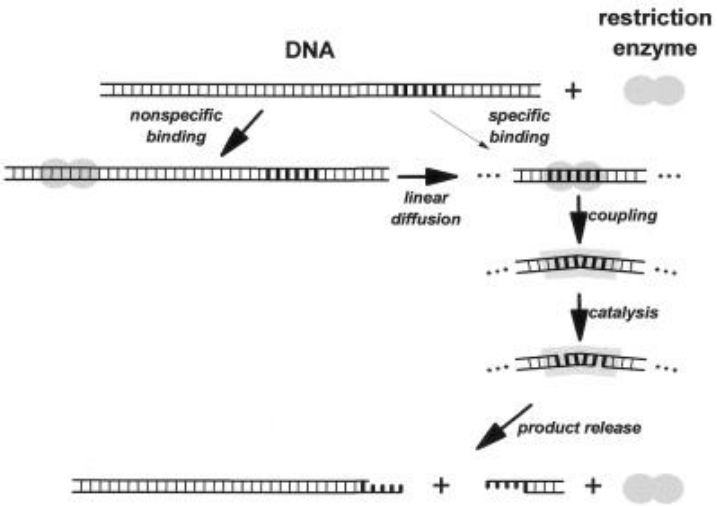


Figure 1.3: Schematic representation of the steps involved in type II restriction endonuclease recognition binding DNA up to cleavage of double stranded DNA. The restriction enzymes associate with the DNA at specific recognition sites, and catalyze cleavage of the DNA strands a certain length away, known as cleavage sites. The figure has been adapted from Pingoud & Jeltsch [54]

Among the subclasses are the type IIS restriction endonucleases. These enzymes interact

with asymmetric recognition sites, normally 4-7 base pair (bp) long, and a cleavage site where DNA is cut. These cleavage sites are usually found 1-20 bp from the recognition site. Since these enzyme catalyze cleavage of DNA strand outside the recognition site, they could be particularly useful in DNA assembly of smaller fragments. The recognition site could then be part of the fragment of interest and go unharmed by the restriction endonuclease. Overhanging sequences are thereby unrelated to the target sequence. However, these recognition sites are small and often quite abundant in DNA. Therefore, the use of type IIS restriction endonucleases for cutting longer DNA fragments could become problematic. Fragments could contain several of these enzymes recognition sites, resulting in multiple undesired cuts [55].

Two type IIS restriction endonucleases are central in the Start-Stop Assembly; SapI and BsaI. SapI recognizes the sequence 5'-GCTCTTC-3' and cleaves both strands in the DNA molecule. It cleaves the strands one and four nucleotides downstream of the recognition site, respectively. Such a cleaving pattern generates a 3 bp overhang [56]. BsaI recognizes the sequence 5'-GGTCTC-3' and cuts one and five nucleotides downstream of recognition site, respectively. Unlike SapI, BsaI generates a four bp long overhang [57].

DpnI is another type II restriction enzyme. This enzyme belongs to the IIM subclass, where M stands for modification-dependant. The enzyme recognizes methylated adenine and cleaves 5'-GATC-3' sites, which are either entirely or hemi-methylated. DNA is cleaved in the middle of this DNA sequence, resulting in blunt-end fragments. Bacteria such as *E. coli* uses methylation of DNA to separate "self" from non-self DNA. These bacteria often methylate their own DNA using methyltransferases, and specific endonucleases degrade unmethylated non-self DNA. However, this is not a universal rule. Other organisms recognize methylated DNA as foreign DNA, and contain endonucleases such as DpnI, which can degrade the methylated DNA. The use of the DpnI endonuclease is widely used within the field of biotechnology to degrade methylated DNA [58, 59].

1.7.2 Golden Gate Assembly

The Golden Gate Assembly is a one-pot, one-step modular method to assemble DNA fragments into a single plasmid. The method is based on the use of previously described type IIS restriction enzymes and T4 ligases [60]. DNA ligases such as the phage-encoded T4 ligase catalyze the formation of phosphodiester between 3'-hydroxyl and 5'-phosphate groups of DNA strands [61]. As described, type IIS restriction endonucleases cleave DNA leaving cohesive ends, overhangs of defined length and sequence. For Golden Gate Assembly, restriction enzymes such as BsaI, which generate four bp overhangs are typically used.

Additionally, as these enzymes cut outside their recognition sites, the site can be left out of the finished product. Golden Gate cloning allows for the assembly of at least nine DNA fragments on to an acceptor vector with up to 90% of the recombinant clones containing the correct construct. In order for the Golden Gate Assembly to function each end of the DNA fragments of interest must be flanked by recognition sites of type IIS restriction endonucleases. After digestion with correct restriction enzymes, the ends with overlapping sequences can then be ligated [60]. The cohesive ends can be designed to be non-palindromic. Hence, self-ligation, which often is a problem when using classical type II endonucleases and palindromic cohesive ends, can be avoided.

Furthermore, directionality during assembly can be achieved by designing distinct non-palindromic cohesive ends which only anneal to the cohesive end of one of the other parts. This is also why the reaction can be carried out in 'one-pot'. DNA fragments/plasmids, restriction endonucleases and ligases can all be added in one reaction without risking several unwanted products. Due to the explained design conferring directionality, only two possible outcomes are likely. Either the DNA fragments are cut and ligated to their intended cohesive end, or they ligate back onto their donor plasmid/DNA fragment. Due to loss of restriction site when ligated with the intended cohesive end, this assembled DNA fragment should not be cleavable by the restriction enzymes. With increasing reaction time, the concentration of donor fragment/plasmid will, therefore, decrease. Target DNA fragment/plasmid concentration will on the other hand increase [51].

Even though this approach offers a highly efficient DNA assembly, it comes with certain limitations. One of these limitations is the generation of scars within the finished DNA assembly. These scars are the fusion sites, the regions where the cohesive ends ligate with each other, which becomes integrated parts of the product. As an example, if a gene is to be assembled using Golden Gate, scars will be generated between each different part. Scars between the promoter and CDS will be transcribed as part of the mRNA and could negatively impact translation through undesired properties such as structure and accessibility [51].

1.7.3 Start-Stop Assembly

The Start-Stop Assembly was developed to overcome limitations of the already existing modular and bespoke assembly techniques. This multi-part modular DNA assembly system builds on the Golden Gate assembly but is both functionally scarless and usable for combinatorial DNA assembly [51].

As described, scars are normally created during modular multi-part DNA assembly due to the need for sequence identity at the DNA ends. This identity is required for the recognition and ligation of the DNA sequences of interest. To avoid scarring, the Start-Stop Assembly utilizes start and stop codons as fusion sites. These sequences are conserved motifs found at the beginning and end of all CDS, making them suitable fusion sites for DNA assembly. As these codons are already found at the beginning and end of CDS, no additional fusion site has to be used, and therefore no scars are created upon assembly [51].

Like many other DNA assemblies, the Start-Stop Assembly utilizes type IIS restriction endonucleases. However, unlike most methods, the restriction enzymes used here need to generate three bp long overhangs instead of the more standard four bp overhangs. The need for shorter overhangs is due to using three bp long start and stop codons as fusion sites. SapI is a restriction endonuclease with the required property, and can therefore be used for this technique [51].

During Start-Stop Assembly, all the different parts can be stored in separate plasmids, referred to as level 0 constructs. Such parts would include promoters, UTRs/ribosome binding sites (RBS), CDS, and transcription terminators. All these elements can be combined on the same plasmid, creating a functional expression unit. The assembled expression unit represent the next step in the Start-Stop Assembly, known as level 1 constructs. It is in the assembly of such vectors that start and stop codons are used as

fusion sites between CDS and upstream and downstream sequences to avoid additional scarring between them. The level 0 constructs used for the development of a level 1 construct contain five different fusion sites (α, β, γ, δ and ϵ), which all are three bp long as seen in Figure 1.4. The α and ϵ fusion sites are found at the beginning and end of the promoter and terminator, respectively. No typical sequence is found at these locations. Therefor these fusion sites are simply designed to contain different sequences from the rest of the fusion sites, as well as each other. The β fusion site contain the nucleotide sequence CCA. It is situated at the transcription start site between the promoter and RBS. The choice of sequence is based on an alignment of *E. coli* MG1655 transcription start site. No exact consensus sequence has been found in this region. Hence, the chosen sequence is based on nucleotide frequency per position, and reduction of similarity to the other fusion sites. γ and δ fusion sites are found at the junction between CDS and RBS or terminator, respectively. Therefore, γ is defined as ATG, the most abundant start codon, while the δ fusion sites consist of a suitable stop codon [51].

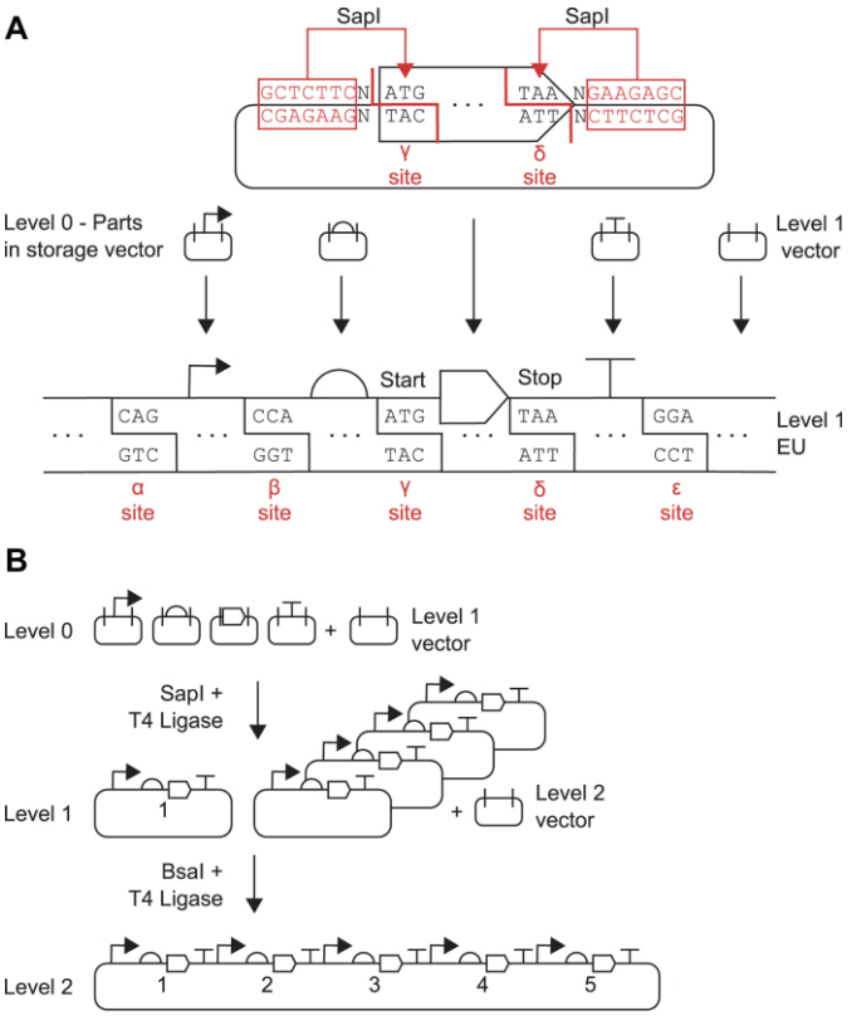


Figure 1.4: Design of the Start-Stop Assembly. A) Illustration of a level 1 expression unit (EU), generated by assembly of level 0 constructs. Examples of three bp fusions sites (α-ϵ) including the use of start (γ) and stop codons (δ) flanking the coding sequence. B) Scheme of steps from level 0 constructs til development of final level 2 constructs containing several gene cassettes. The figure is adapted with modifications from Taylor et al. [51].

In order to assemble these level 1 constructs, each 'part' (promoter, RBS, CDS and terminator) is first put on level 0 storage construct. The parts are synthesized with the fusion sites as well as inwards facing BsaI and SapI recognition sites flanking the part. SapI recognition sites need to be situated one bp away from the fusion site due to cleavage pattern. Additionally, the parts contain storage fusion sites and BsaI recognition sites, enabling BsaI endonucleases to be used with Golden gate Assembly to put each 'part' in Level 0 constructs. SapI can be used to facilitate assembly of promoter, RBS, CDS and terminator onto level 1 constructs, which will not contain scars between the parts. Onward, the system can be extended to develop level 2 constructs. Here, several expression units are gathered on the same constructs. The backbone plasmid used for level 1 construction then has to introduce new fusion sites and type IIS restriction endonuclease recognition sites, such as BsaI. The level 0, 1 and 2 acceptor vectors, providing the backbone which the parts are ligated to, need to contain the correct fusion sites as well as corresponding outward facing BsaI (level 0 and 2) or SapI (level 1) recognition sites. Level 0/2 acceptor vector contains the storage fusion sites, while level 1 acceptor contains α and ϵ fusion sites, enabling ligation to corresponding DNA parts of interest. Unlike level 1 fusion sites, level 0 and 2 fusion sites are located outside regions sensitive to scarring, and regular Golden Gate Assembly can be utilized [51].

1.7.4 Gibson Assembly

Another technique for assembly of DNA fragments is the Gibson assembly. The technique is based on the assembly of two or more overlapping DNA molecules, using 5'-exonucleases, DNA polymerase and DNA ligase. Exonuclease and DNA polymerase activity do not compete. Hence, the reaction can be performed in a single step, optimized at 50°C. As illustrated in Figure 1.5, 5'-exonucleases degrade the DNA strands from the overlapping 5'-end. The cleavage pattern creates single-stranded overhangs. As the different parts have overlapping sequences at the ends, these 3'-overhangs of different parts contain complementary sequences. The overhangs are ligated at the overlapping sequences, and strands are repaired using the DNA polymerase. As both ends of the DNA fragments can be ligated to each other, the Gibson assembly can result in circular DNA molecules [62].

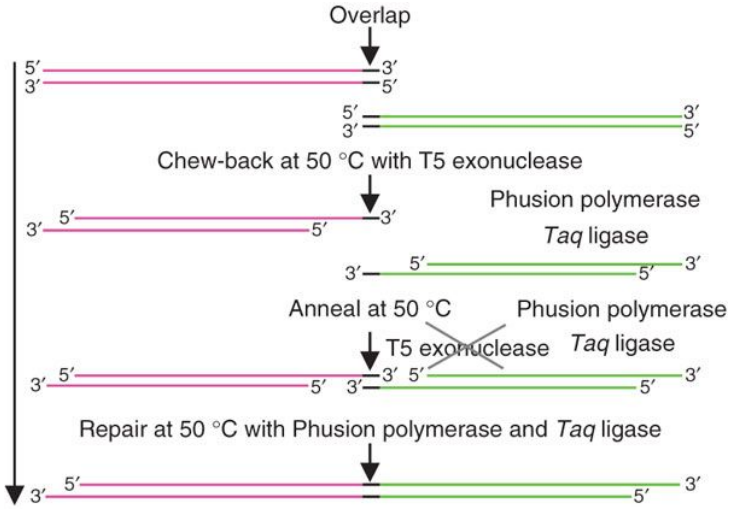


Figure 1.5: Single step assembly of overlapping double stranded DNA fragments using exonuclease, DNA polymerase and DNA ligase. Exonuclease degrade 5'-ends of the DNA molecule, the overhangs of the different molecules anneal and ligase and polymerase bind and repair joined parts. The figure has been adapted from Gibson et al [62].

1.7.5 DNA Amplification by Polymerase Chain Reaction

PCR is a common technique to amplify a high quantity of DNA from a single or few DNA molecules. In this chain reaction, one DNA molecule is used as a template to generate another copy. From there two new copies are made, then four and so onwards. An essential enzyme for the PCR reaction is the DNA polymerase. The polymerases are able to bind nucleotides, the building blocks of DNA. However, in order to function, the polymerases require small single-stranded DNA fragments, called primers, as well as a longer DNA molecule. The polymerase attaches nucleotides to the primer, while the larger DNA molecule function as a template. There are three main steps in a PCR reaction. The first step is the denaturation of double-stranded DNA molecules. Denaturation is carried out at a high temperature, usually 90-97 °C. Afterward, the primers anneal to complementary sequences in the now single-stranded DNA molecules. Annealing takes place at a lower temperature than denaturation, allowing for the hybridization of primers to the DNA strands. In the final step, polymerases cause extension at the ends of the primers, synthesizing new DNA molecules with the original DNA strand as its template. Using thermostable DNA polymerases eliminates the need for supplying new enzymes for each cycle. Therefore, the chain reaction can be repeated for several cycles, generating a high magnitude of DNA [63].

1.8 Bacterial Transformation

In order to obtain new genetic material, bacteria have developed distinct parasexual methods for horizontal gene transfer. These three methods are transformation, conjugation, and transduction [38]. Due to the relevance of this project, the focus will be on the explanation of transformation.

Transformation is the process where an organism takes up DNA from the environment. After uptake, exogenous DNA is incorporated into the organism's own genome, either as plasmids or in the chromosome. Uptake of exogenous DNA requires specialized pro-

teins that assemble into a DNA-uptake complex. The mechanism of how these organisms conduct the transformation varies. Not all bacteria are capable of performing the transformation. Wt-*E. coli* growing under normal conditions does not contain the necessary proteins and can, therefore, not perform the transformation. However, techniques based on chemical and physical treatment of *E. coli* have been developed to enable DNA to permeate the cell membrane. *E. coli* can therefor be made competent and bacterial transformation can be utilized to introduce foreign genetic material [38, 64].

1.9 Genetic Transformation in Microalgae Targeting cpDNA

Green algae such as *C. reinhardtii* have become increasingly popular within the field of plant biology and biotechnology. Therefore, several techniques have been developed for the genetic transformation of these organisms. Here, three techniques which could be utilized are presented. These techniques include electrophoration, glass bead-assisted and particle gun-mediated (biolistic) transformations. While the nuclear transformation of algae usually involves random integration events of the transgenic DNA, chloroplast transformation can be directed through homologous recombination [39, 47].

Biolistic delivery employs gold or tungsten microparticles coated with DNA, which is bombarded at the microalgae at high speed. This allows DNA to pass through the cell wall. Tungsten particles are more commonly used than gold as they are cheaper to produce. However, downsides include shape irregularity and increased degradation of transformant DNA compared to inert gold particles. The microalgae are spread as a monolayer on agar, liquid film or filters, and then bombarded with the DNA covered particles. Transformation efficacy has been reported to range from 0.1 - 100 transformants per million transfected microalgae. Furthermore, biolistic delivery has proven successful in chloroplast transformation, while glass bead and electrophoration are more commonly used for nuclear transformation [65].

The glass bead method is often used in combination with cell wall-deficient mutants of the microalgae. The method is widely used due to its simplicity and reproducibility. Agitations are used, employing a high number of microalgae and a high amount of exogenous DNA. Through vortexing the algae and DNA with glass beads, permeabilization is obtained, which enables the exogenous DNA to enter the microalgae. Even though the technique is most efficient with cell wall- deficient strains, it has proven successful with wt-*C. reinhardtii*. In these cases, the algae are pre-treated with autolysin to degrade the cell wall or cultivated in ammonium-deficient growth medium to limit the formation of a complete cell wall. The efficiency of transformation using the glass bead approach range from $\sim$ 1 - 100 transformants per million transfected *C. reinhardtii* [65].

Electrophoration utilizes an electric pulse to introduce DNA into the target cells. This technique also works better with cell wall deficient mutants. During the procedure, intensity and pulse duration are varied to control the efficiency of the transformations. It has been identified that the intensity of these pulses is inversely proportional to the survival rate of the microalgae. The duration of the pulse is proportional to the resistance of the medium. Therefore, the choice of medium is important as it influences pulse duration, which again affects cell survival rate. Before electrophoration, pre-incubation of the microalgae, together with exogenous DNA is carried out to allow contact between DNA and microalgae. Transformation efficiency has been reported as high as thousands of transformants per million cells when using cell wall-lacking *C. Reinhardtii* [65].

1.10 Experimental Project Goal

To summarize, the master thesis is part of a larger project. The project aims to identify orthogonal aaRS/tRNA and develop a system that allows for the introduction of ncAA into proteins synthesized within the chloroplast of the unicellular green microalga *C. reinhardtii*. To achieve this goal, foreign translation systems are to be introduced. The system utilize aaRS/tRNA pairs from several archaea as shown in Figure 1.6, which can incorporate different ncAA. The hypothesis to be tested is that these aaRS/tRNA are orthogonal to endogenous aaRS/tRNA of *C. Reinhardtii*. The reasoning for the hypothesis is that they have proven orthogonal to prokaryotes, to which the chloroplast shares several similarities with. To minimize tampering with endogenouse protein translation at undesired locations, codon reassignment of the opal and amber stop codons are employed to direct incorporation of these ncAA. Two sites for gene delivery were choose in the plastome to increase chances of success by homologous recombination. These are found within the *psbL* and *psbH* loci.

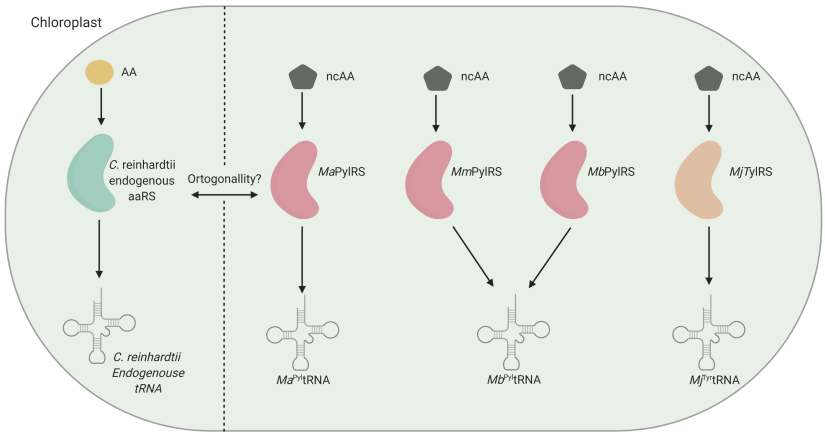


Figure 1.6: The goal of the entire project is to introduce pathways for incorporation of non-canonical amino acids (ncAA) during protein synthesis in the chloroplast of *C. reinhardtii*. To do so, PylRS and TyrRS are to be utilized for codon reassignment of opal and amber codon, respectively. For the codon reassignment to work, *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase (*Ma*PylRS), *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS), *Methanosarcina mazei* pyrrolysyl-tRNA synthetase (*Mm*PylRS), *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*TyrRS) and their cognate tRNA have to be orthogonal towards the endogenous translation system in the chloroplast of *C. reinhardtii* which utilize proteinogenic amino acids (AA).

In addition to orthogonality towards endogenous aaRS/tRNA, several of the chosen aaRS/tRNA pairs are orthogonal towards each other. Therefore, if the orthogonal system's implementation proves successful, the project could be extended to utilize two or more of the aaRS/tRNA pairs simultaneously. Hence, incorporation of two different ncAA into the same protein could be possible.

The goal of this thesis is to construct the genetic vectors necessary to transfer and introduce the aaRS/tRNA genes. Vectors containing two regions flanking the sites for homologous recombination, as well as four different gene cassettes (tRNA, aaRS, selectable marker and reporter), are developed. The goal is to use these vectors to introduce the

necessary genes for the translation system. The six different aaRS/tRNA pairs listed in Table 1.2 are used. Due to different properties, these aaRS are able to incorporate a variety of ncAA. Additionally, two different sites within the plastome are targeted for homologous recombination, both the *psbH* and *psbL* locus. Therefore, a total of twelve different vectors have to be prepared. Individual specifics for these are given in Figure 1.7. To develop vectors that are functionally scarless, the previously described Start-Stop Assembly methods are used when assembling the genes.

The thesis contains several sub-goals. The first milestone is to introduce necessary point mutations to develop described aaRR, tRNA and mVenus variants. Thereafter, it is the assembly of each part into functional scarless expression units, as well as gathering such expression units onto the same final level 2 construct.

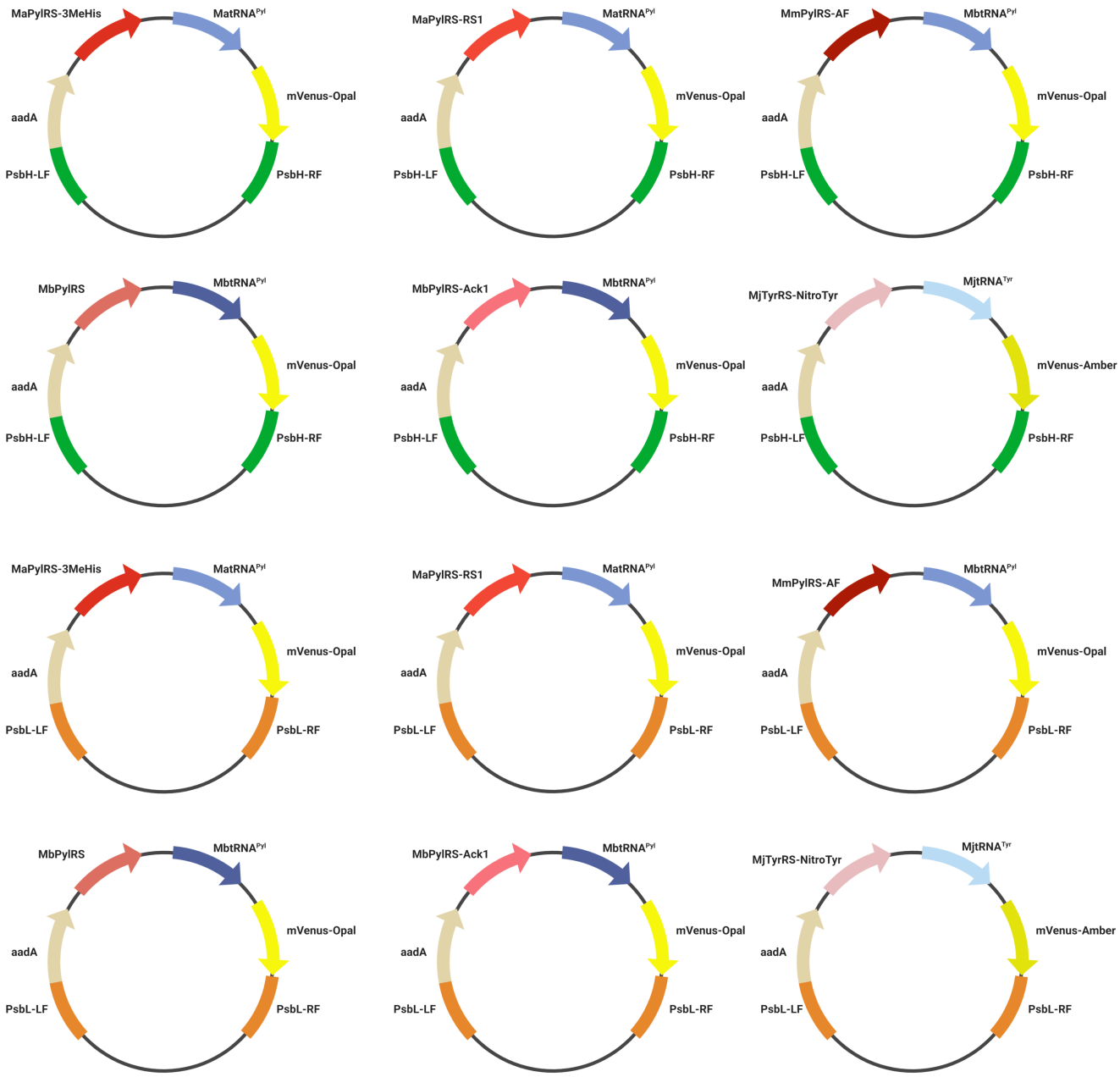


Figure 1.7: Overview of all level 2 constructs which were aimed to be assembled to implement orthogonal systems in *C. reinhardtii*. All vector contain *psbL* or *psbH* regions for homologous recombination, *aadA* for spectromycin resistance, *mVenus* fluorescent reporter gene. Additionally, each vector contain either a variant of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase (*Ma*PylRS), *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS), *Methanosarcina mazei* pyrrolysyl-tRNA synthetase (*Mm*PylRS) or *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*TyrRS) which are to perform the codon reassignment with their cognate tRNA.

1.10.1 System Design

When constructing the gene vectors, considerable focus was put into developing the functionally scarless vectors capable of delivering the transgenes into the cpDNA of *C. reinhardtii*. Here, a detailed description of the design for the system utilized for vector construction is given. The Start-Stop Assembly was utilized as earlier described with some alterations. In the reported Start-Stop Assembly promoter, RBS/5'-UTR, CDS and terminator are assembled as four different parts into one final expression unit. In this project, the promoter and RBS/5'UTR are already on the same part as they are amplified as a single sequence. Hence, only three different parts are assembled on level 1, and the β fusion site is not used.

Firstly, recognition sites for the restriction endonucleases SapI and BsaI and fusion sites are added as flanking regions for most promoters, CDS and 3'UTR used. As seen in Figure 1.8, this allows for the assembly of all parts on to level 0 plasmid, using BsaI in a Golden Gate Assembly. Each part is put on the same backbone, replacing a *tsPurple* gene previously present in the acceptor plasmid. *tsPurple* encode a nonfluorescent purple chromoprotein [66]. This allows for swift selection of correct assembly after cloning, as bacteria still containing plasmids with the *tsPurple* will appear purple.

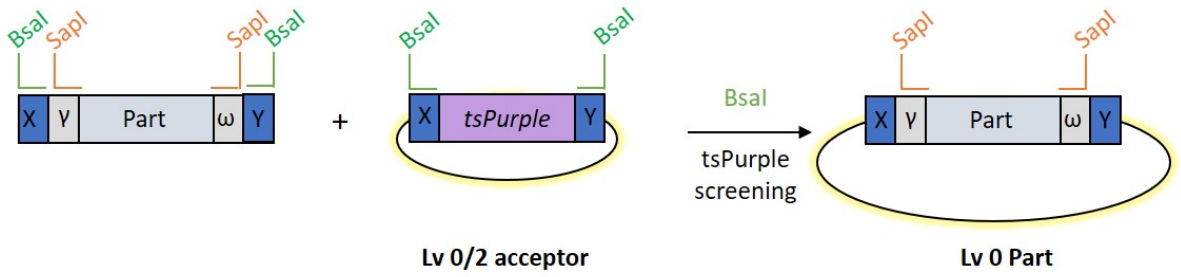


Figure 1.8: Assembly of level 0 vectors. Each part (promoter/5'UTR, coding sequence and 3'UTR) is put on plasmids using BsaI restriction enzymes. BsaI cleavage sites are symbolised as X and Y, and function as fusion sites for the level 0 assembly. SapI cleavage cites are exemplified with γ and ω.

The next step is the assembly of promoters/5'UTR, CDS and 3'UTR on level 1 constructs. The system design allows for restriction ligation procedure using SapI, which binds to the restriction sites on the level 0 vectors and cleaves them at the cut sites. Each SapI cleavage site is located at the beginning and end of the sequence of each part. The cleavage site at the 5' end of the promoter is designed to be a start codon. The cleavage of the 3' end of the CDS is also its start codon. The second cleavage site of the CDS is found at its stop codon. Additionally, the SapI cleavage site at the 3'UTR is the same type of stop codon. This way, a restriction ligation procedure with SapI is utilized to cut and ligate these three parts. The start and stop codons become the overhangs that are ligated, resulting in a promoter/5'UTR, CDS and 3'UTR found directly after each other. No additional scarring is added. The different genes are put on similar level 1 backbones, but with slightly different level 2 fusion sites. Figure 1.9 illustrates a scheme of how this assembly is carried out.

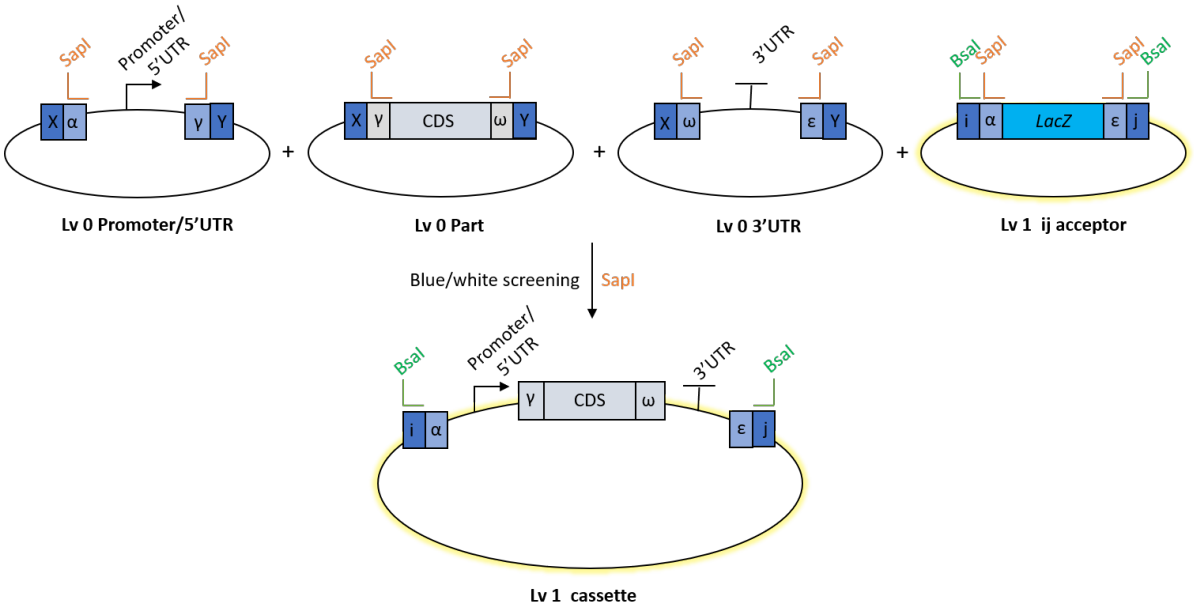


Figure 1.9: Assembly of a level 1 vector containing chloroplast promoter, 5'untranslated region (UTR), coding sequence (CDS) and 3'UTR. The assembly utilize SapI, three different level 0 vectors and a level 1 acceptor plasmid. i and j represent the different BsaI cleavage sites on level 1 acceptors. X and Y are scars from level 0 assembly. α, γ, ω and ϵ are SapI cleavage sites, functioning as fusion sites for level 1 assembly.

To allow for screening of bacterial colonies containing the correct plasmids after cloning, the level 1 acceptor plasmid initially contains a *lacZ* gene. *LacZ* encodes a β-galactosidase, an enzyme that catalyzes the hydrolysis of colorless 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal), where indole is cleaved of. Indole self-dimerize and creates an insoluble blue product. Due to this reaction, bacterial colonies containing β-galactosidase and growing on medium containing X-gal turn blue, while bacteria containing the correct insert will not [67].

The final plasmid, a level 2 vector, can be assembled as shown in Figure 1.10. As the level 1 acceptor vector introduces new BsaI recognition sites, the Golden Gate Assembly can once again be used to assemble the new plasmid. As seen in Figure 1.10, no more than two level 1 constructs contain the same fusion sites, and the left and right fusion sites of each plasmid are different. Regions with similar fusion sites will bind to each other, allowing for the insertion of DNA in the correct order and preventing self-ligation. Seven different fusion sites are involved in the assembly of the level 2 vector, denoted: X, A, B, C, D, Z and Y. These fusion sites contain the BsaI cleavage sites, creating overhangs during restriction. Again, the acceptor vector for level 2 constructs contain *tsPurple*, allowing for quick screening of correct colonies. Following this design, no scars are created between the promoter/5'UTR, CDS and terminator. However, scars are created between the different expression units present in the final level 2 plasmid. Nonetheless, they are situated outside sensitive regions, and therefor not of concern [51].

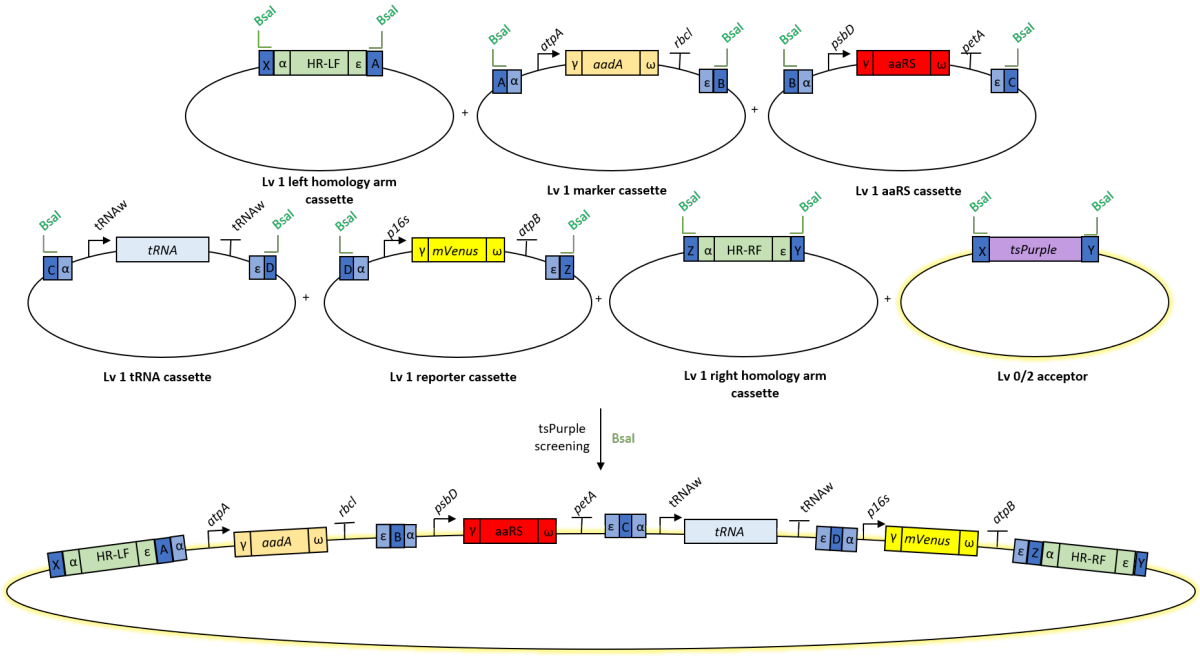


Figure 1.10: Assembly of a level 2 vector using seven different level 1 vectors. Left homology flank (HR-LF), selective marker *aadA*, aminoacyl tRNA synthetase (aaRS), tRNA, reporter gene *mVenus*, and right homology flank (HR-RF) are assembled on the same backbone. BsaI restriction endonucleases and ligation enzymes. Directionality during assembly is conferred by BsaI cleavage sites A, B, C, D, X, Y and Z. α and ϵ symbolise level 1 fusion sites.

2 Materials and Methods

Within this section a detailed description of how the vectors required for codon reassignment in the chloroplast of *C. reinhardtii* constructed. Firstly a description of the microorganisms used and their growth conditions are given. The second part includes a general description of the different techniques used for plasmid construction and cloning. The final part elaborates on how these techniques were utilized to develop the different genetic vectors.

2.1 *E. coli* DH5α

All cloning was carried out using *E. coli* DH5α, a strain engineered for optimized transformation efficacy [68]. Aliquots (200 µl) of the bacterial cultur were stored at -80°C until use. *E. coli* was incubated in lysogeny broth (LB) medium. For the growth of colonies on plates, LB medium was supplied with agar. Incubation of LB agar plates was carried out at 37°C in incubators, while incubation of bacteria in liquid LB was carried out at 37°C and continuously shook at 225 rpm. The content of LB medium is given in Table 2.1.

Table 2.1: Composition of LB medium used for *E. coli* cultivation. Dilution were carried out using distilled water.

Composition	Concentration [g/L]	Comment
Bacterial peptone	10	
Yeast extract	5	
Sodium chloride	5	
Bacterial Agar	3.75	For solid LB only

2.2 C. reinhardtii

For this project, wild-type (wt) *C. reinhardtii* was used for chloroplast DNA (cpDNA) extraction. The algae were cultivated in liquid Tris-Acetate-Phosphate (TAP) medium. Content of TAP medium is given in Table 2.2 Incubation was performed at 25°C in the presence of a light source.

Table 2.2: Composition of 1 L Tris-Acetate-Phosphate (TAP) medium used for *C. reinhardtii* cultivation. Content of phosphate buffer II and solution A is listed in Appendix A.

Component	Volume [ml]	Comment
1 M TRIS base	20	
Phosphate Buffer II	1	
Solution A	10	
Hutner's trace element	1	
Glacial acetic acid	~ 1	Added to pH = 7
Distilled water	967	

2.3 General Techniques

All laboratory work with bacteria and algae were carried out using sterilized equipment and media. Sterilization was achieved thorough autoclavation (20 min, 120°C). All liquid media were stored at room temperature. Agar plates were stored at 4°C. Enzymes, primers and plasmids were stored at -20°C, while bacterial cultures were stored at -80°C. Eppendorf® Microcentrifuge 5424 was used for centrifugation of all small samples (≤ 1.5 mL), while Eppendorf® Centrifuge 5430 R was used for larger samples (>1.5 mL). DNA concentrations were measured using NanoDrop™ Microvolume Spectrophotometer set to measure dsDNA.

2.4 Crude Yeast Genomic DNA Exctraction from Wild-type *C. reinhardtii*

cpDNA was extracted from wt-*C. reinhardtii* using crude yeast genomic DNA extraction (YGE). *C. reinhardtii* was grown in TAP medium until OD730 reached 1. A cell pellet was obtained through centrifugation (600 rcf, 5 min, 25°C) of the algae culture (150 µl). The pellet was re-suspended in lithium acetate (0.2 M, 100 µl) followed by incubation (70°C, 5 min). Ethanol (96%, 300 µl) was added and the sample was mixed using a vortex mixer, then centrifugated (13000 rpm, 3 min) to re-obtain cell pellet. The pellet was washed with ethanol (70%, 100 µl), then dissolved in distilled water (100 µl). Lastly, the supernatant containing cpDNA was obtained by centrifugation (13000 rpm, 15 s).

2.5 Heat-shock Bacterial Transformation

Bacterial transformation using competent *E. coli DH5α* cells were carried out in order to clone DNA fragments of interest. The procedure was initiated by thawing of competent *E. coli DH5α* cells (200 µl) for 10 minutes on ice. Purified plasmid (0.5 µl) or DNA assembly reaction mixture (10 µl) was added. After 30 minutes of incubation on ice, heat-shocked was carried out (42 °C, 45 sec) in a water bath, followed by cooling on ice again (1 min). Afterward, LB (1 mL) was added and the culture was again incubated (37 °C, 1 h). Lastly, cells were plated out (∼ 300 µl) on LB agar plates containing appropriate antibiotics, and incubated (37 °C, overnight).

2.6 Polymerase Chain Reaction for DNA Amplification

Polymerase chain reaction (PCR) was used for both DNA amplification and site-specific point mutations. Two different protocols for PCR were used during the experimental work, referred to as Q5 PCR and colony PCR. Colony PCR was used to screen bacterial colonies for verification of construct assembly. Q5 PCR was used for mutagenesis and whenever amplified DNA was to be further used. All PCR products were analyzed through agarose gel electrophoration. Different DNA ladders were used depending on PCR product size. These included 1 kb GeneRuler DNA ladder (Thermo Scientific), 1 kb DNA ladder (New England Biolabs) and 100 bp plus GeneRuler DNA ladder (Thermo Scientific).

2.6.1 Colony PCR

During colony PCR, bacterial colonies were picked from plates and diluted in distilled water (10 µl) using distilled water. From this dilution 1 µl was used for the PCR reaction. The reaction mixture was prepared as listen in Table 2.3, using Taq DNA polymerase and 10X Standard Taq Reaction Buffer. Total volume was adjusted to 10 µl using distilled water. The annealing temperature for all primers was found using New England Biolab (NEB) T_M calculator [69]. PCR conditions were as described in Table 2.4, with an extension phase set to 1 min/kb of the PCR product.

Table 2.3: Content of PCR mixtures used for Q5 and colony PCR, which were diluted with distilled water to 25 µl and 10 µl, respectively. Taq DNA polymerase and Standard Taq Buffer was used for colony PCR, while Q5 DNA polymerase and Q5 Reaction Buffer was used for Q5 PCR. *DNA template for colony PCR was prepared by diluting each bacterial colony in 10 µl distilled water.

Reagent	Q5 PCR	Colony PCR
DNA template	<10 ng	1 µl*
10 µl Forward primer	1.25 µl	0.2 µl
10 µl Reverse primer	1.25 µl	0.2 µl
10 mM dNTP	0.5 µl	0.2 µl
DNA polymerase	0.25 µl	0.05 µl
10X reaction buffer	5 µl	1 µl

2.6.2 Q5 PCR

Q5 PCR was performed using the same components in Table 2.3 and conditions in Table 2.4. Total volume was adjusted to 25 µl using distilled water. A touch-down procedure was utilized whenever the standard Q5 PCR proved unsuccessful. Here, two different cycles of denaturation, annealing and extension, were used. The temperature of the annealing step was gradually increased per round in the first set of cycles. The second set of cycles ran at the constant annealing temperature. After verification by Gel electrophoresis, the remaining product was cleaned using *QIAquick PCR purification kit*.

Table 2.4: Set-up for Q5 and colony PCR in a thermocycler. Temperature of annealing step (T) depend on the individual primers. Denaturation, annealing and extension were performed over several cycles. Extension time depend on PCR product size.

Step	Q5 PCR			Colony PCR		
	Time	T [°C]	Cycles	Time	T [°C]	Cycles
Initial denaturation	60 s	98		10 min	95	
Denaturation	15 s	98		1 s	95	
Annealing	15 s		x35	15 s		x25
Extension	30 s/kb	72		60 s/kb	70	
Final extension	5 min	72		5 min	70	

2.7 Agarose Gel Electrophoresis

PCR products and other DNA fragments were separated and analyzed through gel electrophoresis. Colony PCR product (10 µl), Q5 PCR product (1 µl) or DNA restriction product (10 µl) were mixed with 6X Purple Loading Dye (2 µl). Q5 PCR products were further diluted using distilled water (9 µl). Samples and appropriate standard DNA ladder were added to 0.8% agarose gels, submerged in 1x Tris-Acetate-EDTA (TAE) buffer. The electrophoresis ran at 90-120V for 30-60 min, depending on gel and DNA fragment size. Following gel electrophoresis, the gels were scanned using Molecular Imager® ChemiDoc™ XRS+ imaging system. The composition of the agarose gel is given in Table 2.5.

Table 2.5: Content of 0.8% agarose Tris-Acetate-EDTA (TAE) gel stock solution used for gel electrophoresis.

Component	Amount
Agarose	3.2 g
GelRed Nucleic Acid Stain	20 µl
1x TAE buffer	400 ml

2.8 DpnI Digestion of PCR Templates

Digestion using DpnI enzymes were carried out after PCR amplification of bacterial DNA. The digestion was carried out to remove methylated template DNA. The reaction mixture contained DpnI (1 µl), CutSmart Buffer (5 µl) and the PCR product. The sample was diluted to a total volume of 50 µl using distilled water and heated at 37°C overnight.

2.9 Gibson Assembly

Gibson assembly was used to assemble DNA fragments into circular plasmids. For each reaction, Gibson reaction mixture (10 µl) was mixed with each of the inserts in equimolar amounts to a total volume no higher than 15 µl. The volumes were adjusted to 20 µl using distilled water and the reaction was carried out in a heating block (50 °C, 1 hour). Details on the reaction mixture are given in Appendix A.

2.10 Golden Gate Assembly for Level 0/2 Construction

Golden Gate DNA assembly was utilized for DNA assembly of level 0 and 2 constructs. High-fidelity BsaI restriction enzyme (BsaI-HFv2), hereby referred to as BsaI, optimized for this type of reaction was used to cleave the plasmids and DNA fragments. Plasmid pSS191_lv0/2_tsPurple containing an ampicillin resistance gene was used as the acceptor vector in both levels, where *tsPurple* was replaced by inserts of interest.

Specifics of the reaction mixtures are given in Table 2.6. For assembly of level 2 construct, the reaction mixture contained pSS191_lv0/2_tsPurple (1 nM) and up to six donor plasmids (3 nM). Details on plasmid reaction volumes and calculations are given in Appendix B. The reaction was carried out over two steps repeated for 30 cycles in a thermocycler. In each cycle, the sample was heated to 37 °C for five minutes then cooled to 16 °C for another five minutes. Following DNA assembly, half of the reaction mixture (10 µl) was used directly for transformation into *E. coli* and then plated out (300 µl) on LB agar plates containing ampicillin (100 mg/ml). Additional screening was carried out based on coloration of purple *tsPurple+* colonies and white *tsPurple-* colonies.

Table 2.6: Reaction mixture used for assembly of level 0 and 2 constructs.

Component	Level 0	Level 2
BsaI-HFv2	0.75 µl	1 µl
T4 ligase	1.25 µl	1 µl
10X T4 buffer	2 µl	2 µl
pSS191_lv0/2_tsPurple	75 ng	1 nM
PCR product/ donor plasmid	75 ng	3 nM

2.11 Start-Stop Assembly

Start-Stop Assembly was used to generate level 1 constructs containing both promoter/5'untranslated region (UTR), coding sequence (CDS) and 3'UTR. Three level 1 acceptor plasmids with slightly different fusion sites (B, C, D and Z) were utilized as backbones for the different expression units. This to confer directionality when assembling all the expression units on level 2 constructs. Which acceptor plasmid to be used is later specified per assembly. These vectors initially contain *LacZ*, which is replaced by the inserts.

During Start-Stop Assembly, level 1 acceptor plasmid (1 nM) and donor plasmids (3 nM) were added to a reaction mix containing SapI enzymes, T4 ligase and 10X T4 buffer. The amounts of each component are given in Table 2.7 and the total volume of each sample was adjusted to 20 µl. A detailed description for calculation of plasmid volumes are given in Appendix B. The reaction was carried out at 37 °C for five minutes, then decreased to 16 °C for another five minutes. This cycle was repeated 30 times. Following

DNA assembly, half of the reaction mixture was used directly for transformation and plated out on LB agar plates containing tetracyclin (12 µg/ml) and Blue/white select™ (40 µg/ml 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal), 40 µg/ml isopropyl β- d-1-thiogalactopyranoside). White *Lac- E. coli* colonies were picked for analysis by colony PCR, while blue *Lac- E. coli* colonies were discarded.

Table 2.7: Reaction mixture used for Start-Stop Assembly of level 1 constructs. volume of level 0 constructs containing backbone, promoter/5'untranslated region (UTR), coding sequence and 3'UTR depend on plasmid size and concentration. A detailed description of calculations are given in Appendix B.

Component	Amount
SapI	1 µl
T4 ligase	1 µl
10X T4 ligase buffer	2 µl
Donor plasmids	3 nM
Acceptor plasmid	1 nM

2.12 PNK Ligation

Polynucleotide kinase (PNK) ligation was utilized to ligate DNA strands containing phosphate groups at 5' ends. Purified PCR products were mixed with T4 PNK, T4 ligase and T4 ligase buffer as described in Table 2.8. The solution was heated (37°C, 30 min), followed by heat-inactivation (65°C, 20 min) of enzymes prior to the transformation in *E. coli*.

Table 2.8: Composition of reaction mixture for PNK ligation

Component	Volume [µl]
T4 PNK	1
T4 ligase	1
10X T4 ligase buffer	1
PCR product	7

2.13 Restriction Check to Verify Correct Inserts

Restriction checks were carried out to verify that the constructs contained correct restriction sites and inserts of the correct size. Plasmids of interest were mixed with BsaI and CutSmart Buffer. Each degradation was carried out (>3h, 37°C), then analyzed by gel electrophoration. Content of reaction mixture is given in Table 2.9.

Table 2.9: Reaction mixture for restriction check of plasmids. Solutions were diluted using distilled water to a total volume of 20 µl

Component	Amount
Plasmid	200 ng
BsaI	1 µl
CutSmart Buffer	2 µl

2.14 DNA Sequencing

All sequencing was performed by Eurofins Genomics. Samples were sent containing the plasmid (400 ng) and an appropriate primer (10 µl), 2.5 µl) diluted in distilled water to a total volume of 10 µl. Results were aligned with expected sequence using the software Benchling.

2.15 Stock Solutions

Whenever a construct was assembled and cloned, stock solutions were generated. *E. coli* transformed with the construct was incubated in LB (5 ml) supplied with the appropriate antibiotic. The cell cultures (1 ml) were centrifuged (2 min, 8000 rpm) and the pellets were re-suspended in 20% glycerol diluted in LB. Stocks were stored at 80°C.

2.16 Amplification of Flanks for Homologous Recombination within *psbL*

DNA regions flanking both sides of the *psbL* insertion site were attempt amplified from *C.reinhardtii's* cpDNA using PCR. Several attempts to amplify the left flanking region of psbL were conducted. Attempts involved previously described Q5 PCR, both with and without the addition of dimethyl sulfoxide to prevent secondary structure formation of the DNA template. Additionally, PCR with touch-down procedure during annealing step was conducted to improve chances of primer annealing. Further details on the PCR reactions are given in Table 2.10

Table 2.10: PCR conditions for amplification of the right (HR-RF) and left (HR-LF) flanks for homologous recombination in the *psbL* locus. PCR of *PsbL* HR-LF was performed at different conditions.

Product	Primers	T [°C]	Comment
psbL HR-RF	556 & 557	61	
psbL HR-LF	554 & 555	61	
psbL HR-LF	554 & 555	52→ 61	Touch-down procedure

2.17 Assembly of aminoacyl tRNA Synthetases on pUC8 Backbone

From here on, PCR referre to Q5 PCR unless otherwise stated. CDS with flanking SapI recognition sites and BioBrick prefix and suffix of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase-3MeHis (*Ma*PylRS-3MeHis), *Methanosarcina mazei* pyrrolysyl-tRNA synthetase-AF (*Mm*PylRS-AF), *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS) and *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase-NitroTyr (*Mj*TyrRS-NitroTyr) were provided by PhD candidate M. Fages-Laurtad. These CDS, as well as pUC8 plasmids, were PCR amplified with BioBrick primers. DNA sequence of all primers is given in Appendix C. Inserts and pUC8 backbone were assembled utilizing Gibson assembly. After transformation and cloning, colony PCR was performed with two different sets of primers. All PCR conditions are given in Table 2.11.

Table 2.11: Primers and annealing temperatures (T) used for amplification of each aminoacyl tRNA synthetase (aaRS) with biobrick prefix and suffix (BB) (1), colony PCR of assembled plasmids containing the aaRS and pUC8 backbone (pUC8_aaRS)(2, 3) and construction of pUC8_MaPylRS-RS1 (4, 5).

Number	PCR	Template	Primers	T [°C]
1	Q5 PCR	aaRS	BB1 & BB2	68
2	Colony PCR	pUC8_aaRS	BB1 & BB2	58
3	Colony PCR	pUC8_aaRS	384 & 427	47
4	Q5 PCR	pUC8_MaPylRS-3MeHis	513 & 516	58
5	Q5 PCR	pUC8_MaPylRS-3MeHis	514 & 515	58

pUC8_MaPylRS-RS1 (Y126M, M129G and V168T) was generated from pUC8_MaPylRS-3MeHis as show in Figure 2.1. pUC8_MaPylRS-3MeHis contain *Ma*PylRS-3MeHis CDS on pUC8 backbone. Two PCR reactions were carried out to introduce the point mutations, followed by DpnI digestion, cleaning and Gibson assembly. After transformation, identification of correct assembly was carried out through colony PCR and sequencing. PCR conditions are given in Table 2.11. Bacterial colonies with positive colony PCR results were incubated in LB (5 ml) supplied with ampicillin (100 µg/mL) overnight. Plasmids were isolated using *QIAprep Spin Miniprep Kit*.

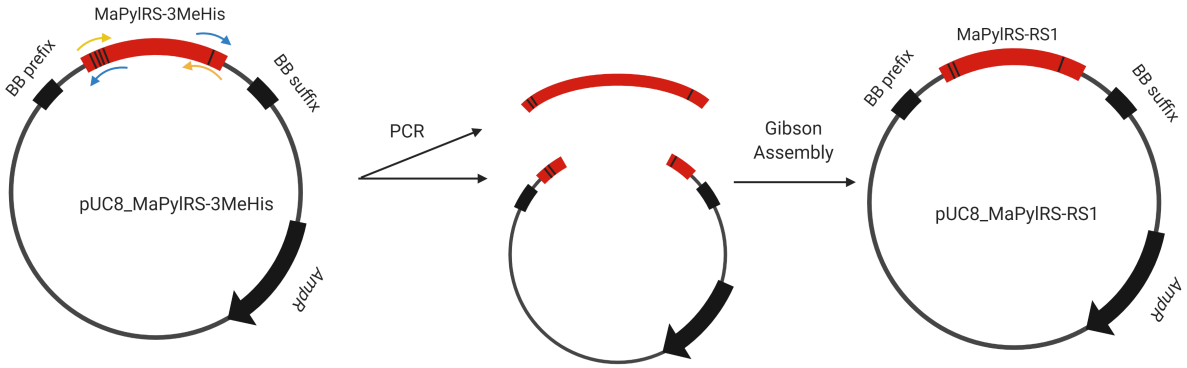


Figure 2.1: Scheme for construction of pUC8_MaPylRS-RS1 from pUC8_MaPylRS-3MeHis. Point mutations were introduced using PCR and DNA fragments were assembled using Gibson assembly. Primer used together are illustrated as blue and yellow arrows, respectively. Point mutations are illustrated with black lines.

2.18 Assembly of Level 0 Storage Constructs

Here, a detailed description of how all level 0 vectors were constructed is given. First, a general description common for all level 0 constructs is provided, followed by specific details for the individual constructs. A list of plasmids provided directly by PhD candidate M. Fages-Laurtad is given in Table 2.12. These were used during plasmid construction on all levels.

Table 2.12: List of all constructs provided by PhD candidate M. Fages-Laurtad. Each level 1 construct is listed with fusion sites (BsaI cleavage sites) conferring directionallity during level 2 construct assembly.

Construct level	Name	Comment
Level 0	pSS124.Lv0_p16S-psaA5'	Contain p16S promoter and psaA 5'-UTR.
	pSS151_Lv0_CDS.mVenus	Contain coding sequence of *mVenus*.
	pSS159_pLv0_petA.T	Contain *petA* 3'UTR.
	pSS164_Lv0_ter.atpB	Contain *atpB* 3'UTR.
	pSS191.Lv0/2_tsPurple	Acceptor plasmid containing *tsPurple*.
	pHT201.Lv0.tRNAW	Contain a $tRNA^{trp}$ gene.
Level 1	pSS104_pLv1_BC	Acceptor plasmid containing a *LacZ* gene with B and C fusion sites.
	pSS106_pLv1_CD	Acceptor plasmid containing a *LacZ* gene, C and D fusion sites.
	pSS104_pLv1_DZ	Acceptor plasmid containing a *LacZ* gene, D and Z fusion sites.
	pHT146_pLv1_XA-psbH.LF	Contain the left flank for homologouse recombination with *psbH*locus as well as X and F fusion sites.
	pHT150_pLv1_ZY-psbH.RF	Contain the right flank for homologouse recombination with *psbH* locus as well as Z and Y fusion sites.
	pLv1_AB:atpA_aadA_rbcl	Contain a *atpA* promoter/5'UTR, *aadA* coding sequence, *rbcl* 3'UTR as well as A and B fusion sites.

Promoters/5'UTR and CDS were all amplified using Q5 PCR from provided plasmids or from *C. reinhardtii's* cpDNA. During PCR amplification, BsaI and SapI recognitions sites, as well as specifically tailored fusion sites (restriction enzyme cleavage sites), were added as flanking regions to each part. Each level 0 construct was assembled by Golden Gate Assembly, using the PCR product of interest and pSS191.Lv0/2_tsPurple as a backbone.

After DNA assembly, the constructs were cloned in *E. coli*. The DNA was introduced into the bacteria through transformation and incubated on LB agar plates supplied with ampicillin (100 µg/ml). With the exception of level 0 tRNA, three steps of screening were conducted to identify bacterial colonies containing correct plasmids. Firstly, purple *tsPurple+* colonies were discarded. Secondly, a colony PCR was performed for white *tsPurple-* colonies. Primer 384 and 427, which anneal to the level 0/2 backbone, were used for all colony PCR of level 0 constructs. The annealing temperature was set to 47°C. Colonies with positive results were inoculated in liquid LB (5 mL) with ampicillin (100 µg/ml) and incubated overnight. Plasmids were isolated from the remaining bacterial culture using *QIAprep Spin Miniprep Kit* and then sequenced.

2.18.1 Assembly of Level 0 aaRS

Level 0 constructs containing *Mb*PylRS (p001_MbPylRS) and *Mj*TyrRS-NitroTyr (p003_TyrRS-NitroTyr) were assembled, due to lacking results of assembly with pUC8. First, the CDS where amplified with new primers, as listed in Table 2.13, introducing necessary BsaI recognition site. Then, these parts were assembled on level 0 constructs.

Table 2.13: Primers and annealing temperatures (T) used during Q5 PCR to introduce BsaI and SapI recognition sites.

Product	Primers	T [°C]
*Mb*PylRS	509 & 510	58
*Mj*TyrRS-NitroTyr	511 & 512	58

A level 0 containing *Mb*PylRS-Ack1 (p002_MbPylRS-Ack1) was generated by PCR mediated site-directed mutagenesis of p001_MbPylRS. A scheme of the process is given in Figure 2.2. Firstly, PCR with two sets of primers was used to introduce mutation L266V, L270I and L274A. After DpnI digestion, the parts were re-assembled using Gibson Assembly. The intermediate was cloned and analyzed by colony PCR. The second round of

PCR was performed to introduce the final point-mutations (D76G and C313T). Gibson Assembly was again performed to assemble the plasmid. PCR conditions are given in Table 2.14.

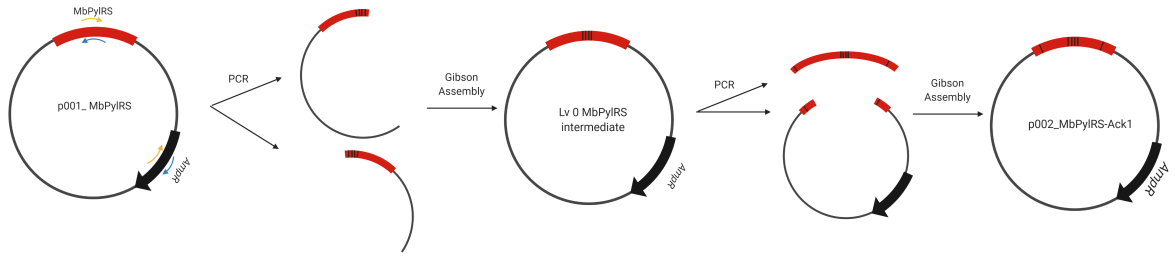


Figure 2.2: Scheme for construction of p002_MbPylRS-Ack1 from p001_MbPylRS. Point mutations are introduced using PCR and DNA fragments are assembled using Gibson assembly. Primer used together are illustrated as blue and yellow arrows, respectively.

Table 2.14: PCR set-up for construction of p002_MbPylRS-Ack1, with primers and template used as well as the annealing temperature (T) for the PCR. Reaction was carried out over two steps, generating a level 0 *Mb*PylRS intermediate after the first round of PCR and Gibson assembly.

Product	Primers	Template	T [°C]
Lv 0 *Mb*PylRS-intermediate	519 & 17	p001_MbPylRS	58
	520 & 18	p001_MbPylRS	58
p002_MbPylRS-Ack1	517 & 522	Lv 0 *Mb*PylRS-intermediate	67
	518 & 521	Lv 0 *Mb*PylRS-intermediate	67

2.18.2 Assembly of Level 0 mVenus

Variants of the *mVenus* CDS were prepared to use as the reporter gene for the finished level 2 constructs. These mVenus variants include one variant with an opal codon at position 151 (*mVenusOpal*), another with an amber codon at position 3 (*mVenusAmber*), and a third variant with both codon inserts (*mVenusAO*). Furthermore, a 5' flanking sequence of six consecutive GTG codons was added to each sequence, creating a histidine tag at the synthesized proteins.

These sequences were generated over three steps. First, PCR was carried out for site-directed mutagenesis, introducing the stop codons at the correct locations. Mutagenesis was performed over two rounds of PCR, using pSS151_Lv0_CDS.mVenus as a template. The PCR products were DpnI digested, purified and then combined using a Gibson assembly. The second step was to introduce 5' his tag, achieved through PCR using primers introducing the necessary DNA sequence. After DpnI digestion and PCR purification, the third round of PCR was conducted in order to re-introduce BsaI and SapI recognition sites lost after addition of the histidine tag. Finally, the inserts were re-assembled on level 0 constructs (p007_mVenusAmber, p008_mVenusOpal and p009_mVenusAO) using Golden Gate Assembly. Information about each PCR is given in Table 2.15.

Table 2.15: PCR set-up used to generate level 0 constructs p007_mVenusAmber, p008_mVenusOpal and p009_mVenusAO mutants with his-tags. Set-up is given with DNA template, primers and annealing temperature (T).

Product	Primers	Template	T [°C]	Comment
mVenusAmber	505 + 17	pSS151_Lv0_CDS.mVenus	58	To introduce amber codon, SapI and BsaI recognition sites.
	506 + 18	pSS151_Lv0_CDS.mVenus	58	
	581 + 582	p007_mVenusAmber	62	To introduce 5' his-tag
	581 + 176	p007_mVenusAmber	61	To re-introduce SapI and BsaI recognition sites at 5' end
mVenusOpal	507 + 17	pSS151_Lv0_CDS.mVenus	58	To introduce opal codon, SapI and BsaI recognition sites
	508 + 18	pSS151_Lv0_CDS.mVenus	58	
	581 + 582	p008_mVenusOpal	62	To introduce 5' his-tag
	581 + 176	p008_mVenusOpal	61	To re-introduce SapI and BsaI recognition site at 5' end.
mVenusAO	505 + 508	pSS151_Lv0_CDS.mVenus	58	To introduce Amber codon, opal codon, SapI and BsaI recognition sites
	506 + 507	pSS151_Lv0_CDS.mVenus	58	
	581 + 582	p009_mVenusAO	62	To introduce 5' his-tag
	581 + 176	p009_mVenusAO	61	To re-introduce SapI and BsaI recognition sites at 5' end.

2.18.3 Assembly of Level 0 Cloroplast Promoters

C. Reinhardtii cpDNA was extracted using the earlier explained YGE protocol. Afterwards, chloroplast promoters and 5'UTR were amplified by PCR, with primers introducing flanking SapI and BsaI recognition sites. Promoter/5'UTR for *psbD* and *rbcl* were amplified by Q5 PCR protocol. The *psbB* promoter75'UTR was amplified by touch-down PCR. Further details on the PCR reactions are given in Table 2.16. After PCR amplifications, level 0 constructs were assembled using the Golden Gate Assembly containing promoter/'5UTR of *psbB* (p004_psbB), *psbD* (p005_psbD) and *rbcl* (p006_rbcl).

Table 2.16: PCR set-up for isolation and amplification of chloroplast promoters/5'-untranslated region(UTR) from chloroplast DNA. T indicates annealing temperature of each reaction. *psbB* promoter/5'UTR was amplified with a touch-down procedure.

Promoter/5'UTR	Primers	T [°C]
psbD	546 & 547	60
rbcl	535 & 536	60
psbB	537 & 538	54-60 °C

2.18.4 Assembly of Level 0 tRNA Constructs

Level 0 constructs containing the CDS of *Mb*tRNA$^{Pyl}_{UCA}$ (p011_MbtRNA) were synthesised by PCR, using pHT201.Lv0.tRNAW as template. A simplified overview of the process is given in Figure 2.3. PCR primers 523 and 524 contained the the *Mb*tRNA$^{Pyl}_{UCA}$ DNA sequence, replacing CDS of $tRNA^{trp}$. Annealing temperature was set to 61°C. The PCR products were DpnI digested overnight, purified and self-ligated through PNK ligation. Ligation product (10 µl) was used directly for transformation and cloning in *E. coli*. Attempts to generate *Ma*tRNA$^{Pyl}_{UCA}$ (primer 525 & 526) and*Mj*tRNA$^{Tyr}_{CUA}$ (primer 527 & 528), were carried out in the same manner and the same PCR conditions.

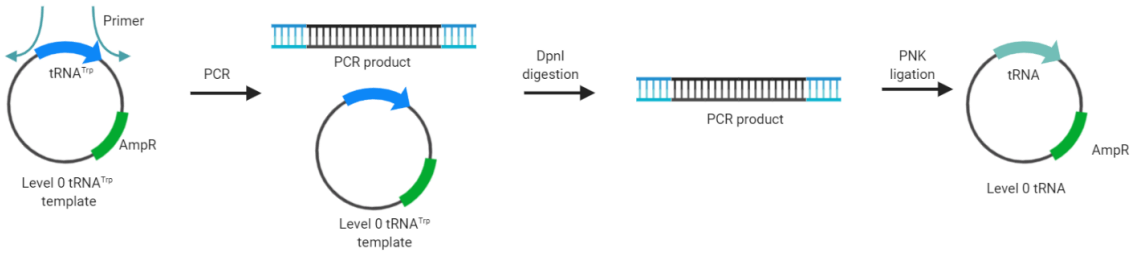


Figure 2.3: Scheme for construction of level 0 constructs containing *Ma*tRNA$^{Pyl}_{UCA}$, *Mb*tRNA$^{Pyl}_{UCA}$ and *Mj*tRNA$^{Tyr}_{CUA}$ from a level 0 template construct containing tRNA^{Trp} (pHT201.Lv0.tRNAW). Each vector carry a ampicllin resistance gene (*AmpR*) as part of the backbone.

2.19 Assembly of Expression Units on Level 1 Constructs

Level 1 constructs containing the different CDS, chloroplast promoter/5'UTR and 3'-UTR, were all constructed using the Start-Stop Assembly. Each vector containing a different type of CDS (aaRS, tRNA, reporter or selective marker) was put on backbones with slightly different fusion sites (A, B, C, D or Z) to confer correct assembly during level 2 construction. All level 1 backbones contained a gene conferring tetracycline resistance, enabling positive selection in the presence of the antibiotic. A list on all level 1 constructs assembled directly by Start-Stop Assembly, and which level 0/pUC8 donor plasmids and level 1 acceptor plasmids were used, are listed in Table 2.17.

Table 2.17: List of all level 1 vectors constructed using Start-Stop Assembly, and which plasmids were used for the assembly of promoter/5'untranslated region (UTR), coding sequence (CDS), 3'UTR and backbone.

Level 1 construct	Promoter/5'-UTR	CDS	3'-UTR	Backbone
p101BC:PsbD_MaPylRS-3MeHis_petA	p005_psbD	pUC8_MaPylRS-3MeHis	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p102BC:PsbD_MaPylRS-RS1_petA	p005_psbD	pUC8_MaPylRS-RS1	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p103BC:PsbD_MbPylRS_petA	p005_psbD	p001_MbPylRS	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p104BC:PsbD_MbPylRS-Ack1_petA	p005_psbD	p002_MbPylRS-Ack1	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p105BC:PsbD_MmPylRS-AF_petA	p005_psbD	MmPylRS-AF (PCR product)	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p106BC:PsbD_MjTyrRS-NitroTyr_petA	p005_psbD	p003_MjTyrRS-NitroTyr	pSS159_pLv0_petA.T	pSS104_pLv1_BC
p107CD:MbtRNA		p010_MbtRNA		pSS106_pLv1_CD
p110DZ:p16s_mVenusAmber_atpB	pSS124.Lv0_p16S-psaA5'	p007_mVenusAmber	pSS164_Lv0_ter.atpB	pSS104_pLv1_DZ
p111DZ:p16s_mVenusOpal_atpB	pSS124.Lv0_p16S-psaA5'	p008_mVenusOpal	pSS164_Lv0_ter.atpB	pSS104_pLv1_DZ
p112DZ:p16s_mVenus-AO_atpB	pSS124.Lv0_p16S-psaA5'	p009_mVenusAO	pSS164_Lv0_ter.atpB	pSS104_pLv1_DZ

Following Start-Stop Assembly, the plasmids were transformed and cloned in *E. coli* and incubated on appropriate LB agar plates. As the backbone of each vector contained a *LacZ* gene, initial screening was carried out based on colour of the bacterial colonies. White *LacZ-* colonies were chosen for further analysis by colony PCR or restriction digestion, while blue *LacZ+* colonies were discarded. Primers and annealing temperatures for colony PCR are given in Table 2.18. Promising candidates were cultivated in liquid LB (5 ml) containing 12 µg/mL tetracyclin. Plasmids were isolated using *QIAprep Spin Miniprep Kit* and sent for sequencing.

Table 2.18: Primers and annealing temperature (T) for colony PCR of *E. coli* transformed with level 1 constructs.

Level 1 construct	Primer	T [C°]
p102BC:PsbD_MaPylRS-RS1_petA	516 & 545	57
p103BC:PsbD_MbPylRS_petA	522 & 545	57
p104BC:PsbD_MbPylRS-Ack1_petA	522 & 545	57
p111DZ:p16s_mVenusOpal_atpB	457 & 546	47

2.19.1 Construction of Level 1 tRNA

Start-Stop Assembly, transferring tRNAw and *Mb*tRNA$^{Pyl}_{UCA}$ from level 0 to level 1, was performed with level 1 CD backbone. The process of generating level 1 constructs with the genes for *Ma*tRNA$^{Pyl}_{UCA}$ (p108CD:MatRNA) and *Mj*tRNA$^{Tyr}_{CUA}$ (p109CD_MjtRNA), was performed exactly as described for level 0 construction of these plasmids, with the modification of using Level 1 CD tRNAw as template instead of level 0 tRNAw. Additional restriction checks with BsaI were performed on level 1*Mb*tRNA$^{Pyl}_{UCA}$.

2.20 Assembly of Level 2 Intermediate

An intermediate (p207_psbH_aadA_tsPurple) for level 2 assembly was constructed containing the pUC8 backbone, left and right flank for homologous recombination within the *psbH* locus, the *aadA* expression unit and *tsPurple*. PCR was used to amplify the relevant sequence, except *tsPurple*, from a provided level 2 template. Primer 599 and 600 with an annealing temperature of 56°C was utilized. The PCR product was closed by ligating both ends to *tsPurple* in a PNK ligation. Following cloning in *E. coli*, purple *tsPurple+* colonies were picked and incubated overnight in liquid LB (5 ml) supplied with ampicillin (100 µg/ml). Plasmids were isolated using *QIAprep Spin Miniprep Kit* and analyzed by restriction digestion.

2.21 Assembly of Level 2 Constructs

Level 2 constructs were assembled utilizing Golden Gate Assembly containing *psbH* homology flanking sequence, as well as *aadA*, aaRS, tRNA and *mVenus* expression units. Two different approaches were used during level 2 construction assembly. Either, all seven units were assembled from seven different plasmids or the level 2 intermediate was assembled with appropriate aaRS, tRNA and *mVenus* genes. Plasmids used for the different assemblies are given in Table 2.20-2.23. After transformation in *E. coli*, cultivation was carried out on LB agar plates containing 100 µg/mL ampicillin. Three rounds of colony PCR were carried out with the set-up given in Table 2.19. Based on the analysis, colonies were incubated overnight in liquid LB (5 ml) supplied with ampicillin (100 µg/ml), then isolated using *QIAprep Spin Miniprep Kit*.

Table 2.19: Colony PCR primers and annealing temperature (T) used to amplify *aadA*, aminoacyl tRNA synthetase (aaRS), tRNA and *mVenus* expression units within the level 2 construct.

Primers	T [°C]	Amplified region
129 & 130	57	*aadA*
511 & 512	54	aaRS & tRNA
545 & 457	53	tRNA & *mVenus*

Table 2.20: Plasmids used for Golden Gate Assembly of the different level 2 constructs. Each plasmid either donate the expression unit of the aminoacyl tRNA synthetase (aaRS), tRNA, reporter or selectable marker, the left homology arm (HR-LF), the right homology arm (HR-RF) or provide a backbone for the assembled construct.

	p201_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal	p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal	p203_psbH_MbPylRS_MbtRNA_mVenusOpal
aaRS	p101BC_psbD_MaPylRS-3MeHis_petA	p102BC_psbD_MaPylRS-RS-1_petA	p103BC_psbD_MbPylRS_petA
tRNA	p108CD_MatRNA	p108CD_MatRNA	p107CD_MbtRNA
Reporter	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB
Selectable marker	pLv1_AB:atpA_aadA_rbcl	pLv1_AB:atpA_aadA_rbcl	pLv1_AB:atpA_aadA_rbcl
HR-LF	pHT146_pLv1_XA-psbH.LF	pHT146_pLv1_XA-psbH.LF	pHT146_pLv1_XA-psbH.LF
HR-RF	pHT150_pLv1_ZY-psbH.RF	pHT150_pLv1_ZY-psbH.RF	pHT150_pLv1_ZY-psbH.RF
Backbone	pSS191.Lv0/2_tsPurple	pSS191.Lv0/2_tsPurple	pSS191.Lv0/2_tsPurple

Table 2.21: Plasmids used for Golden Gate Assembly of the different level 2 constructs. Each plasmid either donate the expression unit of the aminoacyl tRNA synthetase (aaRS), tRNA, reporter or selectable marker, the left homology arm (HR-LF), the right homology arm (HR-RF) or provide a backbone for the assembled construct.

	p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal	p205_psbH_MmPylRS-AF_MbtRNA_mVenusOpal	p201_psbH_MjTyr-NitroTyr_MjtRNA_mVenusAmber
aaRS	p104BC_psbD_MbPylRS-Ack1_petA	p105BC_psbD_MmPylRS-AF_petA	p106BC_psbD_MjTyrRS-NitroTyr_petA
tRNA	p107CD_MbtRNA	p107CD_MbtRNA	p109CD_MjtRNA
Reporter	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB	p110DZ_p16s_mVenusAmber_atpB
Selectable marker	pLv1_AB:atpA_aadA_rbcl	pLv1_AB:atpA_aadA_rbcl	pLv1_AB:atpA_aadA_rbcl
HR-LF	pHT146_pLv1_XA-psbH.LF	pHT146_pLv1_XA-psbH.LF	pHT146_pLv1_XA-psbH.LF
HR-RF	pHT150_pLv1_ZY-psbH.RF	pHT150_pLv1_ZY-psbH.RF	pHT150_pLv1_ZY-psbH.RF
Backbone	pSS191.Lv0/2_tsPurple	pSS191.Lv0/2_tsPurple	pSS191.Lv0/2_tsPurple

Table 2.22: Plasmids used in combination with the level 2 intermediate for Golden Gate Assembly of the different level 2 constructs. Donor plasmids provide the expression units for the aminoacyl tRNA synthetase (aaRS), tRNA or reporter. The level 2 intermediate provides the selectable marker and both the left and right flank for homologous recombination.

	p201_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal	p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal	p203_psbH_MbPylRS_MbtRNA_mVenusOpal
aaRS	p101BC_psbD_MaPylRS-3MeHis_petA	p102BC_psbD_MaPylRS-RS-1_petA	p103BC_psbD_MbPylRS_petA
tRNA	p108CD_MatRNA	p108CD_MatRNA	p107CD_MbtRNA
Reporter	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB
Backbone	p207_psbH_aadA_tsPurple	p207_psbH_aadA_tsPurple	p207_psbH_aadA_tsPurple

Table 2.23: Plasmids used in combination with the level 2 intermediate for Golden Gate Assembly of the different level 2 constructs. Donor plasmids provide the expression units for the aminoacyl tRNA synthetase (aaRS), tRNA or reporter. The level 2 intermediate provides the selectable marker and both the left and right flank for homologous recombination.

	p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal	p205_psbH_MmPylRS-AF_MbtRNA_mVenusOpal	p201_psbH_MjTyr-NitroTyr_MjtRNA_mVenusAmber
aaRS	p104BC_psbD_MbPylRS-Ack1_petA	p105BC_psbD_MmPylRS-AF_petA	p106BC_psbD_MjTyrRS-NitroTyr_petA
tRNA	p107CD_MbtRNA	p107CD_MbtRNA	p109CD_MjtRNA
reporter	p111DZ_p16s_mVenusOpal_atpB	p111DZ_p16s_mVenusOpal_atpB	p110DZ_p16s_mVenusAmber_atpB
Backbone	p207_psbH_aadA_tsPurple	p207_psbH_aadA_tsPurple	p207_psbH_aadA_tsPurple

3 Results

3.1 Amplification of PsbL Regions for Homologous Recombination

The transgenes were to be introduced in the plastome of *C. reinhardtii* through homologous recombination. Two sites for transgene insertion were chosen within the *psbH* and *psbL* loci. PhD candidate M. Fages-Laurtad provided necessary sequences for homologous recombination within the *psbH* locus. Hence, only regions flanking the insertion site in *psbL* had to be amplified by polymerase chain reaction (PCR) from the chloroplast DNA (cpDNA) of *C. reinhardtii*.

The right flanking region of *psbL* was successfully amplified from cpDNA by PCR. As seen from the gel electrophoresis in Figure 3.1, amplified band size aligned with the expected value of 1061 bp. However, all attempts to amplify the left flank of *psbL* failed. The focus of the project was therefor shifted towards only developing vectors for homologous recombination within the *psbH* locus.

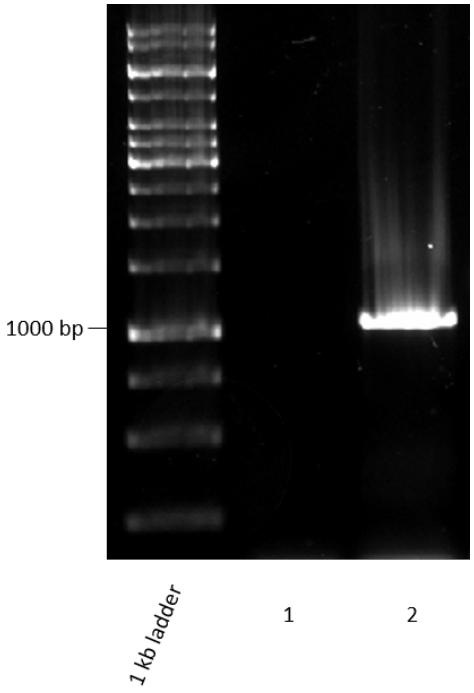


Figure 3.1: 0.8% agarose gel electrophoresis (40 min, 95V) after PCR-mediated amplification of *psbL* regions. No DNA detected in lane 1, expected to contain *psbL* left flank. Lane 2 contain *psbL* right flank. GeneRuler 1 kb DNA ladder (Thermo Scietific) was utilized as a standard.

3.2 Assembly of Storage Plasmids Containing Aminoacyl tRNA Synthetases

The coding sequence (CDS) of six aminoacyl tRNA synthetases (aaRS) were prepared to be used for codon reassignment. Multiple aaRS were included to increase chances of identifying aaRS/tRNA pairs orthogonal to the endogenouse aaRS/tRNA of *C. reinhardtii*. Additionally, it increased the quantity and diversity between non-canonical

amino acids (ncAA), which can utilized. CDS of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase-3MeHis (*Ma*PylRS-3MeHis), *Methanosarcina mazei* pyrrolysyl-tRNA synthetase-AF (*Mm*PylRS-AF), *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS) and *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase-NitroTyr (*Mj*TyrRS-NitroTyr) were amplified by PCR. These products were combined with a pUC8 backbone in a Gibson assembly, which after cloning was analyzed by colony PCR and gel electrophoresis. Results from the analysis are given in Figure 3.2.

Primer BB1 and BB2 were used for the first round of colony PCR. Based on the results given in Figure 3.2A, a single colony (lane 3) containing *Ma*PylRS-3MeHis (888 bp) was identified. The correct DNA sequence was confirmed by sequencing. pUC8_MmPylRS, pUC8_MbPylRS and pUC8_MjTyrRS were not identified. However, growth of *E. coli* was observed in the presence of ampicillin, indicating the presence of antibiotic resistance. Therefore, colony PCR was repeated with primers annealing to the pUC8 backbone, which contain a ampicilline resistance gene. Results of the following gel electrophoresis are given in Figure 3.2B. Nevertheless, all results were negative for the aaRS (>1000 bp). Instead, all DNA fragment sizes correlated with the expected size of the self-ligated pUC8 backbone (∼650).

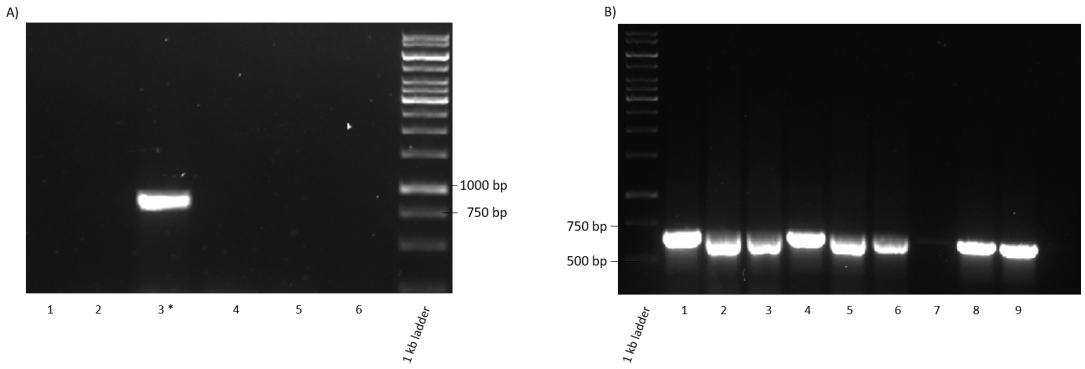


Figure 3.2: 0.8% gel electrophoresis (60 min, 120V) of colony PCR products, screening for assembly of plasmid pUC8 with inserted aminoacyl tRNA synthetases. GeneRuler 1 kb DNA ladder (Thermo Scietific) was utilized as a standard. A) Amplification using BioBrick primers. Lane 1,2,4,5 and 6: no DNA amplification. Lane 3: pUC8_MaPylRS-3Mehis. B)PCR primers annealing to pUC8 backbone Lane 1-6: self-ligated pUC8. Lane 7: no DNA amplification. The lane marked with a star (*) indicate correct assembly, verified through sequencing.

3.2.1 Construction of pUC8_MaPylRS-RS1 Utilizing Site-directed Mutagenesis

pUC8_MaPylRS-RS1 was constructed using pUC8_MaPylRS-3MeHis as a template. By reverting present mutations to wild-type and introducing point mutations (Y126M, M129G and V168T) the CDS was changed to that of MaPylRS-RS1. Mutations were introduced through two rounds of PCR, and the resulting DNA fragments were combined through Gibson assembly.

Figure 3.3 shows the gel electrophoresis of both PCR products and colony PCR after plasmid assembly and cloning in *E. coli*. The PCR products were expected to be 162 bp and 3125 bp, respectively. As seen from Figure 3.3B, obtained PCR product align with expected fragment size. Results from the colony PCR in Figure 3.3C indicates the

presence of *Ma*PylRS-RS1 in all tested colonies. The amplified DNA fragments were in compliance with the sought fragment size of 1763 bp. Sequencing confirmed the correct DNA sequence.

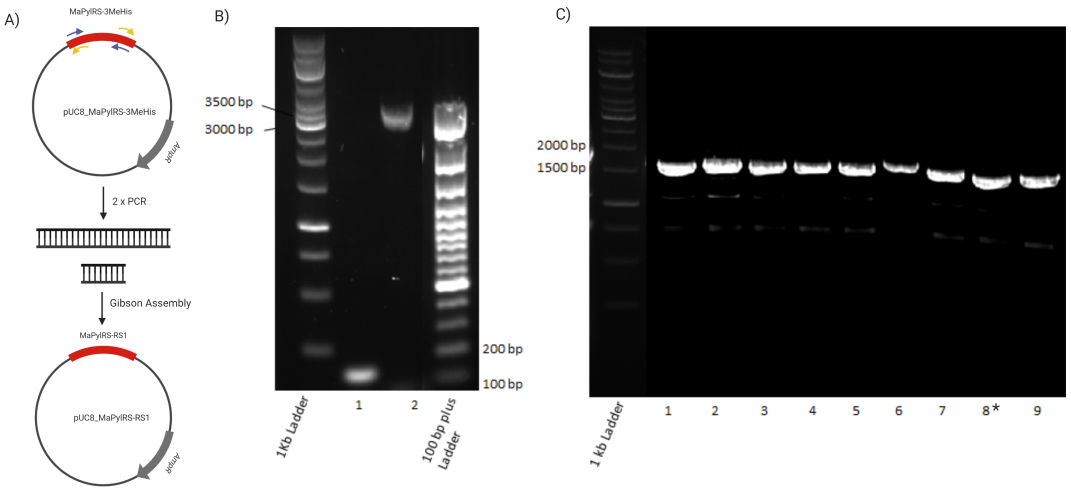


Figure 3.3: A) Synthesis of pUC8_*Ma*PylRS-RS1 by PCR mediated site-specific point mutation of pUC8_*Ma*PylRS-3MeHis. Two PCR was carried out with different primer pairs, indicated with blue and yellow arrows, respectively. The PCR products were assembled using Gibson Assembly. B) Gel electrophoresis (40 min, 95V) of the two PCR products. C) Gel electrophoresis (60 min, 120V) of colony PCR products of *Ma*PylRS-RS1_pUC8 cloned in *E. coli* (lane 1-9). GeneRuler 1 kb and 100 bp Plus DNA ladder (Thermo Scietific) were utilized as standards. Lane marked with a star (*) contain correct plasmids confirmed by seqcuencing.

In total, the CDS of four different aaRS were attempted to be combined with the pUC8 backbone. Of these, only pUC8_MaPylRS-3MeHis was successfully constructed, from which pUC_MaPylRS-RS1 was synthesized. These plasmids are shown in Figure 3.4. Both contain *ampR* conferring ampicillin resistance and SapI recognition sites for further assembly on level 1 constructs. Due to the placement of these recognition sites, the SapI cleavage sites are the ATG start codon and TAA stop codon of the aaRS gene.

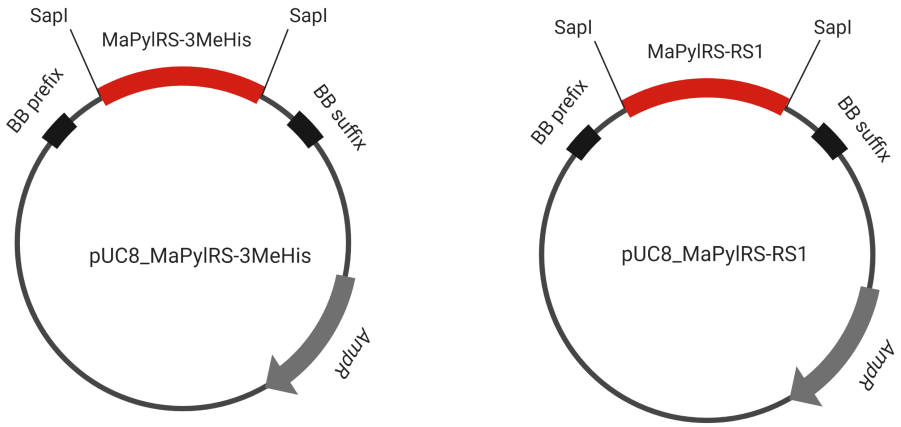


Figure 3.4: Structure of pUC8_MaPylRS-3MeHis and pUC8_MaPylRS-RS1_pUC8. The coding sequence of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase 3MeHis (*Ma*PylRS-3MeHis) and *Ma*PylRS-RS1 were assembled with pUC8 backbone at biobrick (BB) prefix and suffix using Gibson assembly. Both plasmid contain *AmpR* conferring ampicillin resistance. SapI cleavage sites are located at the start and stop codons of the coding sequence.

3.3 Assembly of Level 0 Storage Vectors

Attempts of assembling the remaining aaRS, tRNA, promoters/5'UTR and monomeric *Venus* (*mVenus*) on level 0 constructs were carried out. Construction for all but the tRNAs was achieved using Golden Gate Assembly, utilizing BsaI restriction enzyme and pSS191.Lv0/2_tsPurple as an acceptor plasmid. *tsPurple* was replaced by the insert of interest. Before assembly, all parts were PCR amplified with flanking BsaI and SapI recognition sites as well as appropriate cleavage sites for Start-Stop Assembly.

Three different variants of *mVenus* were generated. *mVenusAmber* containing an amber stop codon at position 3, *mVenusOpal* with an additional opal stop codon at position 151 and *mVenusAO* containing both the amber and opal stop codons at position 3 and 151, respectively. Furthermore, a flanking DNA sequence (5'-CACCACCACCACCACCACTAA-3') was added at the 3' end. This sequence provided six consecutive histidine codons providing a histidine-tag to the translated protein. The histidine-tag was added to identify synthesised mVenus by eventual mass spectrometry with more ease. The ochre stop codon (TAA) was chosen as the stop codon for the gene, as not to be affected by amber and opal codon reassignment.

The combined promoters and 5'UTR of *psbD*, *psbB* and *rbcl* were amplified from cpDNA of *C. reinhardtii* by PCR. Amplification of correct fragment size was evaluated by gel electrophoresis. As seen from Figure 3.5, product sizes align with predicted fragment size for promoter/5'-UTR of psbB (340 bp), psbD (213 bp) and rbcl (220 bp).

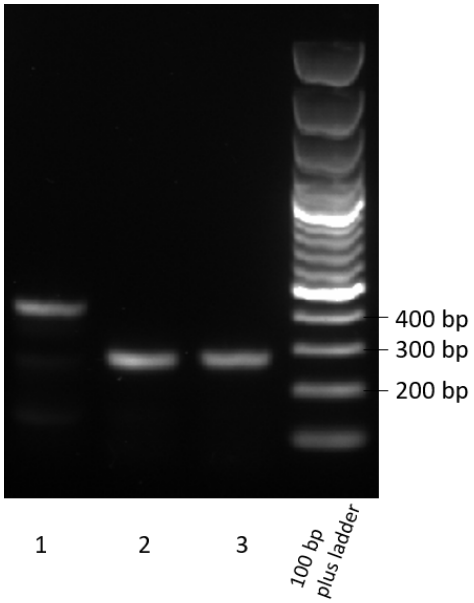


Figure 3.5: 0.8% Agarose gel electrophoresis (95 V, 40 min) after PCR mediated amplification of psbB (lane 1), psbD (lane 2) and rbcl (lane 3) promoters/5'-UTR. GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific) was utilized as standard.

The assembly of *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS) *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*TyrRS) with pUC8 backbones failed. Therefore, the CDS of these aaRS were re-amplified with flanking BsaI and SapI recognition sites, instead of BB prefix and suffix. Resulting DNA bands, separated and visualized by gel electrophoresis, were consistent with the desired fragment size. Table 3.1 contains a list of all assembled level 0 constructs.

Table 3.1: List of all assembled level 0 constructs containing either a promoter/5'untranslated region,the coding sequence of an aminoacyl tRNA synthetase (aaRS), tRNA or a reporter gene. With the exception of p010_MbtRNA, all constructs are listed with estimated colony PCR product size.

aaRS	Promoter/5'UTR	Reporter	tRNA
p001_MbPylRS (1938 bp)	p004_psbB (1022 bp)	p007_mVenusAmber (1422 bp)	p010_MbtRNA
p002_MbPylRS-Ack1 (1938 bp)	p005_psbD (1020 bp)	p008_mVenusOpal (751 bp)	
p003_MjTyr-NitroTyr (1599 bp)	p006_rbcl (903 bp)	p009_mVenusAO (1422 bp)	

Each level 0 construct contains SapI restriction sites with specific cleavage sites. Illustrations of level 0 constructs with inserted promoter or CDS are given Figure 3.6. As seen from the figures, level 0 promoters contain α and γ SapI cleavage sites, while all level 0 CDS were designed with γ and ω SapI cleavage sites. These sites were designed to function as fusion sites for level 1 assembly. Table 3.2 gives further details on the 3-bp cleavage sites. As a result of the Golden Gate Assembly of level 0 constructs, each vector carry scarring at junctions between insert and backbone. These scars are the BsaI cleavage sites, symbolized with X and Y. Lastly, *ampR*, introduced as part of the backbone provided ampicillin resistance for selective growth during cloning in *E. coli*.

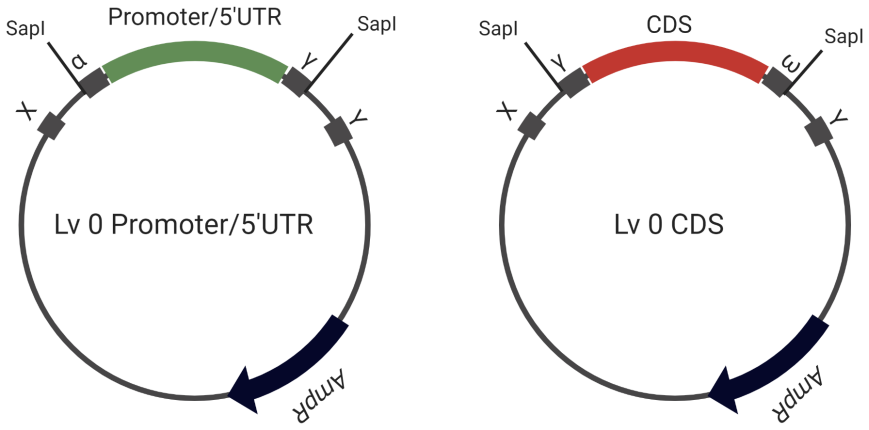


Figure 3.6: General scheme of level (lv) 0 constructs containing promoters/5'UTR or coding sequences (CDS). Each plasmid carry a *AmpR* gene conferring ampicillin resistance. X and Y symbolise scarring left by Golden Gate Assembly. The promoter/5'UTR or CDS is flanked by SapI recognition sites and specially designed cleavage sites (α, γ and ω). Details on SapI cleavage sites are given in Table 3.2.

Table 3.2: List of SapI cleavage sites present in level 0 constructs.

Sap I cleavage site	DNA sequence	Comment
α	5'-CTG-3'	
γ	5'-ATG-3'	Start codon
ω	5'-ATT-3'	Ochre stop codon
ϵ	5'-GGA-3'	

Three steps of screening was carried out to identify constructs with correct assembly. Initial identification of *tsPurple+* and *tsPurple- E. coli* was performed with ease due coloration of the bacterial colonies by *tsPurple*. This is exemplified in Figure 3.7. The figure show growth *E. coli* following transformation with p008_mVenusOpal. White *tsPurple- E. coli* colonies were chosen for further screening, while purple *tsPurple+ E. coli* were discarded.

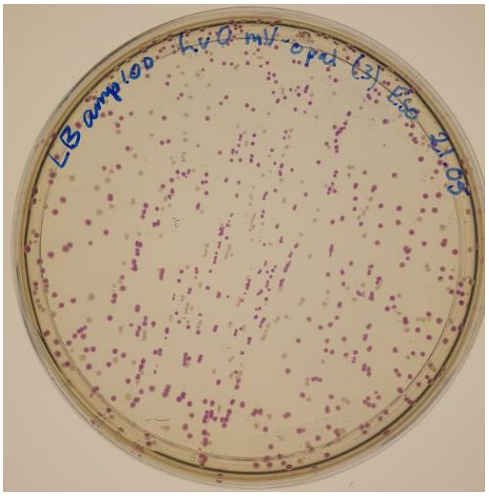


Figure 3.7: Growth of *tsPurple+* (purple) and *tsPurple-* (white) *E. coli* colonies on LB agar supplied with ampicillin (100 µg/ml) following transformation with p009_mVenusOpal.

Colony PCR and gel electrophoresis were conducted to identify the presence of assembled constructs. The results of the agarose gel electrophoresis are given in Figure 3.8 and Figure 3.9. All expected PCR product sizes are listed in Table 3.1. Colony PCR products which aligned with these values were identified for each construct, indicating correct assembly. The correct DNA sequence was confirmed through sequencing. Lanes which contain these samples are marked with a star (*). PCR products of unwanted fragment sizes were not further analyzed. Therefore, no conclusion can be drawn as to what they contain.

In addition to correct DNA bands, each colony containing p001_MbPylRS (Figure 3.8B), p007_mVenusAmber and p009_mVenusAO (Figure 3.9B) contained additional PCR products. These bands were weaker than the expected bands but considered prominent enough to be mentioned. In all cases, the DNA fragments were approximately 700 bp.

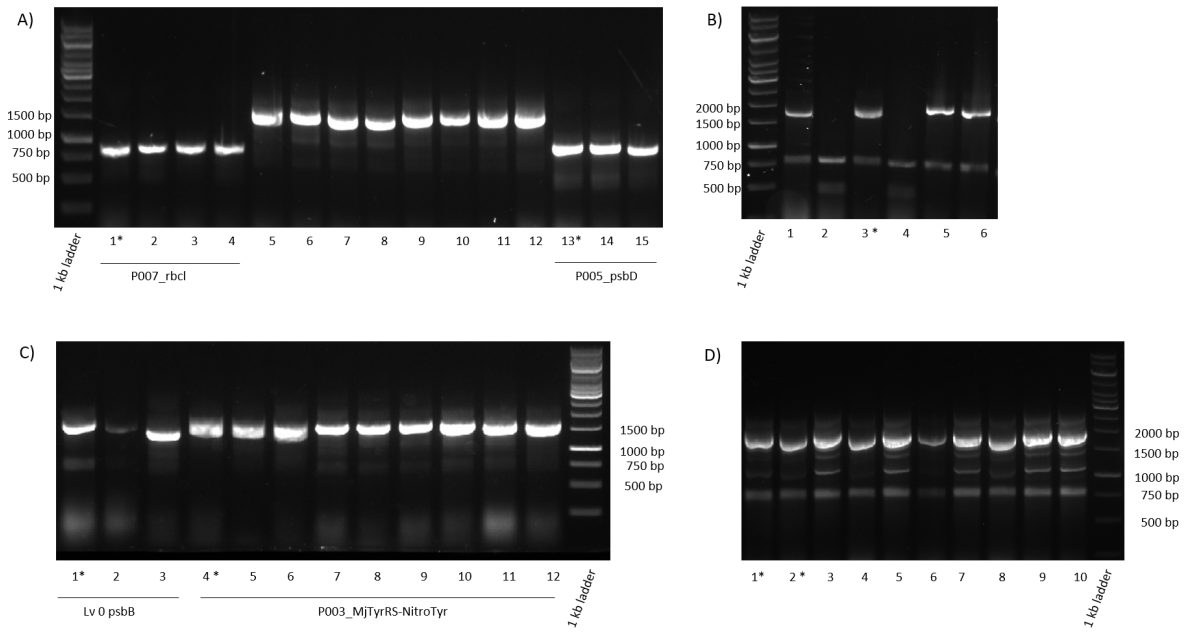


Figure 3.8: 0.8% Agarose gel electrophoresis of colony PCR products. Samples containing correct DNA sequence, verified by sequencing are indicated with a star (*). GeneRuler 1 kb DNA Ladder (Thermo Scientific) was utilized as standard. A) Gel electrophoreis (60 min, 120V) screeing for p007_rbcl (lane 1-4) and p005_psbD (lane 13-15). B) Gel electrophoresis (45 min, 95V), screening for p001_MbPylRS. C) Gel electrophoreis (60 min, 120V) of p004_psbB (lane 1-3) and p003_MjTyrRS-NitroTyr (lane 4-12). D) Gel electrophoreis (45 min, 95V), screening for p002_MbPylRS-Ack1.

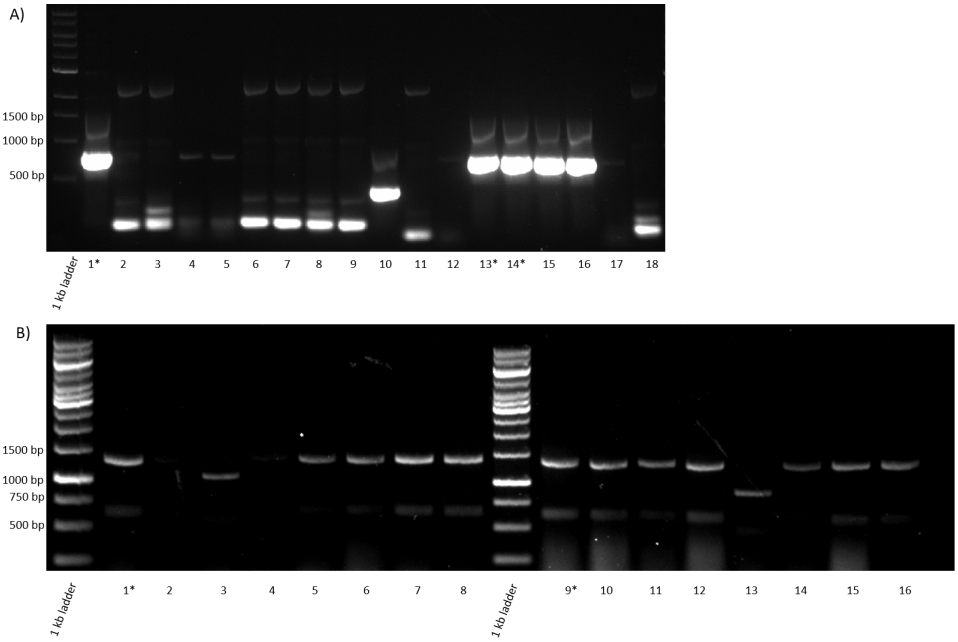


Figure 3.9: 0.8% Agarose gel electrophoresis of colony PCR products. Samples containing correct DNA sequence, verified by sequencing are indicated with a star (*). GeneRuler 1 kb DNA Ladder (Thermo Scientific) was utilized as standard. A) Gel electrophoreis (60 min, 120V), screening for p008_mVenusOpal. B) Gel electrophoresis (45 min, 95V) screening for p007_mVenusAmber (1-8) and p009_mVenusAO (9-16).

3.3.1 Construction of p002_MbPylRS-Ack1 Utilizing Site-directed Mutagenesis

The CDS of mutant version *Mb*PylRS-Ack1 (D76G, L226V, L270I, Y271F, L274A and C313F) was synthesized from wild type (wt) *Mb*PylRS due to its abilities to esterify the ncAA N^{ϵ}-acetyllysin onto the cognate tRNA. Because of a relatively large distance between the point mutations, an intermediate was constructed. PCR mediated site-directed mutagenesis was utilized to introduce the mutations: L226V, L270I, Y271F, L274A. Results from gel electrophoresis of the two PCR products are given in Figure 3.10. The PCR products correlate well with expected fragment sizes of 1499 bp and 1990 bp in lane 1 and 2, respectively. Gibson assembly was utilized to combine the PCR products. Assembly was verified by gel electrophoresis of colony PCR products after cloning in *E. coli*. Results show that all tested colonies contained amplified regions aligning with the expected value of 1938 bp. However, additional products of approximately 750 bp were observed in all samples. These bands are not as intense as the expected bands. Yet, prominent enough to be taken notice of.

The final level 0 construct containing *Mb*PylRS-Ack1 (p002_MbPylRS-Ack1) was generated from the intermediate, as described in Figure 3.11A. The point mutations D76G and C313T were introduced through two PCR reactions. Results from the following gel electrophoresis are shown in Figure 3.11B. The same trend as for the intermediate was observed. As seen in Figure 3.11C, fragments aligning with expected fragment size (1938 bp) were amplified in all tested samples. Furthermore, the additional 750 bp PCR products were still present in all samples. Sequencing of the sample from lane 1 showed that p002_MbPylRS-Ack1 had been correctly assembled with all the point-mutations present. The content of the 750 bp product was not further analyzed.

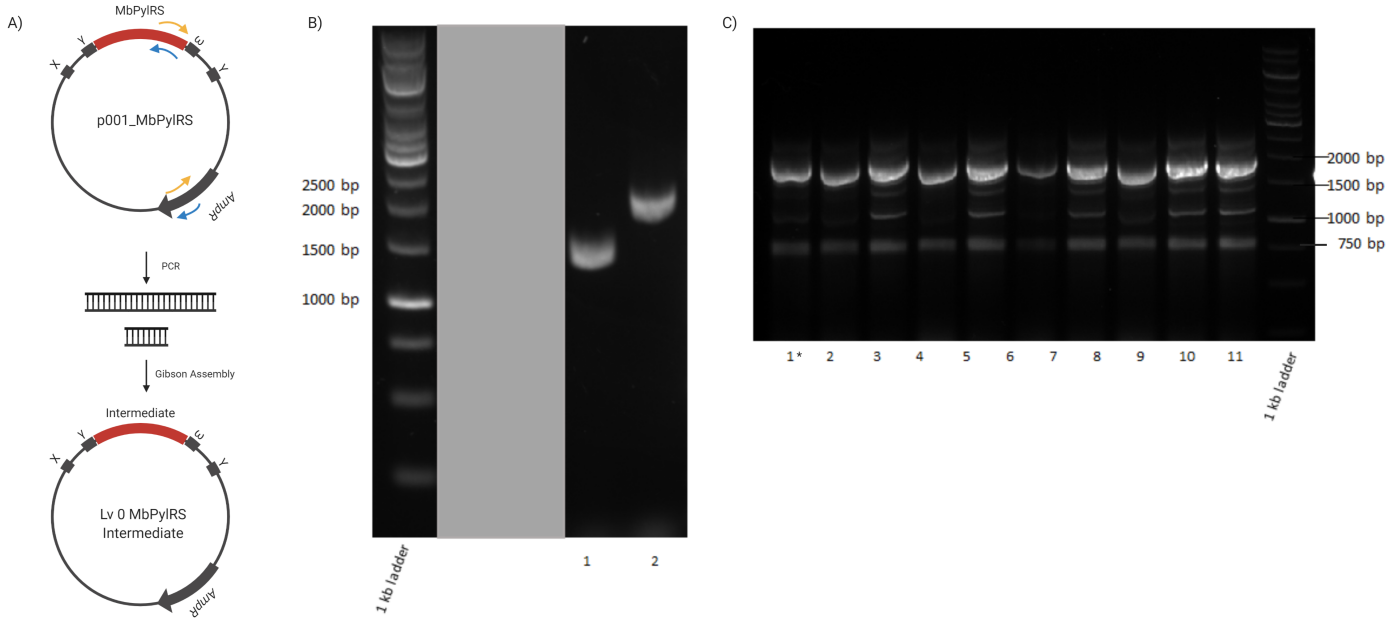


Figure 3.10: A) Schematic representation of intermediate generated from p001_MbPylRS. Point mutations (D76G, L226V, L270I, Y271F, L274A and C313F) were introduced through two PCR reactions, utilizing different primer pairs. These primer pairs are indicated with blue and yellow arrows, respectivly. γ and ω represent SapI cleavage sites, while X and Y are scars from Golden Gate Assembly. B) 0.8% Agarose gel electrophoresis (45 min, 95V) of PCR products confirming amplification of PCR products with correct length, approximately 1499 and 1990 bp, respectively. C) 0.8% Agarose gel electrophoresis (60 min, 120V) of colony PCR products identified product size aligning with expected value (1938 bp) in all samples. The grey area exclude unrelated results. GeneRuler 1 kb DNA Ladder (Thermo Scientific) was utilized as standard.

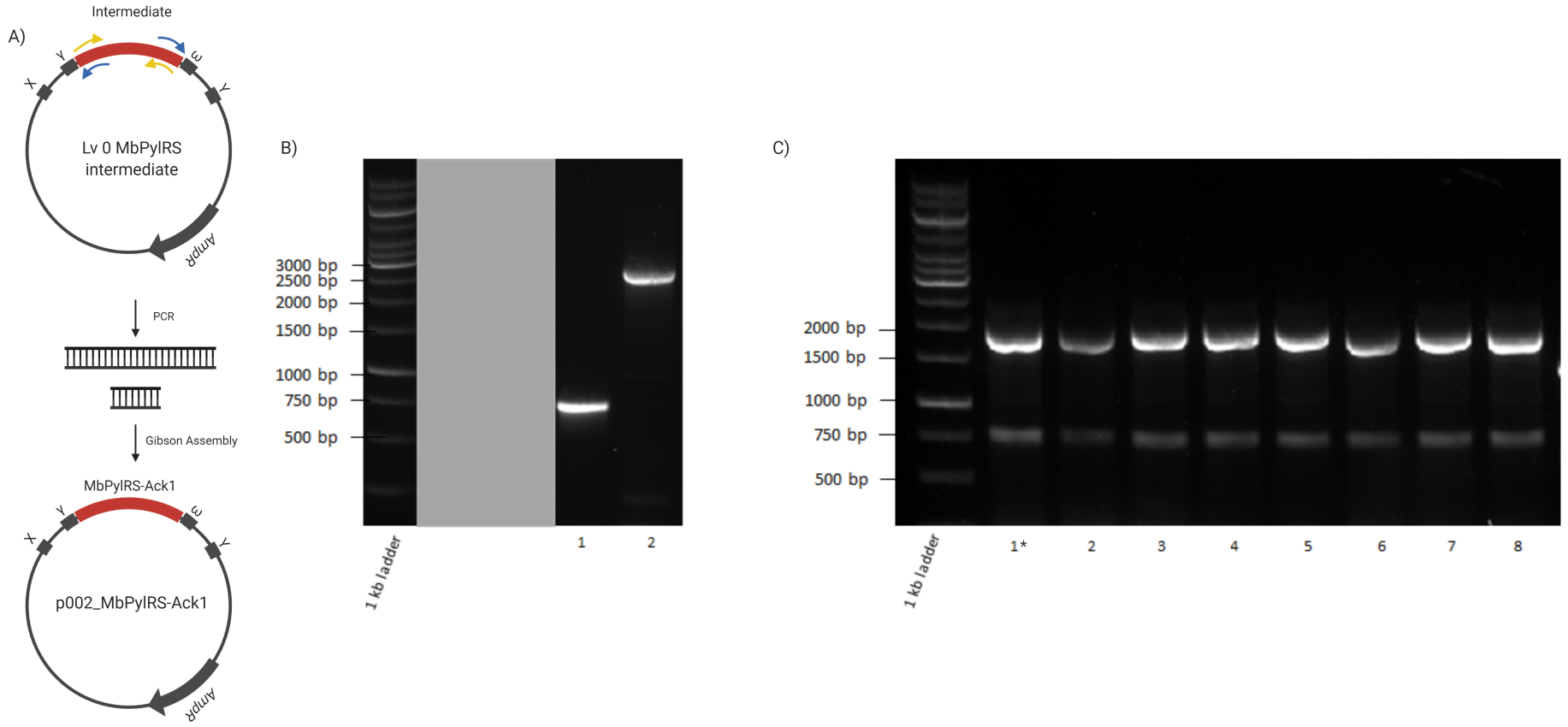


Figure 3.11: A) Overview of the second round of assembly to generate p002_*Mb*PylRS-Ack1 (D76G, L226V, L270I, Y271F, L274A and C313F). Introducing the two point mutations D76G and C313T through two PCR reactions, using different primer pairs indicated with blue and yellow arrows. γ and ω represent SapI cleavage sites, while X and Y are scars from Golden Gate Assembly. B) 0.8% Agarose gel electrophoresis (45 min, 95V) of PCR products confirming PCR amplification of DNA fragments, approximatly 748 and 2715 bp in lane 1 and 2, respectively. C) 0.8% Agarose gel electrophoresis (60 min, 120V) of colony PCR products hold fragments aligning with expected values for p002_MbPylRS-Ack1 (1938 bp) in all lanes. Sample with correct construct assembly, verified by sequencing is marked with a star(*). GeneRuler 1 kb DNA Ladder (Thermo Scientific) was utilized as standard.

3.3.2 PCR-mediated Amplification of tRNA on Level 0 Constructs

The CDS of *Mb*tRNA$^{Pyl}_{UCA}$, *Ma*tRNA$^{Pyl}_{UCA}$ and *Mj*tRNA$^{Tyr}_{CUA}$ were attempted introduced onto level 0 tRNA^{Trp} (pHT201.Lv0.tRNAW), employing PCR with primers containing the necessary DNA sequence, replacing the tRNA^{Trp} CDS. Here, screening was only performed through sequencing, as *tsPurple* is not used. Additionally, colony PCR would not differentiate between the template and the synthesized product. The correct assembly of p010_MbtRNA was confirmed. However, all sequenced level 0 constructs, which contained *Ma*tRNA$^{Pyl}_{UCA}$ and *Mj*tRNA$^{Tyr}_{CUA}$ lacked up to several nucleotide in the fused region. Unlike the other level 0 constructs, p010_MbtRNA did not lack any nucleotides. The plasmid contain a full expression unit; 5'UTR and 3'UTR of tRNA^{Trp} coupled to the CDS of MbtRNA$^{Pyl}_{UCA}$. As seen in Figure 3.12 the plasmid contains α and ϵ SapI cleavage sites, allowing complete assembly on level 1 without the need of more inserts.

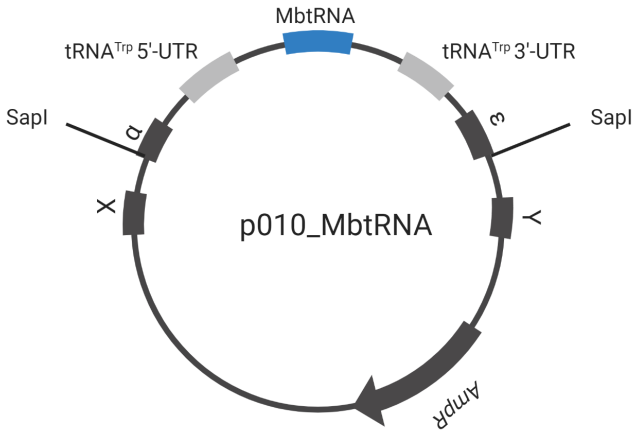


Figure 3.12: Illustration of p010_MbtRNA containing the coding sequence of *Mb*tRNA$^{Pyl}_{UCA}$ as well as 3'untranslated region (UTR) and 5'UTR of tRNA^{Trp}. The plasmid carry a *ampR* gene confering ampicillin resistance. Inserts are flanked by SapI recognition sites and specially designed cleavage sites (α and ϵ). Details on SapI cleavage sites are given in Table 3.2. The plasmid also contain scars (X and Y) which were present in the template, caused by a previous DNA Assembly.

3.4 Assembly of Functionally Scarless Expression Units on Level 1 Constructs

Complete expression units were assembled on level 1 constructs. CDS were coupled to promoters/5'UTR and 3'UTR of highly expressed chloroplast genes and assembled on level 1 constructs. Functionally scarless expression units were obtained utilizing the Start-Stop Assembly. The assembly was performed using the previously constructed level 0 constructs or PCR products and a level 1 acceptor plasmid containing *tetR*, providing tetracyclin resistance. A start codon functioned as a fusion site between promoter/5'UTR and CDS, while ochre stop codon was the fusion site between CDS and 3'UTR. *LacZ* initially present in the acceptor plasmid was replaced by the inserts of interest.

Figure 3.13, shows growth of *E. coli* after transformation with p102_BC:PsbD_MaPylRS-RS1_petA. Similar growth patterns could be seen in on all plates containing *E. coli* level 1 constructs. Initial screening after cloning in *E. coli* was performed by differentiating between blue *LacZ+* and white *LacZ-* *E. coli* colonies.

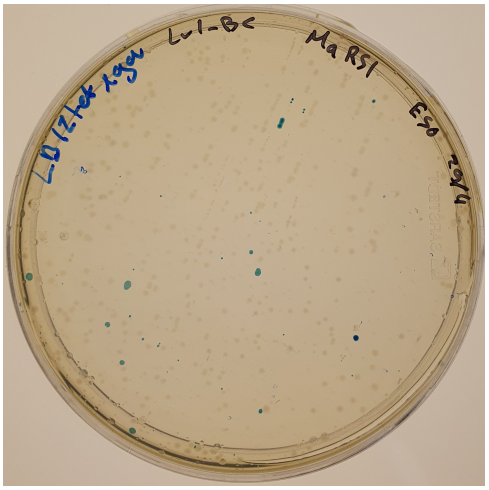


Figure 3.13: Growth of *LacZ+* (blue) and *LacZ-* (white) *E. coli* colonies on LB agar plate supplied with tetracyclin (100 µg/ml), after transformation with p102BC:PsbD_MaPylRS-RS1_petA.

3.4.1 Assembly of Level 1 Reporter Constructs for Assessment of Codon Reassignment

Reporter cassettes were assembled on level 1 constructs for all three mVenus variants. The constructs were all assembled with *p16s* promoter, *psaA* 5'UTR, *atpB* 3'UTR as well as *mVenusAmber* (p110DZ:p16s_mVenusAmber_atpB), *mVenusOpal* (p111DZ:p16s_mVenus-Opal_atpB) or *mVenusAO* (p112DZ: p16s_mVenusAO_atpB). The backbone provid BsaI recognition site, as well as D and Z fusion sites (BsaI cleavage sites) for level 2 assembly. Restriction digestion using BsaI was performed to screen for correct DNA assembly of p110DZ:p16s_mVenusAmber_atpB and p112DZ:p16s_mVenus-AO_atpB. The digestion was expected to result in two fragments of 1291 bp and 2997 bp for all three plasmids. Such fragmentation was observed in lane 1 and 4 of Figure 3.14, containing p110DZ: p16s_mVenusAmber_atpB and p112DZ: p16s_mVenus-AO_atpB, respectively. Sequencing of these plasmids confirmed correct assembly.

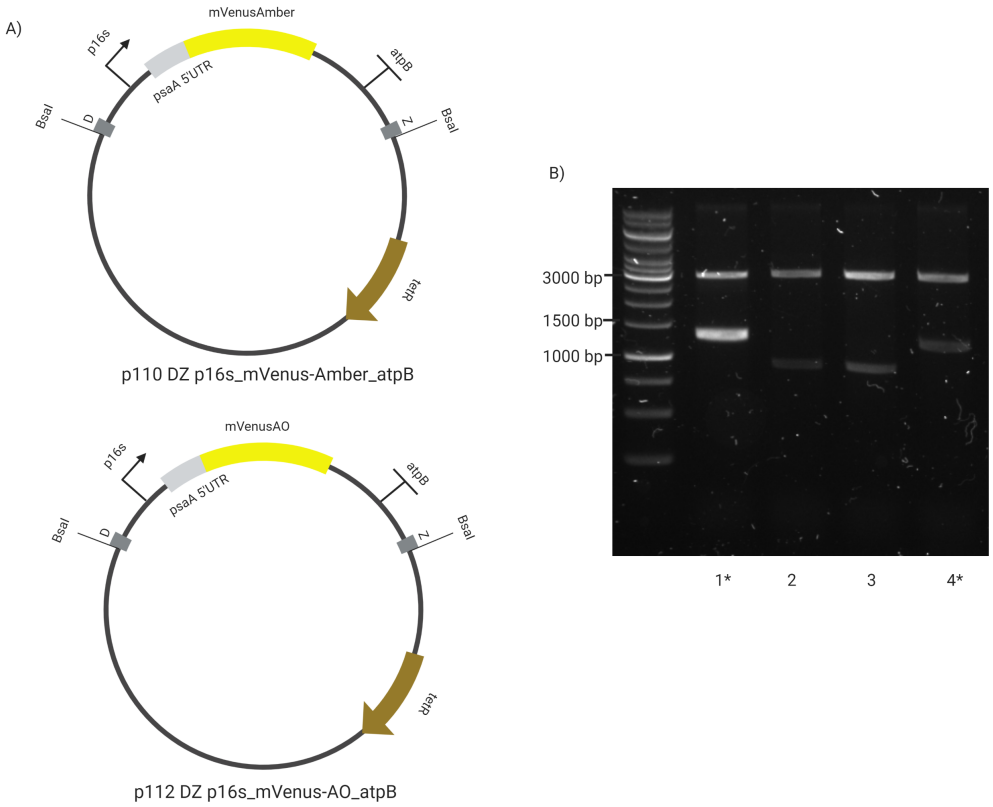


Figure 3.14: A) Illustration of p110DZ:p16s_mVenusAmber_atpB and p112DZ:p16s_mVenus-AO_atpB with D and Z fusion site for level 2 assembly. Each plasmid contain *tetR* providing tetracyclin resistance. The expression unit consist of a *p16s* promoter, *psaA* 5'UTR, *atpB* 3'UTR and the coding sequence of *mVenus* with an additional amber codon (*mVenusAmber*) or amber and opal codon (*mVenusAO*). Constructs were assembled through Start-Stop Assembly. B) Gel electrophoresis (40min, 95V) after restriction digestion using BsaI. Lane 1 and 2: p110DZ: p16s_mVenusAmber_atpB. Lane 3 and 4: p112DZ: p16s_mVenusAO_atpB. Expected DNA band size was 1291 bp and 2997 bp for both constructs. Samples containing correct assembly, confirmed by seqeuncing are marked with a star (*). GeneRuler 1 kb DNA Ladder (Thermo Scientific) was utilized as a standard.

Colony PCR was performed to screen for p111DZ:p16s_mVenus-Opal_atpB. Results from the following gel electrophoresis are given in Figure 3.15. Based on PCR product size , samples in lane 3, 4, 6 and 17 could contain p111DZ:p16s_mVenus-Opal_atpB (1088 bp), which was confirmed for plasmids in lane 3 and 4 through sequencing. The content of the remaining samples ($\sim$ 500 bp) was not further analyzed. Hence, no conclusion of its content was drawn.

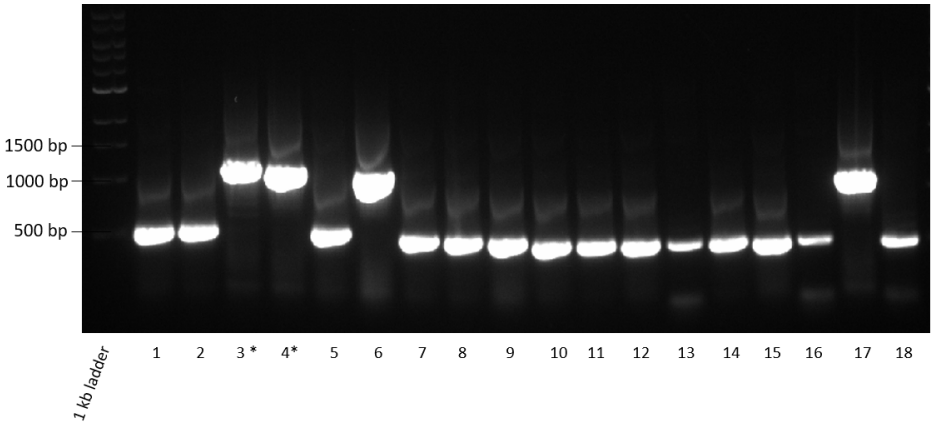


Figure 3.15: Gel electrophoresis (60 min, 120V) of colony PCR products after cloning of p111DZ:p16s_mVenus-Opal_atpB in *E. coli*. Lane 3, 4, 6 and 17: p111DZ:p16s_mVenus-Opal_atpB (1088 bp). Lanes marked with a star (*) contain correctly assembled plasmid, confirmed by sequencing. 1kb DNA Ladder (NEB) was utilized as standard.

3.4.2 Assembly of Level 1 aaRS Constructs

All aaRS CDS were combined with the chloroplast promoter/5'UTR of *psbD* and *petA* 3'UTR on a level 1 acceptor plasmid. Schemes of all level 1 aaRS constructs are given in Figure 3.16. The backbone provided *TetR*, two BsaI recognition sites as well as type B and C cleavage sites.

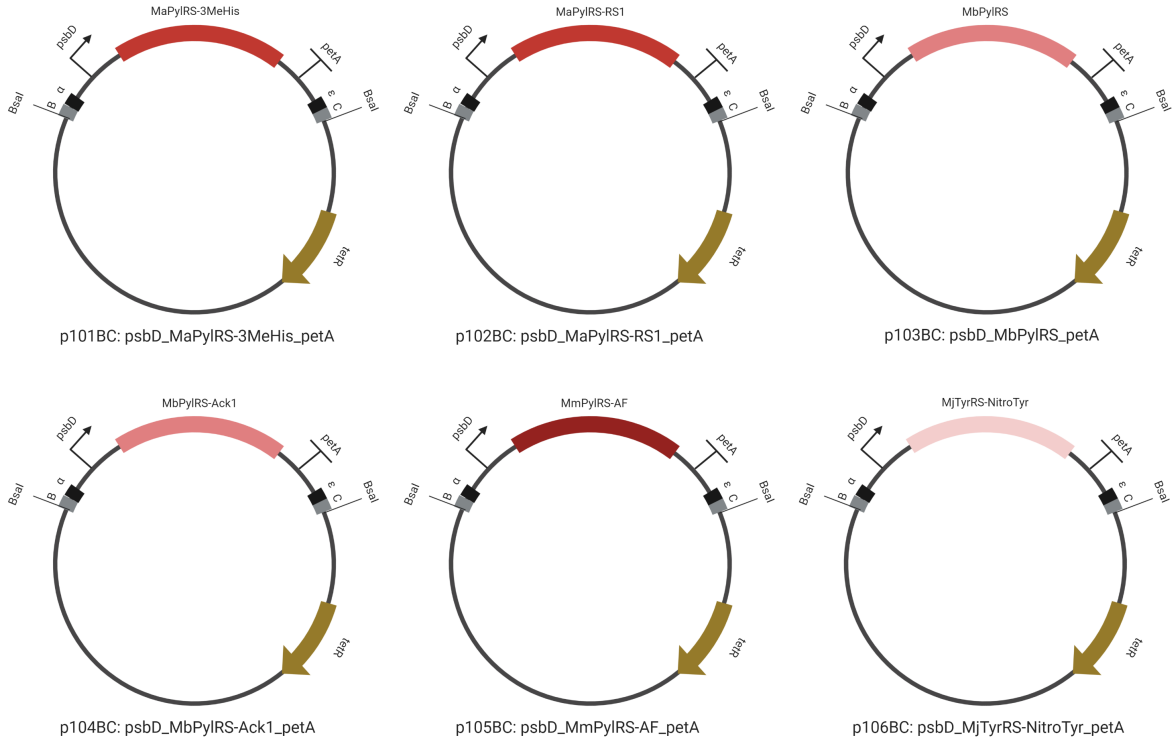


Figure 3.16: Illustration of all level 1 aminoacyl tRNA synthetase (aaRS) constructs. Each construct contains a *psbD* promoter/5'UTR and *petA* 3'UTR, in addition to their respective aaRS CDS. Six aaRS were assembled on respective constructs: *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase 3MeHis (*Ma*PylRS-3MeHis), *Ma*PylRS-RS1, *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS), *Mb*PylRS-Ack1, *Methanosarcina mazei* pyrrolysyl-tRNA synthetase AF (*Mm*PylRS-AF) and *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase NitroTyr (*Mj*TyrRS-NitroTyr). Fusion sites between backbone and promoter/5'UTR and 3'UTR are symbolised with α and ϵ, respectively. BsaI recognition and cleavage sites (B and C) as well as *tetR* for tetracyline resistance were introduced as part of the backbone structure.

Restriction checks using BsaI or colony PCR was conducted to screen for plasmids containing correct inserts. Amplified or DNA fragment size from restriction digestion of all level 1 aaRS are listed in Table 3.3. Results from the gel electrophoresis after restriction digestion are shown in Figure 3.17A. Lane 1 and 2 were expected to contain p101BC:psbD_MaPylRS-3MeHis_petA. One sample was cut as expected (lane 1), while the other sample (lane 2) gave an additional band around 2000 bp. As seen from the same figure, results from both p106BC:psbD_MjTyrRS-NitroTyr_petA and p105BC:psbD_MmPylRS_petA aligned with expected values. Colony PCR and gel electrophoresis was used to identify p102BC:psbD_MaPylRS-RS1_petA, p104BC:psbD_MbPylRS_petA and p105BC:psbD_MbPylRS-Ack1_petA. These results are given in Figure 3.17B and 3.17C. All DNA fragments aligned with expected values, indicating correct construct assembly. Plasmids containing correct assembly, without significant mutations, were identified for each plasmid through sequencing. Blank lanes indicate failed PCR amplification. No further analysis was conducted for these, and samples were discarded.

Table 3.3: List of DNA fragment size after restriction digestion using BsaI, or amplified DNA region by colony PCR of level 1 constructs containing expression units for aminoacyl tRNA synthetases.

Type	Construct	Fragment 1 [bp]	Fragment 2 [bp]
Restriction check	p101BC:psbD_MaPylRS-3MeHis_PetA	1293	2997
	p105BC:psbD_MmPylRS_petA	1830	2997
	p106BC:psbD_MjTyrRS-NitroTyr_petA	1386	2997
Colony PCR	p102BC:psbD_MaPylRS-RS1_petA	735	
	p103BC:psbD_MbPylRS_petA	1165	
	p104BC:psbD_MbPylRS-Ack1_petA	1165	

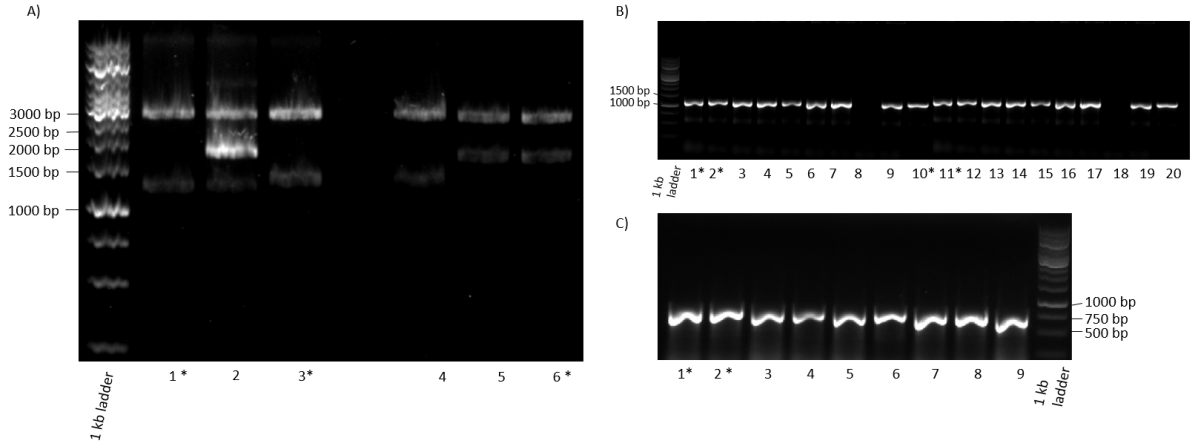


Figure 3.17: A) 0.8% Agarose gel electrophoresis (40 min, 95V) after restriction digestion using BsaI. The gel is loaded with restricted candidates assumed to be p101BC:psbD_MaPylRS-3MeHis_petA (lane 1 and 2), p106BC:psbD_MjTyrRS-NitroTyr_petA (lane 3 and 4) and p105BC:psbD_MmPylRS_petA (lane 5 and 6). Bands from all samples, except lane 2, match with expected fragment size. Expected DNA fragment sizes are listed in Table 3.3. B) 0.8% Agarose gel electrophoresis (45 min, 120V) of colony PCR screening for p103BC:psbD_MbPylRS_petA (lane 1-9) and p104BC:psbD_MbPylRS-Ack1_petA (lane 10-20). C) 0.8% Agarose gel electrophoresis (45 min, 120V) of colony PCR screening for p102BC:psbD_MaPylRS-RS1_petA. Lanes marked with a star (*) contain samples with correct assembly, confirmed through sequencing. GeneRuler 1kb DNA Ladder was used as standard in all the gels.

3.4.3 Development of Level 1 tRNA Constructs

Synthesis of *Ma*tRNA$^{Pyl}_{UCA}$ and *Mj*tRNA$^{Tyr}_{CUA}$ on level 0 failed. In order to save time and resources level 0 tRNA^{Trp} was transferred to level 1, allowing for synthesis of the other tRNA CDS directly on level 1. Constructs containing *Ma*tRNA$^{Pyl}_{UCA}$ (p108DZ:MatRNA) and *Mj*tRNA$^{Tyr}_{CUA}$ (p109CD:MjtRNA) were then constructed by PCR amplifying the tRNA CDS onto the level 1 template, and subsequently closing the PCR products by PNK ligation. Six of each plasmid were sent for sequencing. Correct sequence, in one p109CD:MjtRNA and two p108DZ:MatRNA were confirmed, without missing nucleotides in the ligated region. Illustrations of each level 1 construct containing a tRNA gene is presented in Figure 3.18 All level 1 tRNA constructs contain expression unit consisting of the respective CDS of each tRNA, as well as 5'UTR and 3'UTR of tRNA^{trp}. Genes for *Ma*tRNA$^{Pyl}_{UCA}$ and *Mb*tRNA$^{Pyl}_{UCA}$ contain an opal anticodon, while a version containing

an amber codon was chosen for the gene of *Mj*tRNA$^{Tyr}_{CUA}$. Additionally, level 1 constructs contain *tetR*, two BsaI recognition sites and type C and D fusion sites.

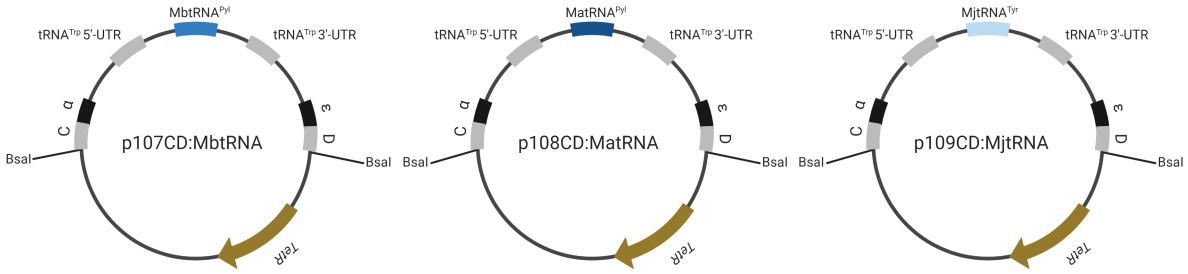


Figure 3.18: Illustration of all level 1 tRNA constructs. Each construct contains 5' untranslated region (UTR) and 3'UTR from tRNA^{Trp}, in addition to their respective tRNA gene. Three different tRNA were assembled on respective constructs p107CD:MbtRNA contain *Methanosarcina barkeri* tRNA (*Mb*tRNA^{Pyl}), p108CD:MatRNA contain *Methanomethylophilus alvus* tRNA (*Ma*tRNA^{Pyl}), while p109CD:MjtRNA hold a *Methanocaldococcus jannaschii* tRNA (*Mj*tRNA^{Tyr}). Each plasmid contain BsaI recognition and cleavage sites (B and C), as well as *tetR* providing tetracycline resistance. α and ϵ indicate fusion sites, used during Start-Stop Assembly of p107CD:MbtRNA.

p107DC_MbtRNA was constructed through Start-Stop assembly, extracting MbtRNA$^{Pyl}_{UCA}$ from p010_MbtRNA and fusing it to a level 1 backbone. Sequencing identified lack of one BsaI recognition site on the level 1 construct. Therefore, a second assembly was conducted. This time a restriction check was performed with BsaI before sequencing. Figure 3.19 shows results from the gel electrophoresis of the restricted plasmids. Lane 1 and 4 contain two fragments each, indicating the presence of two BsaI recognition sites within the construct. The size of the largest band is consistent with the expected value of 2997 bp. The second band was expected to be 228 bp. However, due to the choice of ladder, the size could not be estimated more specific than being lower than 500 bp. Nonetheless, sequencing of these plasmids confirmed the correct sequence with both BsaI recognition site intact.

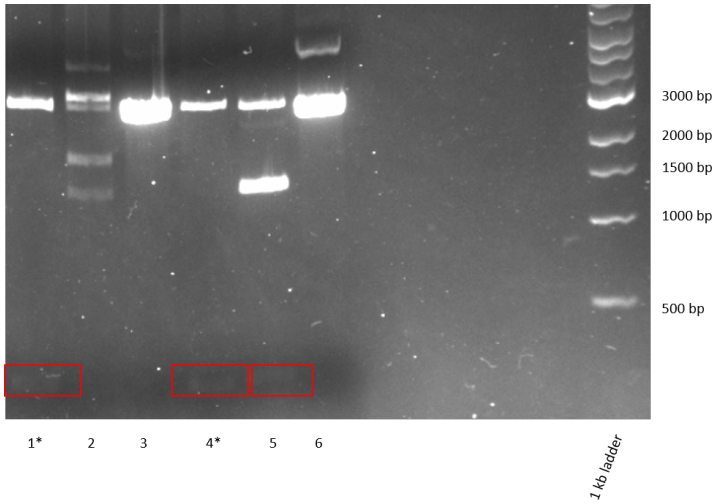


Figure 3.19: 0.8% Agarose Gel electrophoresis (40 min, 95V) after restriction digestion of p107CD_MbtRNA using BsaI. Samples in lanes marked with a star (*) correct sequence confirmed by sequencing. These lanes contain two bands, where the largest align with expected value of 2997 bp. Weak bands smaller than 500 bp are marked with red boxes. 1kb DNA Ladder (NEB) was utilized as standard.

3.5 Construction of Level 2 Intermediate p207_psbH_aadA_tsPurple

Plasmid p207_psbH_aadA_tsPurple containing both flanks for homologous recombination within the *psbH* locus, *aadA* expression cassette and *tsPurple* was constructed. The plasmid was created to function as an intermediate during level 2 assembly to reduce the number of different parts that need to be combined in a single assembly.

Figure 3.20 show a scheme of how the plasmid was constructed, as well as results from restriction digestion by BsaI. *AmpR*, *psbH* homology flanks and the *aada* gene were amplified from a level 2 construct, provided by PhD candidate M. Fages-Laurtad. Outward flanking BsaI recognition and cleavage sites were added during the PCR. *tsPurple* was added through PNK ligation to both ends of the PCR product, simultaneously closing the plasmid. Correct assembly was identified by restriction digestion and sequencing. All restricted samples showed bands correlating with expected values of 5861 bp and 975 bp, indicating correct assembly, which was confirmed by sequencing.

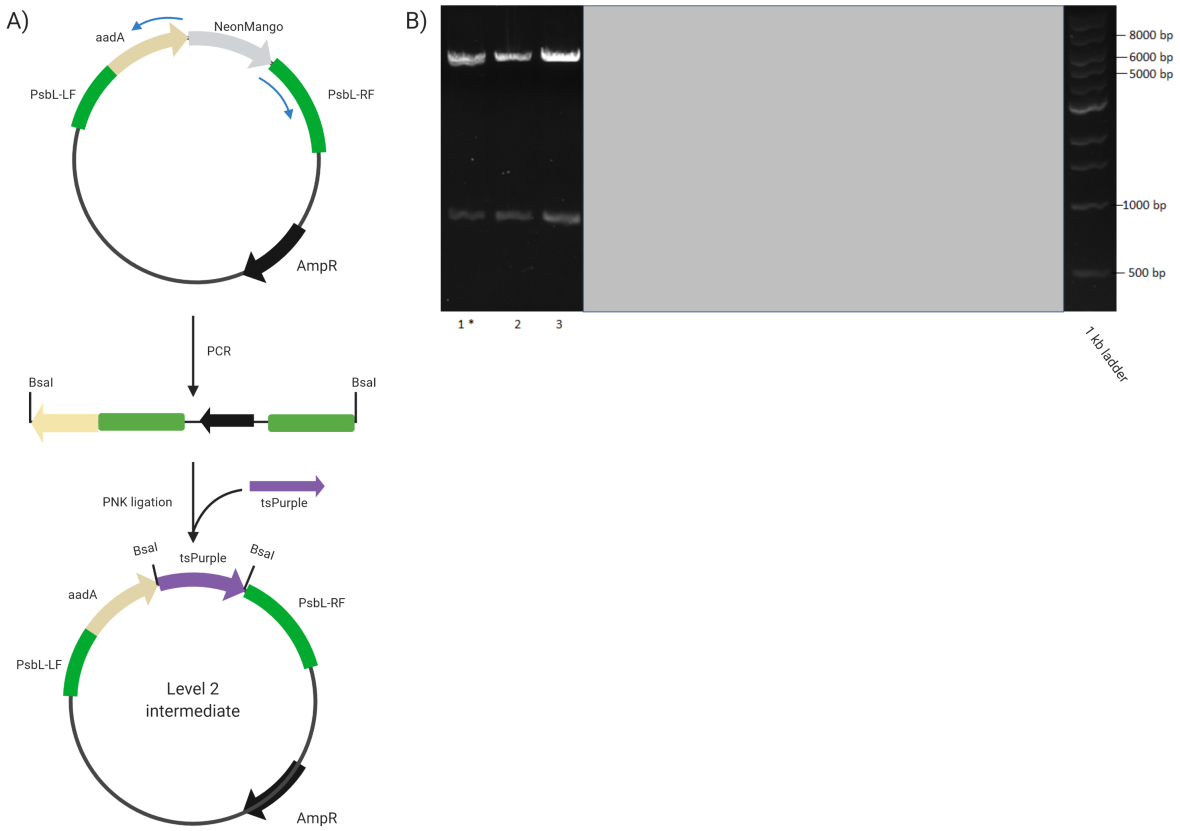


Figure 3.20: A) *aadA* expression unit, and two *psbH* flanks were amplified from a level 2 construct through PCR, using primers illustrated with blue arrows. *tsPurple* was attached by PNK ligation. B) Verification was carried out by restriction digestion utilizing BsaI followed by 0.8% agarose gel electrophoresis (120V, 60 min). All three samples show DNA fragments aligning with expected value of 975 bp and 5861 bp. A grey box cover unrelated samples on the same gel. Lanes marked with a star (*) contain plasmid with correct assembly confirmed by sequencing. 1kb DNA Ladder (NEB) was utilized as standard.

3.6 Assembly of Level 2 Constructs for Codon Reassignment

Both flanks for homologous recombination, selection cassette, reporter cassette, as well as the expression unit for aaRS and tRNA were combined on level 2 constructs. Illustrations of the plasmids are given in Figure 3.21. Two different Golden Gate Assembly approaches were tested simultaneously. One assembly utilized all six level 1 constructs and pSS191.Lv0/2_tsPurple. The second approach utilized the level 2 intermediate (p207_psbH_aadA_tsPurple) along with level 1 constructs containing reporter, aaRS and tRNA cassette, combining only four different units.

Expression units for aaRS were assembled together with the expression unit of their cognate tRNA, with the exception of *Mm*PylRS-AF. This aaRS was assembled with $Mb\text{tRNA}^{Pyl}_{UCA}$, based on previous research on compatibility between the aaRS/tRNA pair. Furthermore, *mVenusOpal* expression unit was included in constructs containing $\text{PylRS/tRNA}^{Pyl}_{UCA}$, as the tRNA contain an opal anticodon. *mVenusAmber* expression unit was combined with $Mj\text{TyrRS/tRNA}^{Tyr}_{CUA}$ due to presence of amber anticodon in this

tRNA. However, assembly of p206_psbH_MjTyrRS_MjtRNA_mVenusAmber was not confirmed.

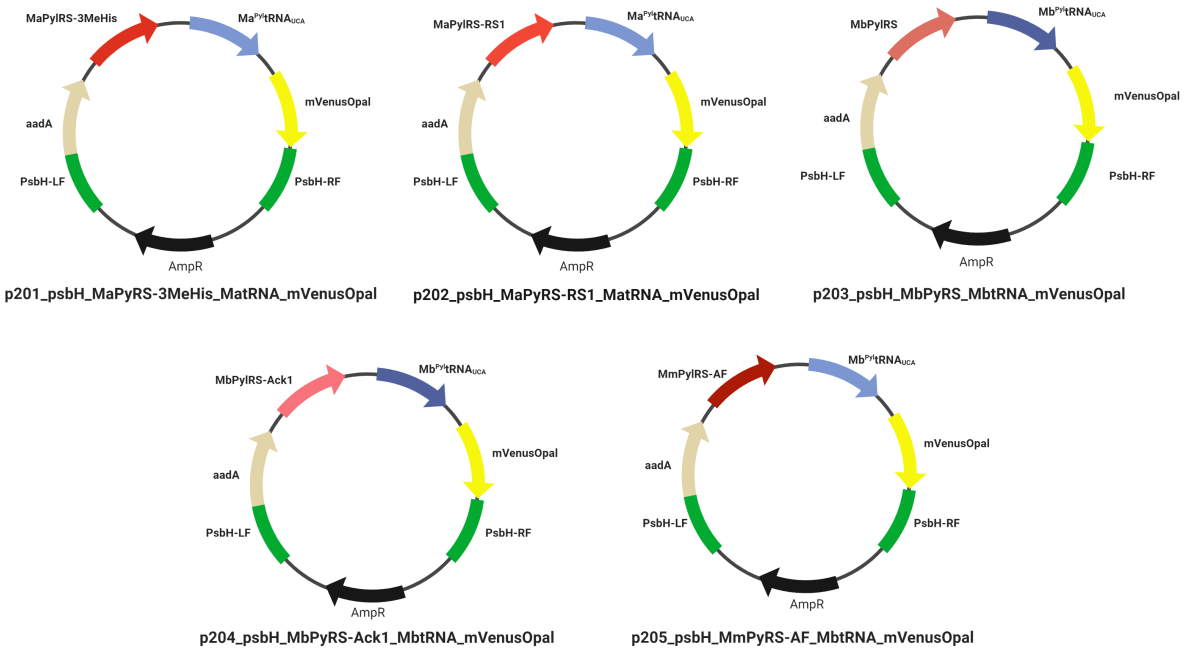


Figure 3.21: Level 2 constructs for gene delivery of pyrrolysyl-tRNA synthetase (PylRS) or tyrolysyl-tRNA synthetases (TyrRS) and their cognate tRNA to chloroplast DNA of *C. reinhardtii*. Each construct contains *aadA* for spectinomycin/streptomycin resistance and a reporter gene *mVenus* containing an additional amber (mVenusAmber) or opal (mVenusOpal) stop codon. Inserts are flanked *psbH* right flank (psbH-RF) and left flank (psbH-LF) for homologous recombination. *AmpR*, introduced as part of the backbone confer ampicillin resistance.

Figure 3.22 shows growth of *E. coli* after transformation with level 2 constructs of assembly using seven parts. Figure 3.23 display growth after assembly utilizing the intermediate. As can be observed from these plates, the assembly of seven units, resulted in what seems to be prominent growth of purple *tsPurple+ E. coli* colonies. Observation of bacterial growth after assembly of four units indicates dominant growth of white *tsPurple- E. coli* colonies. The exception is growth following cloning of the level 2 containing MjTyrRS-NitroTyr/tRNA^{Tyr} assembled using all seven units and the level 2 containing *Ma*PylRS-3MeHis/tRNA^{Pyl} assembled from the intermediate. Here, dominant growth was opposite of the described trend.

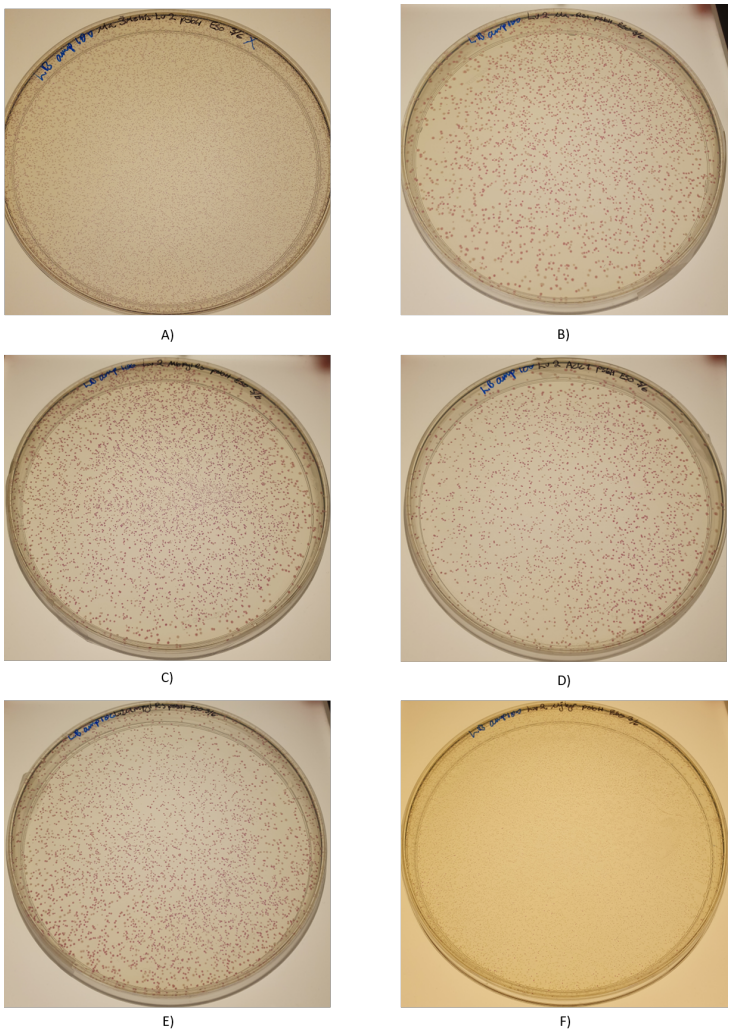


Figure 3.22: Growth of *E. coli* on LB agar supplied with 100 µg/ml ampicillin after transformation with level 2 constructs. Golden Gate Assembly was performed using the seven different parts. A-E show dominant growth of purple *tsPurple+* strains, while F) contain dominant growth of white *tsPurple-* strains. *E. coli* were transformed with the following products:
p001_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal (A),
p002_psbH_MaPylRS-RS1_MatRNA_mVenusOpal (B),
p003_psbH_MbPylRS_MbtRNA_mVenusOpal (C),
p004_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal (D),
p005_psbH_MmPylRS-AF_MbtRNA_mVenusOpal (E),
p006_psbH_MjTyrRS-NitroTyr_MjtRNA_mVenusAmber (F).

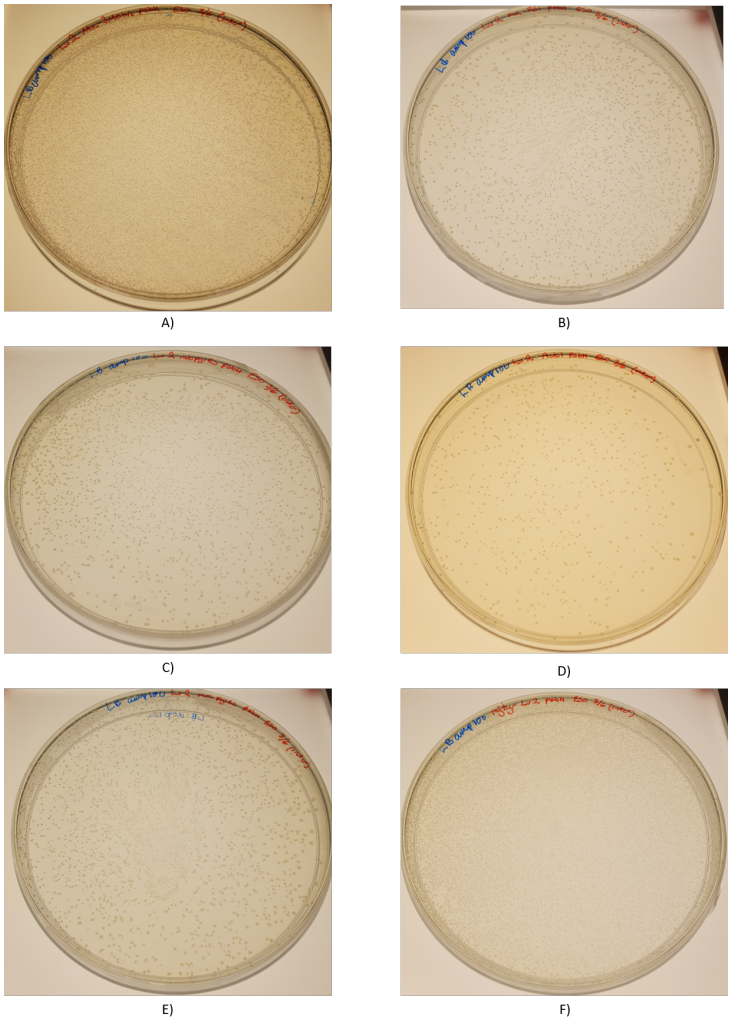


Figure 3.23: Growth of *E. coli* on LB agar supplied with 100 µg/ml ampicillin after transformation with level 2 constructs. Golden Gate Assembly was performed using a level 2 intermediate (p207_psbH_aadA_tsPurple). A) contain dominant growth of purple *tsPurple+* strains, while B-F dominant growth of white *tsPurple-* strains. *E. coli* were transformed with the following products: p001_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal (A),
p002_psbH_MaPylRS-RS1_MatRNA_mVenusOpal (B),
p003_psbH_MbPylRS_MbtRNA_mVenusOpal (C),
p004_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal (D),
p005_psbH_MmPylRS-AF_MbtRNA_mVenusOpal (E),
p006_psbH_MjTyrRS-NitroTyr_MjtRNA_mVenusAmber (F).

Based on results from colony PCR, assembly using the level 2 intermediate appeared to have been completed for all but p206_psbH_MjTyr-NitroTyr_MjtRNA_mVenusAmber. Results form analysis by colony PCR and gel electrophoresis of level 2 constructs, assembled using the intermediate, are given in Figure 3.24, with PCR product sizes listed in Table 3.4. Three different regions (x, y and z) were amplified as illustrated in Figure 3.24A. Plasmids where all three regions were amplified, were identified for all constructs, expect p206_psbH_MjTyr-NitroTyr_MjtRNA_mVenusAmber. Here amplification of region x failed, while region z gave inconclusive results. Amplification of all three regions indicates a successful assembly. This is due to directionality during assembly, where all parts are required to be present for a complete plasmid assembly.

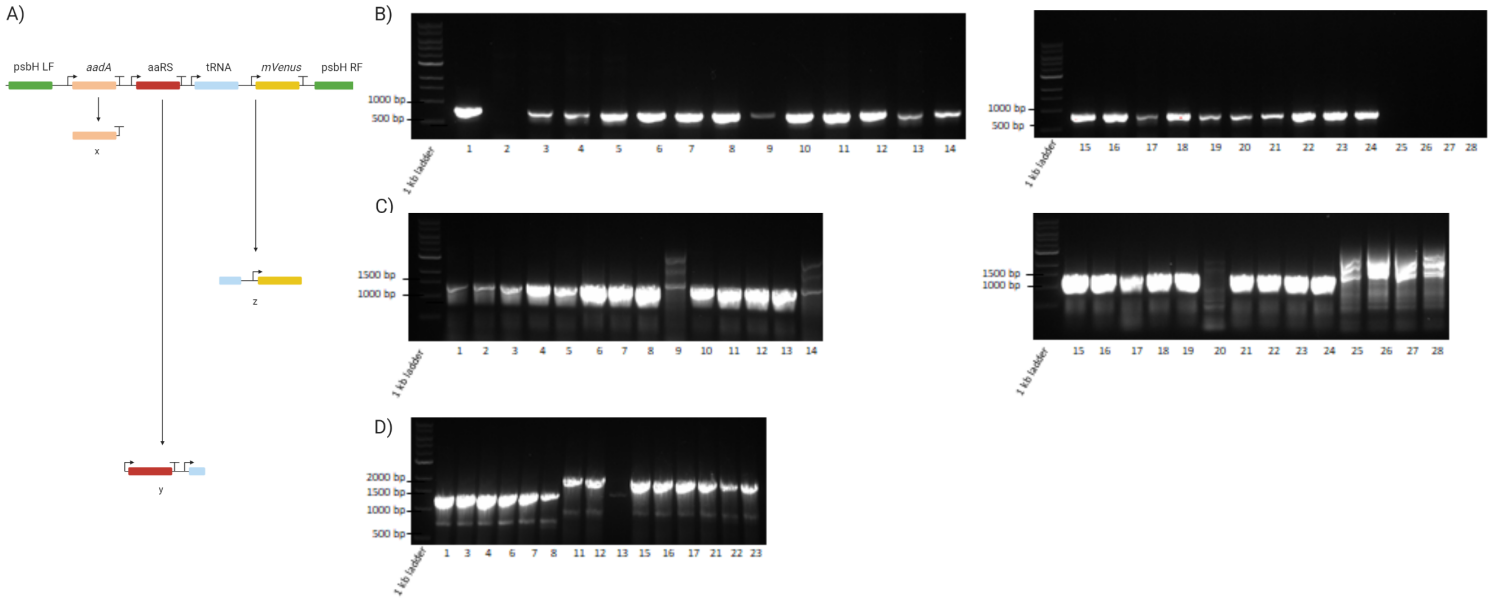


Figure 3.24: Analysis by colony PCR and gel electrophoresis (120V, 30min). DNA assembly was made using the level 2 intermediate. Three different regions were amplified (A); x, y, z. Gel electrophoresis of region x (B), z (C) were first conducted. Then region y (D) was analysed for plasmids showing positive results for x and z amplification. Each gel is loaded with 1kb DNA Ladder (NEB). and PCR product from p201_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal (1-5), p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal (6-10), p205_psbH_MmPylRS-AF_MbpyltRNA_mVenusOpal (11-14), p203_psbH_MbPylRS_MbtRNA_mVenusOpal (15-19), p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal (20-24) and p206_MjTyrRS-NitroTyr_MjtRNA_mVenusAmber (25-28).

Table 3.4: Size of colony PCR product of region x, z and y in each level 2 construct. x and z are of equal size for all constructs, while y vary.

Region	Construct	Size [bp]
x	All	801
z	All	1316
y	p001_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal	1308
	p002_psbH_MaPylRS-RS1_MatRNA_mVenusOpal	1308
	p003_psbH_MbPylRS_MbtRNA_mVenusOpal	1740
	p004_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal	1740
	p005_psbH_MmPylRS-AF_MbtRNA_mVenusOpal	1845
	p006_psbH_MjTyr-NitroTyr_MjtRNA_mVenusAmber	1401

Figure 3.25 show gel electrophoresis of colony PCR products from the seven pieced DNA assembly. By comparison to expected values in Table 3.4, assembly of p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal (lane 6), p205_psbH_MmPylRS-AF_MbpyltRNA_mVenusOpal (lane 11), p203_psbH_MbPylRS_MbtRNA_mVenusOpal (lane 17) and p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal (lane 22-24) were identified. Results from lane 26-30 Figure 3.25C were unreadable, which appeared to be caused by problems with the agarose gel. Even so, *aada* (Figure 3.25B) was not amplified in sample 26-30 either, and therefor the samples were discarded regardless of amplification of *mVenusAmber*.

p201_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal,
p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal,
p205_psbH_MmPylRS-AF_MbpyltRNA_mVenusOpal,
p203_psbH_MbPylRS_MbtRNA_mVenusOpal,
p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal and
p205_psbH_MmPylRS-AF_MbtRNA_mVenusOpal from the assembly utilizing the intermediate were isolated. These samples were stored, and ready to be utilized for future aspects building on the project.

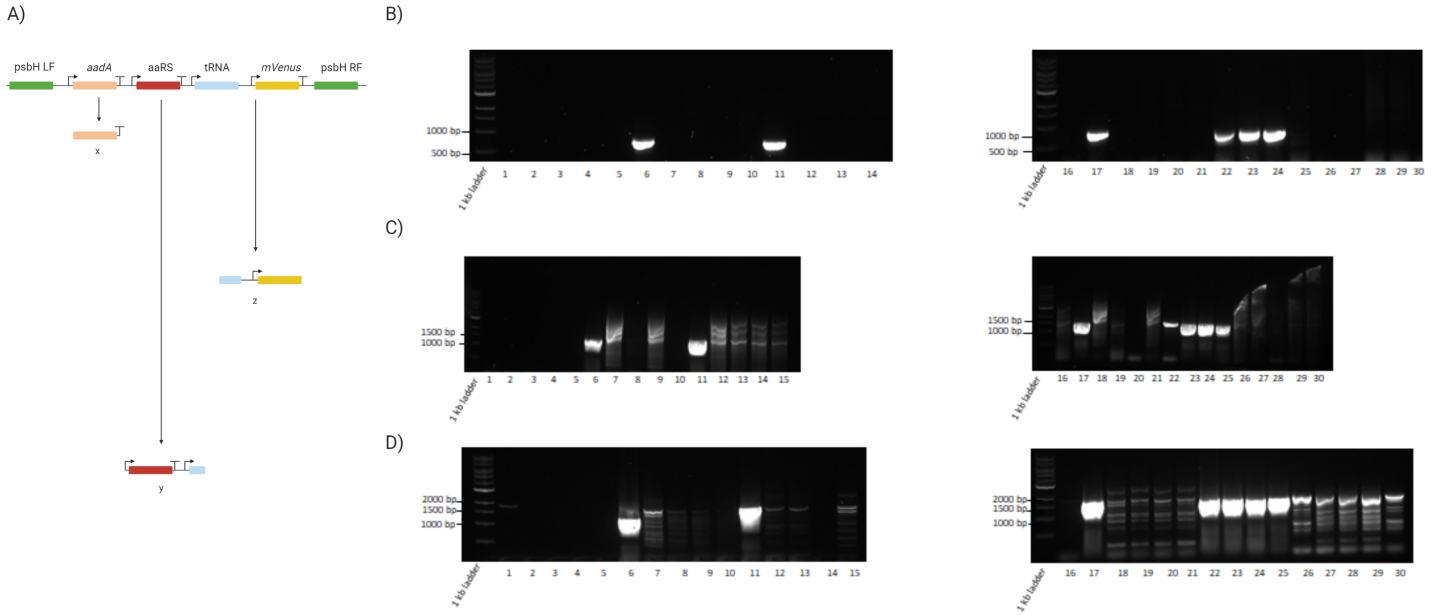


Figure 3.25: Analysis by colony PCR and 0.8% agarose gel electrophoresis (120V, 30min). DNA assembly was made using seven different parts. Three different regions were amplified (A); x, y and z. B) Gel electrophoresis of amplified region x. C) Amplification of region z, problems with gel affected lane 26-30. D) Amplification of region y. Each gel is loaded with 1 kb DNA Ladder (NEB) and PCR product from p201_psbH_MaPylRS-3MeHis_MatRNA_mVenusOpal (1-5), p202_psbH_MaPylRS-RS1_MatRNA_mVenusOpal (6-10), p205_psbH_MmPylRS-AF_MbpyltRNA_mVenusOpal (11-15), p203_psbH_MbPylRS_MbtRNA_mVenusOpal (15-20), p204_psbH_MbPylRS-Ack1_MbtRNA_mVenusOpal (21-25) and p206_MjTyrRS-NitroTyr_MjtRNA_mVenusAmber (26-30).

4 Discussion

The aim of the thesis was to establishes the molecular tools to identify and utilize orthogonal translation systems for the incorporation of ncAA during protein synthesis in the chloroplast of the microalgae *Chlamydomonas reinhardtii*. This was achieved by develop vectors containing genes of different aminoacyl tRNA synthetases (aaRS) together with their cognate tRNA for codon reassignment of amber and opal codon in the chloroplast of *C. reinhardtii*. These codons were to be reassigned from promoting translation termination, to the incorporation of different non-canonical amino acids (ncAA). Additionally, the vectors contain a selectable marker to drive selection towards homoplasmy, homologous flanking sequences for site-specific recombination, and a reporter gene to evaluate if the systems function. Two loci were to be targeted for homologous recombination, *psbL* and *psbH*.

4.1 Design of Expression Units for High Expression Rate

Every coding sequence (CDS) was combined with the promoter and 5'untranslated region (UTR) of a highly expressed chloroplast gene to provide a high and steady expression rate for each gene. Each was chosen due to the previously reported effect in transgenic expression. The promoter of p16s was used for the *mVenus* genes, the *atpA* promoter for *aadA*, while the promoter of *psbD* was chosen for each aaRS. As previously stated, the *psbA* promoter has reported a high expression rate and been used as a promoter for transgenic expression in D1-deficient strains [6]. However, the aim is to introduce the system in wild type (wt) *C. reinhardtii*. Therefore, this enhanced expression rate was not taken into consideration. The promoter of *psbB* and *rbcl* were amplified and assembled on level 0 constructs to function as back-ups. For future work, if the expression rate of the generated gene cassettes fails or is low, a possible solution could be to re-assemble the constructs with the remaining promoters. Previous research indicates that choice of terminator/3'UTR for transgenes has little effect on mRNA and protein accumulation in the chloroplast, as long as a terminator/3'UTR is present [40]. The reasoning for using different 3'UTR for the different expression units was, therefore, to prevent undesired homologous recombination between these parts.

Before this project, no research was found on utilizing orthogonal aaRS/tRNA for codon reassignment in the chloroplast of *C. reinahardtii*. Therefore, multiple aaRS/tRNA pairs were prepared to increase the chances of identifying pairs that are orthogonal towards the algae' endogenous aaRS/tRNA. These aaRS/tRNA pairs were chosen based on the phylogenetic distance between green algae and the archaea from which the aaRS/tRNA was isolated. Additionally, each aaRS/tRNA pair have been proven orthogonal towards the endogenous aaRS/tRNA of *E. coli*. Due to the similarities between prokaryotes and the chloroplast, it was hypothesized that the chosen aaRS/tRNA and endogenous aaRS/tRNA would be orthogonal. Furthermore, *Methanosarcina mazei* pyrrolysyl-tRNA synthetase AF (*Mm*PylRS-AF)/tRNA^{Pyl} are orthogonal towards other eukaryotic cells as well, thereby strengthening it's potential for orthogonality in the algal chloroplast.

Each of the PylRS are mutually orthogonal towards *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase NitroTyr (*Mj*TyrRS-NitroTyr)/^{Tyr}tRNA. Additionally, *Mm*PylRS-AF/ tRNA^{Pyl} is mutually orthogonal towards both variants of *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase (*Ma*PylRS)/tRNA^{Pyl} with the altered *Ma*tRNA^{Pyl}. There-

fore, these aaRS/tRNA could be used simultaneously, incorporating two different ncAA. Although, a requirement for the two PylRS would be a reassignment of two different codons. This could be achieved by tRNAs containing different anticodons. *mVenusAO* containing both an amber and opal codon was prepared to be utilized as a reporter for the eventual combination of opal and amber reassignment.

4.2 Amber and Opal Codon Reassignment

The plastome of *C. reinhardtii* do not contain the opal codon, but utilize the amber codon to a certain extent [5]. Two of these amber codons are directly followed by an ochre codon. Those, reassigning the amber codon to incorporate ncAA could potentially only introduce one additional amino acid (AA) before the translation is terminated. The consequences, and perhaps cytotoxicity, of such alterations, are unknown. A third amber codon lies within the *psbL* locus. A sub-goal was to introduce the orthogonal translation system directly after this amber codon. Without addressing this codon, and still going forward with the amber codon reassignment, the gene product could potentially attain eleven additional AA with unknown after-effects. This is an event that could take place if the expression units are only inserted in the *psbH* locus. *Mj*TyrRS-NitroTyr/^{Tyr}tRNA was to be utilized for amber codon reassignment. However, a level 2 assembly with all required parts was not achieved. The other PylRS/tRNA pairs were designed for opal codon reassignment. Therefore, moving forward with the opal codon reassignment, the amber codon within the *psbL* locus is not of importance. As no opal codon are found in the chloroplast DNA (cpDNA), this stop codon could be freely reassigned without affecting endogenous protein synthesis. For future directions, if the amber codon is to be reassigned, the mentioned *psbL* amber codons should be replaced by ochre codons, or otherwise dealt with.

4.3 Evaluation of Plasmid Construction

All pUC8, level 0 and level 1 constructs required to construct the genetic vectors were successfully developed, including those for amber codon reassignment. Some assemblies were not completed as planned. These include *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase (*Mb*PylRS), *Mm*PylRS and *Mj*TyrRS on pUC8 backbones, as well as level 0 constructs containing *Ma*tRNA and *Mj*tRNA. Nonetheless, ways were found around these obstacles, allowing for assembly of the final plasmids. *Mb*PylRS and *Mj*TyrRS were assembled on level 0 constructs rather than pUC8, while *Mm*PylRS were assembled directly on level 1. *Ma*tRNA and *Mj*tRNA were PCR amplifed onto a level 1 construct instead of level 0.

The aim of developing a genetic vector containing required genes for amber codon reassignment was not achieved. Why assembly of a level 2 construct containing *Mj*TyrRS-NitroTyr/tRNA$^{Tyr}_{CUA}$ failed was not identified. No assembled p206_psbH_MjTyrRS-NitroTyr_MjtRNA_mVenusAmber was detected by colony PCR in any *E. coli* colonies following cloning. Neither from the assembly using all necessary level 1 constructs, nor utilizing the level 2 intermediate. It is uncertain why this assembly failed while the other level 2 assemblies functioned. Three parts separate the *Mj*TylRS-NitroTyr/tRNA$^{Tyr}_{CUA}$ construct from the rest: *Mj*TylRS-NitroTyr, tRNA$^{Tyr}_{CUA}$ and mVenusAmber. The correct sequence of these three genes was confirmed after level 1 assembly. Although, a possible explanation could be that these samples contain contamination interfering with the level

2 assembly. The analysis of the assembled plasmid confirmed presence of the *aadA* and *mVenus* expression units. However, the analysis of the *Mj*TyrRS was inconclusive. Due to the design of the Start-Stop Assembly, a complete plasmid should not be able to assemble without all parts present. Therefor, if there is a problem with recognition or cleavage sites for the restriction endonuclease in one part, it could affect the whole assembly. Plasmid construction and colony PCR were re-attempted several times, without identifying complete assembly of the construct. In case of false-negative results by the colony PCR, samples could also be further screened by restriction digestion and sequencing to evaluate the constructs. This was not conducted due to time constraints.

The five different vectors for opal codon reassignment, targeting the *psbH* locus, were all successfully constructed. *E. coli* grew in the presence of ampicillin after transformation with level 2 constructs, suggesting the presence of *ampR*. Selection cassette, aaRS, tRNA and reporter cassette were all confirmed present by colony PCR. No analysis of *psbH* regions was conducted. However, as ampicillin resistance and *aadA* expression unit were confirmed present, it is likely to believe that the entire intermediate backbone was present, which includes the right and left *psbH* flanks. Although the main parts were analyzed by colony PCR, sequencing of the constructs should be performed before transformation in *C. Reinhardtii* to assure correct construct assembly.

4.4 Utilization of the Start-Stop Assembly

Utilizing the principles of the Start-Stop Assembly, functionally scarless genes were constructed. By avoiding scarring, potential problems scarring could pose on sensitive DNA related to transcription, translation and stability was hindered. During the multi-part DNA assembly of these genes, a total of seven units were combined. These units included four expression units, two regions for homologous recombination and the acceptor plasmid. Previous attempts at using Start-Stop Assembly report the use of up to five expression units during level 2 construct assembly [51]. That being the case, there existed an uncertainty if the level 2 construct assembly with seven units could be accomplished. Therefore, two approaches were tested out simultaneously. These consisted of assembling all seven units, or using a intermediate with the *psbH* right flank, *psbH* left flank and the *aadA* gene. As a result, the number of units used for the assembly was reduced to four, bringing it below the maximum number of parts that have been reported used.

The motivation for testing two different approaches was to assemble the constructs with respect to time and resources effectively. Assembled level 2 constructs were swiftly identified after assembly using the intermediate approach. Evaluating the Start-Stop Assembly was not an aim for the project. However, a notable observation was that an assembly of level 2 constructs using seven different units proved unsuccessful for two of the constructs, and what seemed to be low assembly efficiacy for the ones which were identified. A strong presence of *tsPurple+ E. coli* strains were observed after transformation and cloning, indicating failed DNA assembly of constructs. Moreover, analysis by colony PCR among the few *tsPurple- E. coli* colonies present gave for most inconclusive or negative results. However, results from colony PCR of some samples did indicate correct assembly of four out of the six constructs. Based on these observations, using an intermediate to lower the number of units seems to be a better approach for a DNA assembly of this type and size. Further analysis, testing more assemblies with equal conditions would have to be performed to draw a more specific conclusion.

The use of *tsPurple* and *LacZ* in the different acceptor plasmids provided a quick initial method to screen for assembled constructs after cloning. Several *tsPurple+* and *LacZ+ E. coli* colonies were observed after cloning of the respective plasmids. Indicating a relatively large presence of intact donor plasmid left after the DNA assembly. It is, therefore likely that time and resource consumption spent on colony PCR would be much higher if these genes had not been utilized.

4.5 Trouble Shooting During PCR Amplification and DNA Assembly

During the project, a couple time-consuming issues were met. These difficulties were all related to PCR amplification or DNA assembly. Here, the most troubling challenges are addressed, as well as how or if they were handled.

Several attempts were carried out to amplify the left flank for homologous recombination of the *psbL* locus. Firstly in an attempt to promote annealing of primer to the cpDNA, annealing temperature during PCR was varied. However, the region was still not amplified. PCR with touch-down procedures was also attempted, to span a larger region of annealing temperatures. This approach was found to be rather successful during other problematic PCR-mediated amplifications, but did not work for the *psbL* region. It was speculated whether the problem lay with the primers themselves. New primers with a slightly different annealing sequence were ordered and tested, but still gave no result. The same extract of cpDNA template was used for succesfull amplification of the right flank of *psbL* as well as the promoter of *psbB*, *psbD* and *rbcl*. Nonetheless, a possible explanation could be that the region of interest was inaccessible. The DNA structure and folding in the area of *psbL* was not researched. However, it could potentially be the reason why amplification of *psbL* failed, while other regions were successfully amplified. Generation of secondary structures in the DNA template is know to cause difficulties during PCR amplification. Previous research show that dimethyl sulfoxide can help enhance amplification specificity by relieving secondary structures, especially in GC-rich templates [70]. The left flank only have a GC content of 32%. However, dimethyl sulfoxide was still added to the PCR reaction to try to amplify the region. Although, it did not change the outcome. Due to time constraints, no further attempts to amplify the *psbL* region were attempted. The focus was therefore shifted to focusing singly on homologous recombination within the *psbH* locus.

Out of the four ordered aaRS sequences (*Ma*PylRS-3Mehis, *Mb*PylRS. *Mm*PylRS-AF and *Mj*TyrRS-NitroTyr), only *Ma*PylRS-3MeHis was assembled onto a plasmid with a pUC8 backbone. Several attempts at assembling the backbones and the aaRS using Gibson assembly were carried out. This involved re-amplification of the inserts, but did not change the outcome. pUC8 backbones were tested with other inserts and worked, suggesting the problem lie with the aaRS. As previously stated, colony PCR analysis indicated a product size that aligns with that of self-ligated pUC8. This would also explain why *E. coli* could grow in the presence of ampicillin after transformation, even though the CDS of the aaRS was not detected. Each aaRS were synthesized with BioBrick prefix and suffix to allow assembly with the pUC8 backbone. Hence, PCR amplification could be conducted using the same primers to save time and resources. As assembly on pUC8 failed, different ways were found to assemble storage vectors. Primers annealing to *Mb*PylRS and *Mj*TyrRS as well as introducing BsaI and SapI restriction sites with

appropriate cleavage sites were available. These CDS were therefor assembled at level 0 constructs. All CDS were initially synthesized with SapI and the correct cleavage sites for Start-Stop Assembly. Similar primers for *Mm*PylRS were not available. The CDS of *Mm*PylRS was, therefore, assembled directly on a level 1 construct, instead of spending time and resources on new primers needed for amplification and preparation for assembly on level 0.

Through analysis of colony PCR products, some results showed second prominent bands of approximately 700 bp, in addition to the expected result. Such unexpected PCR products were found in p001_MbPylRS, p002_MbPylRS-Ack1, p007_mVenusAmber and p009_mVenusAO. For the aaRS, these bands were found in every sample, which also contained a product of the correct size. A possible explanation is that the primers used during PCR had a strong affinity for more than one region, thereby amplifying two regions. Another explanation could be the presence of additional plasmids, to which the primers anneal. Although clearly visible, the $\sim 700 - 750$bp bands were not as prominent as the expected band. No additional PCR products of comparable concentration were detected once the inserts were transferred from level 0 to level 1. Therefore, the additional bands at level 0 were not further analyzed. All these additional bands were estimated to fragment size, which correlates well with that of pSS191.Lv0/2_tsPurple, where *tsPurple*, have been excised and the remaining part self-ligated. However, the plasmid was designed with non-palindromic fusion sites (X and Y), which should have prevented self-ligation. If these bands are caused by self-ligation, point-mutations or unexpected cleavage patterns could be possible explanations. That being said, no conclusion can be drawn from the obtained results. To further analyze these DNA fragments, they could have been extracted from the gel and attempted sequenced. Since correct constructs were identified, such an analysis was not prioritized.

4.6 Future Perspectives

Throughout the thesis, assemblies required for codon reassignment of the opal codon towards ncAA incroporation were completed. The next step is to transform *C. reinhardtii* with the vectors and. Then test if the system functions, evaluating if the different PylRS/tRNA^{Pyl} pairs are orthogonal towards the endogenous chloroplast aaRS/tRNA of *C. reinhardtii*.

Here, a possible scenario for establishing and evaluating the *C. reinhardtii* strains is outlined. Firstly, attempts to assemble p206_MjTyr-NitroTyr_MjTyr-tRNA_mVenus-Amber_psbH could have been re-attempted. In addition to the analysis by colony PCR, each level 2 construct should be sequenced in order to be certain of the correct assembly. The next step would be to test the system in *C. reinhardtii* as outlined in Figure 4.1. The first step would be to introduce the gene vectors into the chloroplast of *C. reinhardtii* to allow for homologous recombination in *psbH*. Three techniques are commonly used for transformation in green algae. These are glass bead assisted delivery, electrophoration and biolistic delivery, which are described in Section 1.9. The gene gun-mediated approach could be choosen as it is a more commonly used technique for transformation within the chloroplast. However, this project's research group has been working on an improved protocol for homologous recombination with cpDNA using electrophoration. If successful, this approach could also prove a viable choice.

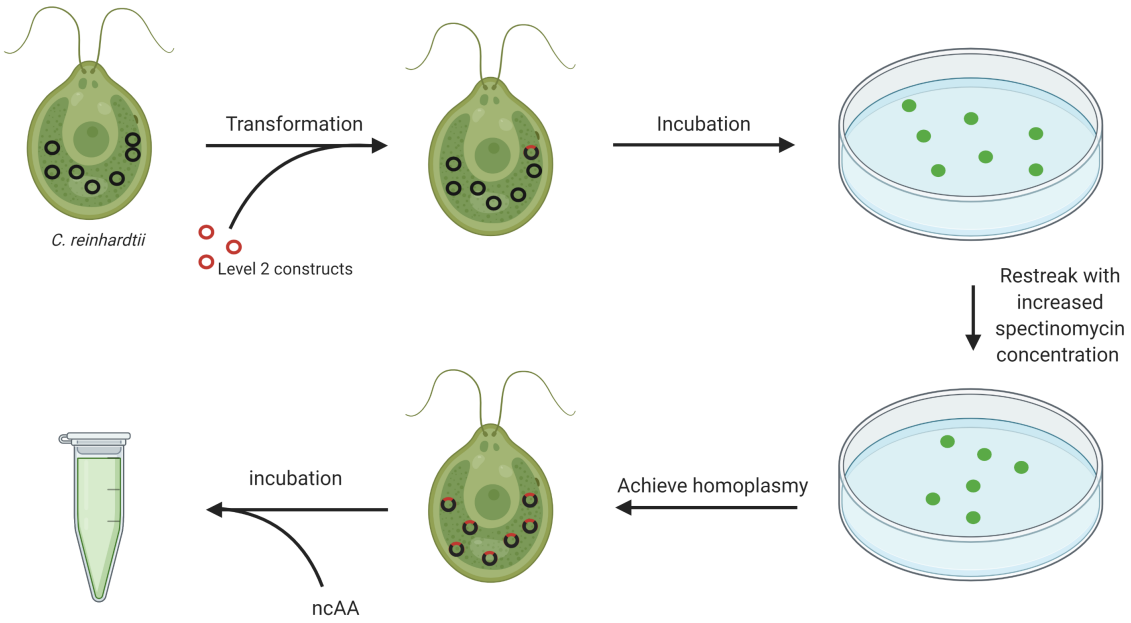


Figure 4.1: Development of homoplasmic *C. reinhardtii* strains through homologouse recombination between *psbH* in the cpDNA and level 2 constructs. Increasing selective pressure to achieve homoplasmic strains then incubation in liquid medium supplied with appropriate non-canonical amino acids (ncAA).

Following transformation *C. reinhardtii* would have been grown on agar plates supplied with spectinomycin for selective growth of recombinant *C. reinhardtii*. To obtain homoplasmic strains, restreaking while gradually increasing spectinomycin concentration could be performed. By increasing antibiotic concentration, the selective pressure would increase, favoring the growth of organisms with a higher copy number of recombinant plastomes. To test for homoplasmy, colony PCR amplifying *psbH* could be performed. Recombinant cpDNA would contain a larger *psbH* loci than the wild-type. It is believed that that homoplasmy would be indicated by one band due to the presence of only the recombinant cpDNA. Heteroplasmy is thought to show two bands, both the shorter wt-cpDNA and the longer recombinant cpDNA. After achieving homoplasmy, or as close to homoplasmy as possible, the algae could be incubated in liquid medium supplied with ncAA. What type of aaRS/tRNA pairs which are compatible with which ncAA are listed in Table 1.2. It was anticipated that this step could prove one of the larger bottle-necks of the project. It is uncertain whether *C. reinhardtii* is fully capable of absorbing the ncAA from the medium. If not, a possible solution could be to develop *C. reinhardtii* strains with biosynthetic pathways for the development of said ncAA. Utilizing such a strain would abolish the need for supplying ncAAs to the medium, as the organisms are capable of synthesizing the ncAA themselves.

The reporter gene, *mVenus*, was included to identify the successful application of the exogenous aaRS/tRNA pair. If the orthogonal system worked, mVenus would be fully synthesized with an additional ncAA as the 3rd (mVenusAmber) or 151st (mVenusOpal) AA. If the codon reassignment of amber or opal codon did not work, fully functional mVenus would not be synthesized. Hence, after incubation in the presence of ncAA, fluorescence could be measured to evaluate the codon reassignment. A control consisting of only wt-*C. reinhardtii*, would also have to be included to correct for potential background

noise. Additionally, mass spectrometry could also be conducted to screen for mVenus. For this reason, histidine-tags were added to mVenus for easier identification.

The *aadA* gene was included as the selectable marker, due to an established use in chloroplast engineering. A possible alteration to the system could be to exchange the antibiotic resistance marker with a photosynthetic gene. As stated in the introduction, selectable markers could be used to re-introduce functional genes essential for the photosystem in mutant strains of *C. reinhardtii* where such genes have been removed or otherwise inactivated. As an alternative to *aadA*, a functional version of *psbH* could potentially be used as a selectable marker for phototrophic restoration. This gene have previously been used as a selectable marker in combination with Δ*psbH C. reinhardtii* strain TN72 [71]. Hence, risk of spreading additional antibiotic resistance during cloning in bacteria could be reduced, compared to when utilizing antibiotic resistance markers.

If the system proves successful, codon reassignment to facilitate ncAA incorporation could prove useful in several ways. By introducing the reassigned codons in endogenous genes, ncAA with distinct properties could be included, enhancing or introducing novel protein functionalities. ncAA could also be used to probe the proteins. Such site-spesific probes could prove useful for analysing protein structure and function [18]. Additionally, developing a method for synthesis of proteins containing ncAA could help elevate the microalgae as a useful production platforms for recombinant protein expression.

5 Conclusion

In this study, molecular tools to analyze orthogonal translation systems within the chloroplast of *C. reinhardtii* have been developed. If proven orthogonal, the system can also be utilized for codon reassignment towards incorporating non-canonical amino acids during translation.

Functionally scarless vectors for chloroplast engineering were constructed utilizing the Start-Stop assembly. Start codons were used as fusion sites between promoter/5'untranslated region (UTR) and coding sequence (CDS), while ochre stop codons were utilized as fusion sites between coding sequences and 3'UTRs. Hence, diminishing the inclusion of genetic scars unrelated to the assembled genes.

These genetic vectors were constructed to introduce genes into the chloroplast DNA, which gene product could potentially reassign opal and amber codons towards the incorporation of non-canonical amino acids. Different genes for Pyrrolysyl tRNA synthetases (PylRS) were genetically combined with genes for the corresponding tRNA. A total of five PylRS were utilized, originating from three different archaea. Two variants of *Methanomethylophilus alvus* PylRS (*Ma*PylRS-3MeHis and *Ma*PylRS-RS1), wild-type and one variant of *Methanosarcina barkeri* PylRS (*Mb*PylRS and *Mb*PylRS-Ack1) and a variant of *Methanosarcina mazei* PylRS (*Mm*PylRS-AF) were utilized. A *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase NitroTyr gene was attempted combined with its cognate amber suppressor tRNA gene. Albeit, the final vector assembly was not accomplished. However, all required expression units were constructed on separate vectors.

Each vector was designed with a selection cassette conferring spectinomycin and streptomycin resistance, to allow for the selection of recombinant organisms and drive the

population towards homoplasmy of recombinant chloroplast DNA. To test the system's functionality, a variant of fluorescent *mVenus* was used as a reporter gene. The gene was engineered to inhabit an additional opal codon. Hence, fluorescence detection could be utilized to evaluate the establishment of orthogonality and codon reassignment. All vectors were designed with flanking regions facilitating homologous recombination within the *psbH* locus in the chloroplast of *C. reinhardtii*. Each CDS were coupled to the promoter, 5'UTR and 3'UTR of highly expressed chloroplast genes to promote high expression rate and enhanced mRNA stability.

The project was concluded with establishing the five molecular tools for opal codon reassignment and expansion of the genetic code utilizing a set of different PylRS/tRNA$^{Pyl}_{UCA}$ variants. Future aspects include transforming *C. reinhardtii* and evaluating the system's functionality. If proven successful, the system can be utilized to probe and thereby analyze proteins, strengthen sought properties or develop novel protein functionalities. All these attributes contribute to further establishing the microalga as an attractive site for recombinant protein production.

References

[1] Elizabeth Specht, Shigeki Miyake-Stoner, and Stephen Mayfield. "Micro-algae come of age as a platform for recombinant protein production". In: *Biotechnology Letters* 32.10 (June 2010), pp. 1373–1383. DOI: 10.1007/s10529-010-0326-5. URL: https://doi.org/10.1007/s10529-010-0326-5.

[2] Miller Tran et al. "Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts". In: *Biotechnology and Bioengineering* (2009), n/a–n/a. DOI: 10.1002/bit.22446. URL: https://doi.org/10.1002/bit.22446.

[3] Muhammad Imran Khan, Jin Hyuk Shin, and Jong Deog Kim. "The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products". In: *Microbial Cell Factories* 17.1 (Mar. 2018). DOI: 10.1186/s12934-018-0879-x. URL: https://doi.org/10.1186/s12934-018-0879-x.

[4] Henry N. Taunt, Laura Stoffels, and Saul Purton. "Green biologics: The algal chloroplast as a platform for making biopharmaceuticals". In: *Bioengineered* 9.1 (Sept. 2017), pp. 48–54. DOI: 10.1080/21655979.2017.1377867. URL: https://doi.org/10.1080/21655979.2017.1377867.

[5] Rosanna E. B. Young and Saul Purton. "Codon reassignment to facilitate genetic engineering and biocontainment in the chloroplast ofChlamydomonas reinhardtii". In: *Plant Biotechnology Journal* 14.5 (Oct. 2015), pp. 1251–1260. DOI: 10.1111/pbi.12490. URL: https://doi.org/10.1111/pbi.12490.

[6] Beth A. Rasala et al. "Improved heterologous protein expression in the chloroplast of Chlamydomonas reinhardtii through promoter and 5' untranslated region optimization". In: *Plant Biotechnology Journal* 9.6 (May 2011), pp. 674–683. DOI: 10.1111/j.1467-7652.2011.00620.x. URL: https://doi.org/10.1111/j.1467-7652.2011.00620.x.

[7] Ralph Bock and Heribert Warzecha. "Solar-powered factories for new vaccines and antibiotics". In: *Trends in Biotechnology* 28.5 (May 2010), pp. 246–252. DOI: 10.1016/j.tibtech.2010.01.006. URL: https://doi.org/10.1016/j.tibtech.2010.01.006.

[8] Raylene Ramos Moura, Luciana Magalhães Melo, and Vicente José de Figueirêdo Freitas. "Production of recombinant proteins in milk of transgenic and non-transgenic goats". In: *Brazilian Archives of Biology and Technology* 54.5 (Oct. 2011), pp. 927–938. DOI: 10.1590/s1516-89132011000500010. URL: https://doi.org/10.1590/s1516-89132011000500010.

[9] Neus Ferrer-Miralles et al. "Microbial factories for recombinant pharmaceuticals". In: *Microbial Cell Factories* 8.1 (2009), p. 17. DOI: 10.1186/1475-2859-8-17. URL: https://doi.org/10.1186/1475-2859-8-17.

[10] Matthew J. B. Burnett and Angela C. Burnett. "Therapeutic recombinant protein production in plants: Challenges and opportunities". In: *PLANTS, PEOPLE, PLANET* 2.2 (Nov. 2019), pp. 121–132. DOI: 10.1002/ppp3.10073. URL: https://doi.org/10.1002/ppp3.10073.

[11] Takayuki Katoh, Yoshihiko Iwane, and Hiroaki Suga. "tRNA engineering for manipulating genetic code". In: *RNA Biology* 15.4-5 (Sept. 2017), pp. 453–460. DOI: 10.1080/15476286.2017.1343227. URL: https://doi.org/10.1080/15476286.2017.1343227.

[12] Tatsuo Yanagisawa et al. “Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode Nε-(o-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification”. In: *Chemistry & Biology* 15.11 (Nov. 2008), pp. 1187–1197. DOI: 10.1016/j.chembiol.2008.10.004. URL: https://doi.org/10.1016/j.chembiol.2008.10.004.

[13] Heinz Neumann, Sew Y Peak-Chew, and Jason W Chin. “Genetically encoding Nϵ-acetyllysine in recombinant proteins”. In: *Nature Chemical Biology* 4.4 (Feb. 2008), pp. 232–234. DOI: 10.1038/nchembio.73. URL: https://doi.org/10.1038/nchembio.73.

[14] Tamara L. Hendrickson, Valérie de Crécy-Lagard, and Paul Schimmel. “Incorporation of Nonnatural Amino Acids Into Proteins”. In: *Annual Review of Biochemistry* 73.1 (June 2004), pp. 147–176. DOI: 10.1146/annurev.biochem.73.012803.092429. URL: https://doi.org/10.1146/annurev.biochem.73.012803.092429.

[15] Pol Arranz-Gibert, Jaymin R. Patel, and Farren J. Isaacs. “The Role of Orthogonality in Genetic Code Expansion”. In: *Life* 9.3 (July 2019), p. 58. DOI: 10.3390/life9030058. URL: https://doi.org/10.3390/life9030058.

[16] Sergey V. Melnikov and Dieter Söll. “Aminoacyl-tRNA Synthetases and tRNAs for an Expanded Genetic Code: What Makes them Orthogonal?” In: *International Journal of Molecular Sciences* 20.8 (Apr. 2019), p. 1929. DOI: 10.3390/ijms20081929. URL: https://doi.org/10.3390/ijms20081929.

[17] Chang C. Liu and Peter G. Schultz. “Adding New Chemistries to the Genetic Code”. In: *Annual Review of Biochemistry* 79.1 (June 2010), pp. 413–444. DOI: 10.1146/annurev.biochem.052308.105824. URL: https://doi.org/10.1146/annurev.biochem.052308.105824.

[18] Kaihang Wang et al. “Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET”. In: *Nature Chemistry* 6.5 (Apr. 2014), pp. 393–403. DOI: 10.1038/nchem.1919. URL: https://doi.org/10.1038/nchem.1919.

[19] Nancye C. Clark et al. “Detection of a Streptomycin/Spectinomycin Adenylyltransferase Gene (aadA) in Enterococcus faecalis”. In: *Antimicrobial Agents and Chemotherapy* 43.1 (Jan. 1999), pp. 157–160. DOI: 10.1128/aac.43.1.157. URL: https://doi.org/10.1128/aac.43.1.157.

[20] Hsu-Ching Chen and Anastasios Melis. “Marker-free genetic engineering of the chloroplast in the green microalgaChlamydomonas reinhardtii”. In: *Plant Biotechnology Journal* 11.7 (May 2013), pp. 818–828. DOI: 10.1111/pbi.12073. URL: https://doi.org/10.1111/pbi.12073.

[21] Lola Esland, Marco Larrea-Alvarez, and Saul Purton. “Selectable Markers and Reporter Genes for Engineering the Chloroplast of Chlamydomonas reinhardtii”. In: *Biology* 7.4 (Oct. 2018), p. 46. DOI: 10.3390/biology7040046. URL: https://doi.org/10.3390/biology7040046.

[22] Gert-Jan Kremers et al. “Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius†, ‡”. In: *Biochemistry* 45.21 (May 2006), pp. 6570–6580. DOI: 10.1021/bi0516273. URL: https://doi.org/10.1021/bi0516273.

[23] James S. Italia et al. “Mutually Orthogonal Nonsense-Suppression Systems and Conjugation Chemistries for Precise Protein Labeling at up to Three Distinct Sites”. In: *Journal of the American Chemical Society* 141.15 (Mar. 2019), pp. 6204–6212. DOI: 10.1021/jacs.8b12954. URL: https://doi.org/10.1021/jacs.8b12954.

[24] Reimo Zoschke and Ralph Bock. “Chloroplast Translation: Structural and Functional Organization, Operational Control, and Regulation”. In: *The Plant Cell* 30.4 (Apr. 2018), pp. 745–770. DOI: 10.1105/tpc.18.00016. URL: https://doi.org/10.1105/tpc.18.00016.

[25] Yan Ling Joy Pang, Kiranmai Poruri, and Susan A. Martinis. “tRNA synthetase: tRNA aminoacylation and beyond”. In: *Wiley Interdisciplinary Reviews: RNA* 5.4 (Apr. 2014), pp. 461–480. DOI: 10.1002/wrna.1224. URL: https://doi.org/10.1002/wrna.1224.

[26] John F. Atkins. “BIOCHEMISTRY: The 22nd Amino Acid”. In: *Science* 296.5572 (May 2002), pp. 1409–1410. DOI: 10.1126/science.1073339. URL: https://doi.org/10.1126/science.1073339.

[27] Julian C. W. Willis and Jason W. Chin. “Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs”. In: *Nature Chemistry* 10.8 (May 2018), pp. 831–837. DOI: 10.1038/s41557-018-0052-5. URL: https://doi.org/10.1038/s41557-018-0052-5.

[28] Václav Beránek, Julian C. W. Willis, and Jason W. Chin. “An Evolved Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells”. In: *Biochemistry* 58.5 (Sept. 2018), pp. 387–390. DOI: 10.1021/acs.biochem.8b00808. URL: https://doi.org/10.1021/acs.biochem.8b00808.

[29] Carla R. Polycarpo et al. “Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase”. In: *FEBS Letters* 580.28-29 (Nov. 2006), pp. 6695–6700. DOI: 10.1016/j.febslet.2006.11.028. URL: https://doi.org/10.1016/j.febslet.2006.11.028.

[30] Jeffery M. Tharp, Andreas Ehnbom, and Wenshe R. Liu. “tRNAPyl: Structure, function, and applications”. In: *RNA Biology* 15.4-5 (Sept. 2017), pp. 441–452. DOI: 10.1080/15476286.2017.1356561. URL: https://doi.org/10.1080/15476286.2017.1356561.

[31] Lei Wang et al. “Expanding the Genetic Code of Escherichia coli”. In: *Science* 292.5516 (Apr. 2001), pp. 498–500. DOI: 10.1126/science.1060077. URL: https://doi.org/10.1126/science.1060077.

[32] Randall A. Hughes and Andrew D. Ellington. “Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA”. In: *Nucleic Acids Research* 38.19 (June 2010), pp. 6813–6830. DOI: 10.1093/nar/gkq521. URL: https://doi.org/10.1093/nar/gkq521.

[33] Sviatlana Smolskaya and Yaroslav A. Andreev. “Site-Specific Incorporation of Unnatural Amino Acids into Escherichia coli Recombinant Protein: Methodology Development and Recent Achievement”. In: *Biomolecules* 9.7 (June 2019), p. 255. DOI: 10.3390/biom9070255. URL: https://doi.org/10.3390/biom9070255.

[34] Richard B. Cooley et al. “Structural Basis of Improved Second-Generation 3-Nitrotyrosine tRNA Synthetases”. In: *Biochemistry* 53.12 (Mar. 2014), pp. 1916–1924. DOI: 10.1021/bi5001239. URL: https://doi.org/10.1021/bi5001239.

[35] Michel Goldschmidt-Clermont. “Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker for site-directed transformation of chlamydomonas”. In: *Nucleic Acids Research* 19.15 (1991), pp. 4083–4089. DOI: 10.1093/nar/19.15.4083. URL: https://doi.org/10.1093/nar/19.15.4083.

[36] Joseph M. Bateman and Saul Purton. “Tools for chloroplast transformation in Chlamydomonas: expression vectors and a new dominant selectable marker”. In:

Molecular and General Genetics MGG 263.3 (Apr. 2000), pp. 404–410. DOI: 10.1007/s004380051184. URL: https://doi.org/10.1007/s004380051184.

[37] Elizabeth H Harris. "Chlamydomonas as a Model Organism". In: 52.1 (June 2001), pp. 363–406. DOI: 10.1146/annurev.arplant.52.1.363. URL: https://doi.org/10.1146/annurev.arplant.52.1.363.

[38] Peter Snustad and Michael J. Simmons. *Genetics International Student Version*. 6th ed. John Wiley & Sons, Inc, 2012. ISBN: 978-1-118-09242-2.

[39] Lior Doron, Na'ama Segal, and Michal Shapira. "Transgene Expression in Microalgae—From Tools to Applications". In: *Frontiers in Plant Science* 7 (Apr. 2016). DOI: 10.3389/fpls.2016.00505. URL: https://doi.org/10.3389/fpls.2016.00505.

[40] Dwight Barnes et al. "Contribution of 5′- and 3′-untranslated regions of plastid mRNAs to the expression of Chlamydomonas reinhardtii chloroplast genes". In: *Molecular Genetics and Genomics* 274.6 (Oct. 2005), pp. 625–636. DOI: 10.1007/s00438-005-0055-y. URL: https://doi.org/10.1007/s00438-005-0055-y.

[41] Sean D. Gallaher et al. "High-throughput sequencing of the chloroplast and mitochondrion of Chlamydomonas reinhardtii to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates". In: *The Plant Journal* 93.3 (Jan. 2018), pp. 545–565. DOI: 10.1111/tpj.13788. URL: https://doi.org/10.1111/tpj.13788.

[42] Severin Sasso et al. "From molecular manipulation of domesticated Chlamydomonas reinhardtii to survival in nature". In: *eLife* 7 (Nov. 2018). DOI: 10.7554/elife.39233. URL: https://doi.org/10.7554/elife.39233.

[43] Poul Erik Jensen and Dario Leister. "Chloroplast evolution, structure and functions". In: *F1000Prime Reports* 6 (June 2014). DOI: 10.12703/p6-40. URL: https://doi.org/10.12703/p6-40.

[44] Bartolomé Sabater. "Evolution and Function of the Chloroplast. Current Investigations and Perspectives". In: *International Journal of Molecular Sciences* 19.10 (Oct. 2018), p. 3095. DOI: 10.3390/ijms19103095. URL: https://doi.org/10.3390/ijms19103095.

[45] Saul Purton et al. "Genetic engineering of algal chloroplasts: Progress and prospects". In: *Russian Journal of Plant Physiology* 60.4 (June 2013), pp. 491–499. DOI: 10.1134/s1021443713040146. URL: https://doi.org/10.1134/s1021443713040146.

[46] Miller Tran et al. "Production of unique immunotoxin cancer therapeutics in algal chloroplasts". In: *Proceedings of the National Academy of Sciences* 110.1 (Dec. 2012), E15–E22. DOI: 10.1073/pnas.1214638110. URL: https://doi.org/10.1073/pnas.1214638110.

[47] Juliane Neupert et al. "Genetic Transformation of the Model Green Alga Chlamydomonas reinhardtii". In: *Methods in Molecular Biology*. Humana Press, 2012, pp. 35–47. DOI: 10.1007/978-1-61779-558-9_4. URL: https://doi.org/10.1007/978-1-61779-558-9_4.

[48] Stephen P Mayfield et al. "Chlamydomonas reinhardtii chloroplasts as protein factories". In: *Current Opinion in Biotechnology* 18.2 (Apr. 2007), pp. 126–133. DOI: 10.1016/j.copbio.2007.02.001. URL: https://doi.org/10.1016/j.copbio.2007.02.001.

[49] Sriram Kosuri and George M Church. "Large-scale de novo DNA synthesis: technologies and applications". In: *Nature Methods* 11.5 (Apr. 2014), pp. 499–507. DOI: 10.1038/nmeth.2918. URL: https://doi.org/10.1038/nmeth.2918.

[50] Ran Chao, Yongbo Yuan, and Huimin Zhao. "Recent advances in DNA assembly technologies". In: *FEMS Yeast Research* (June 2014), n/a–n/a. DOI: 10.1111/1567-1364.12171. URL: https://doi.org/10.1111/1567-1364.12171.

[51] George M Taylor, Paweł M Mordaka, and John T Heap. "Start-Stop Assembly: a functionally scarless DNA assembly system optimized for metabolic engineering". In: *Nucleic Acids Research* 47.3 (Nov. 2018), e17–e17. DOI: 10.1093/nar/gky1182. URL: https://doi.org/10.1093/nar/gky1182.

[52] Susan Buckhout-White et al. "Restriction Enzymes as a Target for DNA-Based Sensing and Structural Rearrangement". In: *ACS Omega* 3.1 (Jan. 2018), pp. 495–502. DOI: 10.1021/acsomega.7b01333. URL: https://doi.org/10.1021/acsomega.7b01333.

[53] Raymond J. Williams. "Restriction Endonucleases: Classification, Properties, and Applications". In: *Molecular Biotechnology* 23.3 (2003), pp. 225–244. DOI: 10.1385/mb:23:3:225. URL: https://doi.org/10.1385/mb:23:3:225.

[54] Alfred Pingoud and Albert Jeltsch. "Structure and function of type II restriction endonucleases". In: *Nucleic Acids Research* 29.18 (Sept. 2001), pp. 3705–3727. DOI: 10.1093/nar/29.18.3705. URL: https://doi.org/10.1093/nar/29.18.3705.

[55] Shaun M. Lippow et al. "Creation of a type IIS restriction endonuclease with a long recognition sequence". In: *Nucleic Acids Research* 37.9 (Mar. 2009), pp. 3061–3073. DOI: 10.1093/nar/gkp182. URL: https://doi.org/10.1093/nar/gkp182.

[56] James C. Samuelson. "The isolation of strand-specific nicking endonucleases from a randomized SapI expression library". In: *Nucleic Acids Research* 32.12 (July 2004), pp. 3661–3671. DOI: 10.1093/nar/gkh674. URL: https://doi.org/10.1093/nar/gkh674.

[57] V. Butkus et al. "A new restriction endonuclease Eco31I recognizing a non-palindromic sequence". In: *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 826.4 (Dec. 1985), pp. 208–212. DOI: 10.1016/0167-4781(85)90008-9. URL: https://doi.org/10.1016/0167-4781(85)90008-9.

[58] Wojciech Siwek et al. "Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R.DpnI". In: *Nucleic Acids Research* 40.15 (June 2012), pp. 7563–7572. DOI: 10.1093/nar/gks428. URL: https://doi.org/10.1093/nar/gks428.

[59] Karolina Mierzejewska et al. "Structural basis of the methylation specificity of R.DpnI". In: *Nucleic Acids Research* 42.13 (June 2014), pp. 8745–8754. DOI: 10.1093/nar/gku546. URL: https://doi.org/10.1093/nar/gku546.

[60] Carola Engler et al. "Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes". In: *PLoS ONE* 4.5 (May 2009). Ed. by Jean Peccoud, e5553. DOI: 10.1371/journal.pone.0005553. URL: https://doi.org/10.1371/journal.pone.0005553.

[61] Rossella Rossi et al. "Functional characterization of the T4 DNA ligase: a new insight into the mechanism of action". In: *Nucleic Acids Research* 25.11 (June 1997), pp. 2106–2113. DOI: 10.1093/nar/25.11.2106. URL: https://doi.org/10.1093/nar/25.11.2106.

[62] Daniel G Gibson et al. "Enzymatic assembly of DNA molecules up to several hundred kilobases". In: *Nature Methods* 6.5 (Apr. 2009), pp. 343–345. DOI: 10.1038/nmeth.1318. URL: https://doi.org/10.1038/nmeth.1318.

[63] Karim Kadri. “Polymerase Chain Reaction (PCR): Principle and Applications”. In: *Synthetic Biology - New Interdisciplinary Science*. IntechOpen, Feb. 2020. DOI: 10.5772/intechopen.86491. URL: https://doi.org/10.5772/intechopen.86491.

[64] Vyacheslav Palchevskiy and Steven E. Finkel. “Escherichia coli Competence Gene Homologs Are Essential for Competitive Fitness and the Use of DNA as a Nutrient”. In: *Journal of Bacteriology* 188.11 (June 2006), pp. 3902–3910. DOI: 10.1128/jb.01974-05. URL: https://doi.org/10.1128/jb.01974-05.

[65] J.M. Coll. “Methodologies for transferring DNA into eukaryotic microalgae: a review”. In: *Spanish Journal of Agricultural Research* 4.4 (Dec. 2006), p. 316. DOI: 10.5424/sjar/2006044-209. URL: https://doi.org/10.5424/sjar/2006044-209.

[66] Jacob E. Lazarus et al. “A New Suite of Allelic-Exchange Vectors for the Scarless Modification of Proteobacterial Genomes”. In: *Applied and Environmental Microbiology* 85.16 (June 2019). Ed. by Christopher A. Elkins. DOI: 10.1128/aem.00990-19. URL: https://doi.org/10.1128/aem.00990-19.

[67] Douglas H. Juers, Brian W. Matthews, and Reuben E. Huber. “*LacZ* b-galactosidase: Structure and function of an enzyme of historical and molecular biological importance”. In: *Protein Science* 21.12 (Nov. 2012), pp. 1792–1807. DOI: 10.1002/pro.2165. URL: https://doi.org/10.1002/pro.2165.

[68] Maxim Kostylev et al. “Cloning Should Be Simple: Escherichia coli DH5-Mediated Assembly of Multiple DNA Fragments with Short End Homologies”. In: *PLOS ONE* 10.9 (Sept. 2015). Ed. by Mark Isalan, e0137466. DOI: 10.1371/journal.pone.0137466. URL: https://doi.org/10.1371/journal.pone.0137466.

[69] *New England Biolab (NEB) Tm Calculator v.1.12.0*. https://tmcalculator.neb.com/#!/main. Accessed: 21.04.2020.

[70] Juliane Strien, Juliane Sanft, and Gita Mall. “Enhancement of PCR Amplification of Moderate GC-Containing and Highly GC-Rich DNA Sequences”. In: *Molecular Biotechnology* 54.3 (Apr. 2013), pp. 1048–1054. DOI: 10.1007/s12033-013-9660-x. URL: https://doi.org/10.1007/s12033-013-9660-x.

[71] Marco Larrea-Alvarez and Saul Purton. “Multigenic engineering of the chloroplast genome in the green alga Chlamydomonas reinhardtii”. In: *Microbiology* (Apr. 2020). DOI: 10.1099/mic.0.000910. URL: https://doi.org/10.1099/mic.0.000910.

A Solutions

Content of stock solution of reaction mixture used for Gibson Assembly is listed in Table A.1, while content of 5x isothermal reaction buffer (5xIRB) is listed in Table A.2.

Table A.1: Content of Gibson reaction mixture stock solution. Splitt into 15 µl aliquots and stored at -20°C. Content of 5x isothermal reaction buffer (5xIRB) is listed in Table A.2

Solution	Used Volume [µl]
5xIRB	320
T5 Exonuclease	0.64
Phusion High-Fidelity DNA Polymerase	20
Taq DNA ligase	160
dH_2O	700

Table A.2: Content of 5x isothermal reaction buffer (5xIRB) used for Gibson Assembly.

Solution	Used Volume [µl]	Final concentration
2 M Tris-HCl	250	500 mM
1 M DTT	50	50 mM
2 M $MgCl_2$	25	50 mM
100 mM NAD	50	5 mM
10 mM dNTP	100	1 mM
50% PEG-8000	500	25%
dH_2O	25	

Phosphate buffer II and solution A were prepared for use in TAP medium for incubation of *C. reinhardtii*. Content of these solutions are given in Table A.3 and Table A.4.

Table A.3: Concentration (C) of components of Phosphate buffer II used in TAP medium. Dilutions were made using distilled water.

Component	C [g/L]
K_2HPO_4	1.08
KH_2PO_4	0.56

Table A.4: Concentrations (C) of components in Solution A used in TAP medium. Dilutions were made using distilled water.

Component	C [g/L]
$CaCl_2 \cdot 2\,H_2O$	1.25
$MgSO_4 \cdot 7\,H_2O$	2.50
NH_4Cl	10.0

B Plasmid Reaction Volumes for Start-Stop Assembly

The amount of each plasmid included in the Start-Stop assembly of level 1 constructs and Golden Gate assembly of level 2 constructs was calculated based on plasmid size and concentration. All concentrations were measured as mass concentration (C_g). These values were converted to molar concentrations (C_m) using Equation (B.1), where N is the plasmids number of base pairs. For simplicity, an average molecular weight (Mw) was estimated to 660 g/mol per deoxyribonucleoside monophosphate pair.

$$C_m = C_g \cdot \frac{1}{N} \cdot \frac{1}{Mw} \tag{B.1}$$

The relation given in Equation (B.2) was used to calculate necessary volume of each plasmid to be added to the reaction. This equation describes the relation between, plasmid stock concentration (C_1), added volume of the plasmid (V_1), total reaction volume (V_2) and final plasmid concentration (C_2). Total reaction volume was 20 µl, and concentration was set to 3 nM acceptor plasmid, and 1 nM of each donor plasmid.

$$C_1 \cdot V_1 = C_2 \cdot V_2 \tag{B.2}$$

All calculated reaction volumes are listed in Table B.1. To demonstrate the calculations, an example for donor plasmid p101BC_psbD_MaPylRS-3MeHis_petA is given, using values also listed in Table B.1. Firstly, molar concentration was found using Equation (B.1).

$$C_{m(p101)} = 56\,\text{ng/µl} \cdot \frac{1}{4290} \cdot \frac{1}{660\text{nmol/ng}} = 0.02\,\text{pmol/µl} \tag{B.3}$$

Then, Equation (B.2) was solved for V_1 using obtained molar concentration

$$V_{1(p101)} = \frac{0\ 003\text{pmol}/\mu\text{l} \cdot 20\mu\text{l}}{0\ 02\text{pmol}/\mu\text{l}} = 3.0\,\text{µl} \tag{B.4}$$

Table B.1: List of plasmids used during Start-Stop and Golden Gate assembly of level 1 and 2 constructs. Each plasmid is listed with plasmid size, measured stock concentration (C_g) converted to molar concentration (C_1), as well as assembly reaction volume (V_1) and final concentration (C_2)

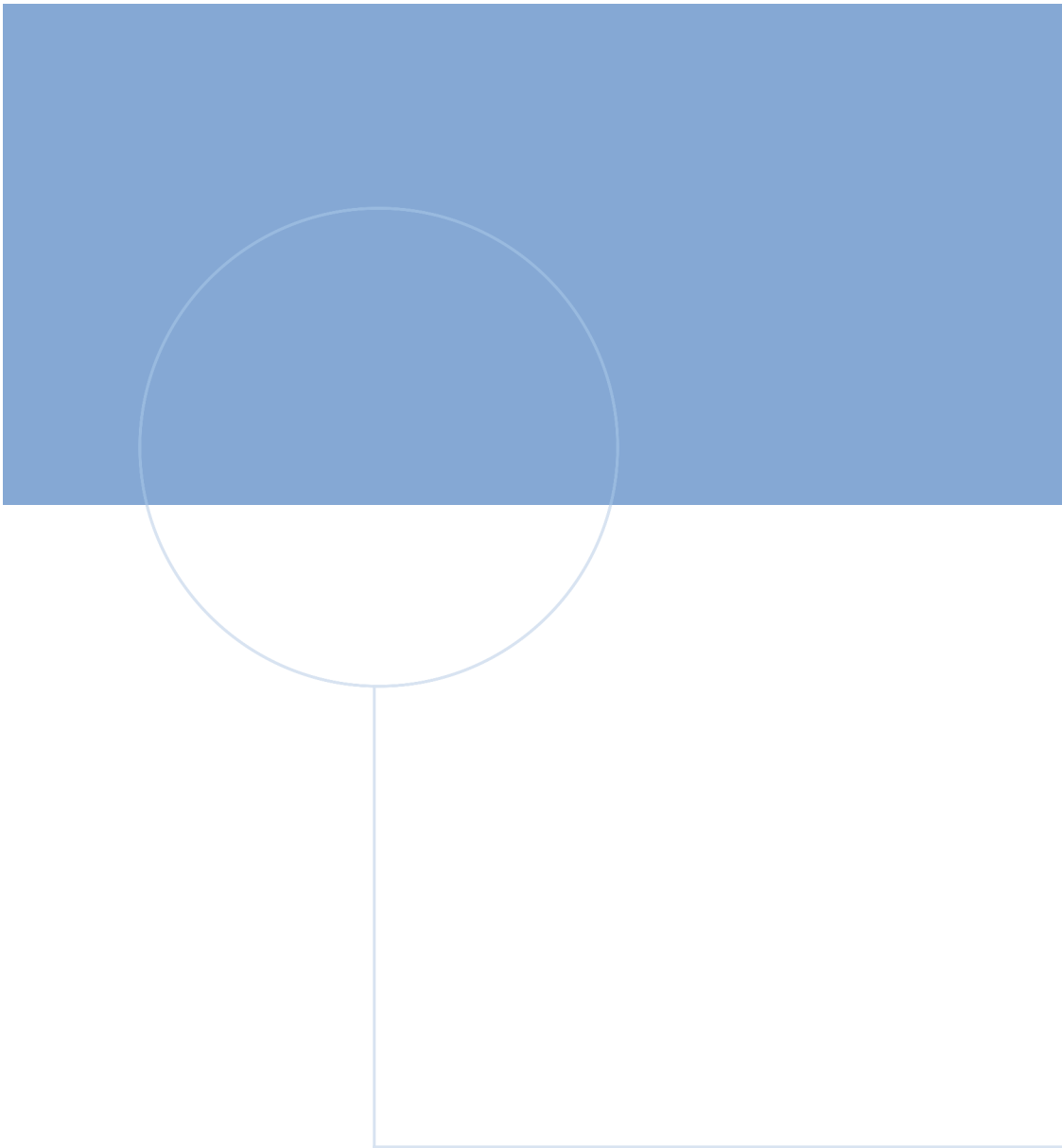
Plasmid	Plasmid size [bp]	C_g [ng/µl]	C_2 [pmol/µl]	C_1 [pmol/µl]	V_1 [µl]
pUC8_MaPylRS-3MeHis	3287	78	0.03	0.04	1.7
pUC8_MaPylRS-RS1	3287	48	0.03	0.02	2.7
p001_MbPylRS	3463	65	0.03	0.03	2.1
p002_MbPylRS-Ack1	3463	68	0.03	0.03	2.0
p003_MjTyrRS-NitroTyr	3124	62	0.03	0.03	2.0
p005_psbD	2419	62	0.03	0.04	1.5
p007_mVenusAmber	2947	75	0.03	0.04	1.6
p008_mVenusOpal	2950	76	0.03	0.04	1.5
p010_ MbtRNA	2488	56	0.03	0.03	1.8
p101BC_psbD_MaPylRS-3MeHis_petA	4290	56	0.03	0.02	3.0
p102BC_psbD_MaPylRS-RS1_petA	4290	62	0.03	0.02	2.7
p103BC_psbD_MbPyl_petA	4722	60	0.03	0.02	3.1
p104BC_psbD_MbPylRS-Ack1_petA	4722	71	0.03	0.02	2.6
p105BC_psbD_MmPylRS-AF_petA	4827	71	0.03	0.02	2.7
p106BC_psbD_MjTyrRS-NitroTyr_petA	4388	61	0.03	0.02	2.8
p107CD_MbtRNA	3285	114	0.03	0.05	1.1
p108CD_MatRNA	3286	72	0.03	0.03	1.8
p109CD_MjtRNA	3290	76	0.03	0.04	1.7
p110DZ_p16s_mVenusAmber_atpB	4288	70	0.03	0.02	2.4
p111DZ_p16s_mVenusOpal_atpB	4291	52	0.03	0.02	3.3
p207_psbH_aadA_tsPurple	6836	115	0.01	0.03	0.8
pSS124.Lv0p16S-psaA5	2560	67	0.03	0.04	1.5
pSS159_pLv0_petA	2460	27	0.03	0.02	3.6
pSS164_pLv0_ter.atpB	2408	34	0.03	0.02	2.8
pSS191.Lv0/2tsPurple	3111	88	0.01	0.04	0.5
pHT201.Lv0tRNAw	2489	65	0.03	0.04	1.5
pSS104_pLv1_BC	3488	24	0.01	0.01	1.9
pSS105_pLv1_CD	3488	24	0.01	0.01	1.9
pSS105_pLv1_DZ	3488	34	0.01	0.01	1.4
pHT146_pLv1_XA-psbH.LF	3810	50	0.03	0,02	3.0
pHT150_pLv1_ZY-psbH.RF	4200	80	0.03	0.03	2.1
pLv1_AB:atpA_aadA_rbcl	4659	57	0.03	0.02	3.2

C Primers for PCR amplification

A complete list of all PCR primers used during the project is given in Table C.1.

Table C.1: List of all PCR primers used for DNA amplification and mutagenesis in the project.

Primer	Sequence
HJ176	CAGTGTTGGGTCTCTGGTCGCTCTTCATTAGTGGTGGTGGTGGT
BB1	GAATTCGCGGCCGCTTCTAGA
BB2	CTGCAGCGGCCGCTACTAGTA
17	AGTGCTGCAATGATACCGCGTGACCCA
18	TGGGTCACGCGGTATCATTGCAGCACT
129	GAACTCGCTCTTCAATGAGGGAAGCGGTGATCGC
130	AGGATTGCTCTTCATTATTTGCCGACTACCTTGGTGATCT
384	TTTCGCCACCTCTGA
427	AGGGCGACACGGAAATGTTG
457	GAACTCGCTCTTCATTAGTGGTGGTGGTGGTGGTG
505	GAATCCAGTGTGGCTCTTCCATGTCTTAGGGTGAAGAATTATTTACTGGT
506	ACCAGTAAATAATTCTTCACCCTAAGACATGGAAGAGCCACACTGGATTC
507	ACTACAACTCTCACAACGTATGAATTACTGCTGATAAACAAAA
508	TTTTGTTTATCAGCAGTAATTCATACGTTGTGAGAGTTGTAGT
509	AGGGTTGGTCTCATGTGGCTCTTCAATGGATAAAAAACCA
510	GTGTTGGGTCTCTGGTCGCTCTTCATTATAAGTTTGTTGAAATACCGTTGTAG
511	AGGGTTGGTCTCATGTGGCTCTTCGATGGACGAATTTGAAATGATTAAGAG
512	GTGTTGGGTCTCTGGTCGCTCTTCATTATAATCTCTTTCTAATTGGCTCTAAAATC
513	TTAGCTCCAAACTTAATGTCTGTTGGTCGTGATTTACGTGAC
514	CGACCAACAGACATTAAGTTTGGAGCTAACATTGGACGAAGTGCACGT
515	GAAGAATTTACTATGCTTAACCTTACTGACATGGGTCCACGTGGT
516	ACCACGTGGACCCATGTCAGTAAGGTTAAGCATAGTAAATTCTTC
517	ACGTTGTCGTGTATCTGGTGAAGATATTAACAACTTCTTA
518	TAAGAAGTTGTTAATATCTTCACCAGATACACGACAACGT
519	GTAGCTCCAACTATTTTCAACTACGCACGTAAATTAGATCGTATCTTACCAGGT
520	TGCGTAGTTGAAAATAGTTGGAGCTACCATTGGACGTAAACATAAGTTTTTATCTAC
521	ACTATGGTAAACTTCTTTCAAATGGGTTCTGGTTGT
522	ACAACCAGAACCCATTTGAAAGAAGTTTACCATAGT
523	TTGAAGTCCGTTCGATCTACATGATCAGGTTTCCAAAATTTGTTATTATAATTTTTATCATGGTATTATTTTAATGTCAATCAG
524	ATCCGTTCAGCCGGGTTAGATTCCCGGGGTTTCCGCCATTTTTCCTAATGTACTTTGTTGTAAAAGTGGCTGGTTT
525	TTTTGAAGACCCGCTGGTCGCCGGACCGTCCCCCAAAATTTGTTATTATAATTTTTATCATGGTATTATTTTAATGTCAATCAG
526	CCTAGCATAGCGGGGTTCGACGCCCCGGTCCCTCGCCATTTTTCCTAATGTACTTTGTTGTAAAAGTGGCTGGTTT
527	TTAGAGTCCGCCGTTCTGCCCTGCTGAACTACCGCCGGAAAATTTGTTATTATAATTTTTATCATGGTATTATTTTAATGTCAATCAG
528	ATCCGCATGGCGCTGGTTCAAATCCGGCCCGCCGGACCATTTTTCCTAATGTACTTTGTTGTAAAAGTGGCTGGTTT
535	AGGGTTGGTCTCATGTGGCTCTTCACATTTATATAAATAAATGTAACTTCTTTTGACGATCC
536	GTGTTGGGTCTCTGGTCGCTCTTCACAGTAGTAAGTAAACTGCGTAAGACGAC
537	GTGTTGGGTCTCTGGTCGCTCTTCACAGCTCTAGCATTCGTTTACAGGCC
538	AGGGTTGGTCTCATGTGGCTCTTCACATAAAAAATTTTTTACTGATTTTTTTACTTAATTATTACAAGGTTAT
545	AGAGTTGGTCTCANATGGTTTCTAAAGGTGAAGAATTATTTACTGGTG
546	GTGTTGGGTCTCTGGTCGCTCTTCACAGTGTGATGACTATGCACAAAGCAGTTCTAG
547	AGGGTTGGTCTCATGTGGCTCTTCACATTGCGTGTATCTCCAAAATAAAAAAACAACTC
554	AGTGTGGCTCTTCGCAGGCCTAGTATGATTTTGTATGTTGTTAATGC
555	GAGGTCGCTCTTCATCCTTAGTTAAAGATATAACTAGAGAAAAGAACAGCTAATACG
556	AGTGTGGCTCTTCGCAGAATTTCTTATTCGATTTTTTTTGTTGATAAACCTTT
557	GAGGTCGCTCTTCATCCGTAAATGGAGCTACGAAAGTATCTAAGTA
581	CTAGTAGGTCTCATGTGTCCAGTGTGGCTCTTCCATG
582	CTAGTAGGTCTCAGGTCTTAGTGGTGGTGGTGGTGTTTGTATAATTCATCCATACCTAAAGTAATACCAG



NTNU

Norwegian University of
Science and Technology