GENETIC ENGINEERING OF GREEN MICROALGAE FOR THE PRODUCTION OF BIOFUEL AND HIGH VALUE PRODUCTS

Joanna Beata Szaub

Department of Structural and Molecular Biology

University College London

A thesis submitted for the degree of Doctor of Philosophy

August 2012

DECLARATION

I, Joanna Beata Szaub confirm that the work presented in this thesis is my own.
Where information has been derived from other sources,
I confirm that this has been indicated in the thesis.

Signed:

ABSTRACT

A major consideration in the exploitation of microalgae as biotechnology platforms is choosing robust, fast-growing strains that are amenable to genetic manipulation. The freshwater green alga *Chlorella sorokiniana* has been reported as one of the fastest growing and thermotolerant species, and studies in this thesis have confirmed strain UTEX1230 as the most productive strain of *C. sorokiniana* with doubling time under optimal growth conditions of less than three hours. Furthermore, the strain showed robust growth at elevated temperatures and salinities.

In order to enhance the productivity of this strain, mutants with reduced biochemical and functional PSII antenna size were isolated. TAM4 was confirmed to have a truncated antenna and able to achieve higher cell density than WT, particularly in cultures under decreased irradiation. The possibility of genetic engineering this strain has been explored by developing molecular tools for both chloroplast and nuclear transformation. For chloroplast transformation, various regions of the organelle's genome have been cloned and sequenced, and used in the construction of transformation vectors. However, no stable chloroplast transformant lines were obtained following microparticle bombardment. For nuclear transformation, cycloheximide-resistant mutants have been isolated and shown to possess specific missense mutations within the *RPL41* gene. Such a mutant allele should prove useful as a dominant marker.

Genetic engineering of the chloroplast genome has been well established for another microalga *Chlamydomonas reinhardtii*. This system was exploited in three biotechnological applications: 1) generation of alkane producing strains by introducing genes encoding for acyl reductase and aldehyde decarbonylase. 2) expression of a vaccine candidate major capsid protein L1 of the human papillomavirus. 3) expression of a potent HIV-inactivating protein cyanovirin-N. In all cases, stable transformant lines were obtained and molecular analysis confirmed the successful integration of the transgenes into the genome. The detailed biochemical analysis of the lines is presented in the thesis.

ACKNOWLEDGEMENT

I would like to acknowledge and extend my gratitude to those who made the completion of this thesis possible:

My supervisor, Dr Saul Purton for his great support and guidance throughout my PhD.

Everyone in the Purton lab, particularly Lamya Al-Haj for her continued support, encouragement and proof reading, Laura Stoffels for her enthusiasm and proof reading, Chloe Economou, Henry Taunt and Rosie Young for proof reading and encouragement, and Janet Waterhouse for her continued support.

Thanya Wannathong for her all-time company and support.

My family in Poland and UK for being there for me whenever I needed them.

My great friends Marta Teperek, Eti Galvin and Kevin Little for their continued support.

ABBREVIATIONS

AMPS	ammonium persulphate
ATP	adenosine 5’-triphosphate
CAI	codon adaptation index
cDNA	complementary deoxyribonucleic acid
CoA	coenzyme A
CYH	cycloheximide
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea
DIG	digoxigenin-dUTP
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	2’deoxynucleoside 5’-triphosphate
DW	dry weight
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid (disodium salt)
FA	fatty acid
FAMEs	fatty acid methyl esters
Fm	maximum fluorescence yield
Fo	minimal fluorescence yield
GC	gas chromatography
GC/MS	gas chromatography–mass spectrometry
gDNA	genomic deoxyribonucleic acid
HPLC	high performance liquid chromatography
IgG	Immunoglobulin G
LHC	light-harvesting complexes
mRNA	messenger ribonucleic acid
NPQ	non-photochemical quenching
OD	optical density
ORF	open reading frame

PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PS	photosystem
rDNA	ribosomal deoxyribonucleic acid
RNA	ribonucleic acid
RNase	ribonuclease
RPL41	ribosomal protein L41
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription-polymerase chain reaction
SD	standard deviation
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TAGs	triacylglycerols
TE	tris EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
tris	tris (hydroxymethyl) aminomethane
tRNA	transfer ribonucleic acid
TSP	total soluble protein
UTR	untranslated region
UV	ultraviolet
v/v	volume for volume
w/v	weight for volume
WT	wild type

LIST OF FIGURES

LIST OF TABLES

LIST OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1. Microalgae

1.1.1. General overview

Microalgae are a polyphyletic group of autotrophic microorganisms that are either unicellular or comprise simple colonial arrangement of cells. This diverse group not only plays a crucial role in primary production in aquatic ecosystems but also participates in the utilization of global atmospheric carbon dioxide.

Figure 1-1 illustrates the global distribution of chlorophyll including not only terrestrial plants but also the algal productivity that is predominantly focused in large repositories of water. Although phytoplankton biomass contributes only 1-2% of the total plant carbon, its ability to fix atmospheric carbon reaches approximately 40% of the total carbon fixed by autotrophic organisms (Falkowski, 1994).

It has been estimated there are between 200,000 and 800,000 algal species present in various ecosystems, of which approximately 35,000 have been classified and described (Ebenezer et al., 2012). The basis for classification of microalgae has been a great challenge for several decades, being based largely on pigmentation and morphology under the light microscope - phenotypes that can be both shared between unrelated algae, and change depending on growth conditions. The introduction of electron microscopy facilitated in-depth studies of algal ultrastructure and resulted in the creation of a new hypothesis of the origin of each algal taxon (Mattox, 1984). However, the morphology of a vegetative cell was revealed to be an inadequate criterion for classification and resulted in the mis-classification of many species, and general disorder in algal taxonomy. Features that are currently applied in algal taxonomy are based on morphological elements. These include cell wall, presence and morphology of a flagellar root system and basal bodies, since these features are recognized as evolutionarily stable (Friedl, 1997).

Figure 1-1 Global distribution of chlorophyll, including both oceanic phytoplankton and terrestrial vegetation.

Reproduced from: http://www.enotes.com/topic/Photosynthesis

Additional methods that give more comprehensive overview of algal phylogeny include molecular analysis of chloroplast and nuclear-encoded genes, such as the evolutionarily conserved 16S, 18S and 26S rRNA (Buchheim and Chapman, 1991, Nakada et al., 2008, Shi et al., 2011, Wu et al., 2001), or sequences of chloroplast genes such as *rbcL* and *atpB* (Nozaki et al., 1999). Finally, a classification method based on ultrastructural features of the plastid organelle, particularly the number of bounding membranes (Sitte, 1993) has been established and this topic will be discussed in detail in subchapter 1.2.1.

1.1.2. Commercial applications of microalgae

Microalgae have been utilized by indigenous cultures for many centuries, yet their commercial cultivation was only initiated in the early 1960s (Spolaore et al., 2006). A continuously increasing world population initiated the search for unconventional protein sources and algal biomass emerged as a suitable alternative (Becker, 2007) since the protein composition of many algal species such as *Spirulina sp*., *Scenedesmus sp*., *Chlorella sp*., and *Synechococcus sp*. exceeds the protein content present in meat, milk or soybean (Spolaore et al., 2006). Additionally, microalgae are a valuable source of many compounds essential in human nutrition such as vitamins (A, B_1, B_2, B_6, B_{12}, C, E, nicotinate, biotin, folic acid and pantothenic acid), pigments (chlorophyll, carotenoids or phycobiliproteins) and polyunsaturated fatty acids (PUFA) such as γ-linolenic acid (GLA), arachidonic acid (AA), eicosapentaenoic acid (EPA) or docosahexaenoic acid (DHA) (Becker, 2007, Spolaore et al., 2006). Currently the only commercially available algae derived PUFA is DHA.

Microalgae for human consumption are produced mainly in the form of tablets, capsules or as powder, but can also be found in ready-made food products such as pasta, snack food or beverages (Liang et al., 2004, Yamaguchi, 1996). Commercial products used in human nutrition are limited to a few species: *Arthrospira sp.*, *Chlorella sp.*, *Dunaliella salina*, *Aphanizomenon flos-aquae* (Spolaore et al., 2006) and *Spirulina sp*. Conversely, microalgal cultures are also commonly exploited for animal feed purposes, which accounts for approximately 30% of the global algal production (Becker, 2007) with *Arthrospira sp.* being the dominant species in animal feed production (Yamaguchi, 1996). Additionally, *Arthrospira sp., Chlorella sp.* extracts have also been applied in cosmetic

industry, where a range of protein-rich skincare, hair care and sun protection products have been developed in recent years (Stolz and Obermayer, 2005)

1.1.3. Chlorella sorokiniana

Chlorella is a genus of single cell non-flagellated spherical green algae varying from 2-10 μm in diameter. Chlorella species are widely distributed and present in fresh and seawater, soil and air. Since the first successful isolation as a pure culture in 1890 by Bejerinck, this genus has been extensively used as a model organism in many biochemical studies, such as photosynthesis (Myers, 1946, Myers and Cramer, 1947), nitrate reduction (Kessler, 1953, Cramer and Myers, 1948), respiration (Eny, 1950, Syrett, 1951), and cell growth (Takeda and Hirokawa, 1979, Samejima and Myers, 1958, Shi et al., 2000). There are over 20 characterised *Chlorella* species with over 100 identified strains (Wu et al., 2001, Furnas, 1990) and *Chlorella sorokiniana* is the major species used in the research presented in this work. Figure 1-2 presents a photograph of a liquid culture of the *C. sorokiniana* UTEX1230 strain taken using an optical microscope [Leica].

Chlorella sorokiniana was isolated and identified for the first time in 1953 (Sorokin and Myers, 1953a) as a putative *Chlorella pyrenoidosa* strain. For many years this species was described as a thermophilic mutant of *Chlorella pyrenoidosa* (Kunz, 1972) until the late 1980s and the beginning of 1990s when DNA work performed on different strains of *Chlorella* officially identified it as a separate species (Dorr and Huss, 1990, Huss and Jahnke, 1994, Kessler, 1985). The taxonomic identification of *Chlorella* species is considerably difficult as all species from this genus are characterised by an asexual reproductive cycle combined with the lack of obvious morphological features. Currently for the laboratory purposes the most efficient identification method of *Chlorella* sp. is determining the differences in conserved and variable regions within the chloroplast-encoded 16S rDNA sequence (Burja et al., 2001, Wu et al., 2001) combined with analysis of the nucleus-encoded 18S rRNA gene (Volker Huss, 1999).

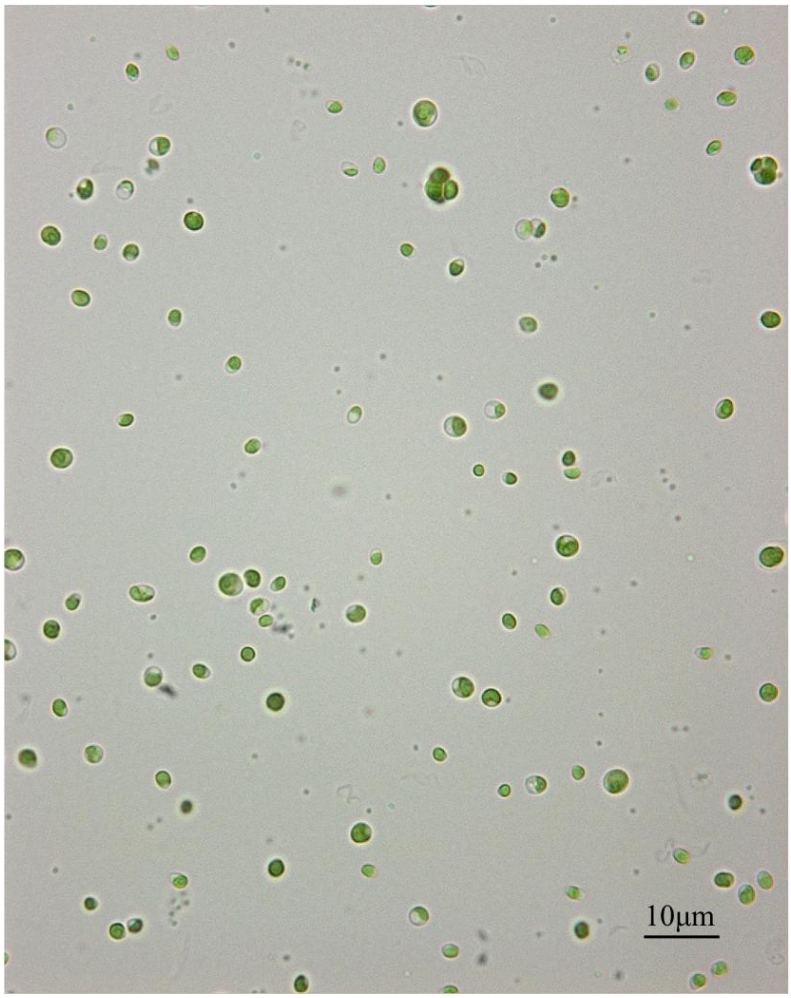


Figure 1-2 A culture of *Chlorella sorokiniana* strain UTEX1230
The picture was captured using Leica optical microscope at the Cell and Developmental Biology Department at UCL.

Chlorella sorokiniana is potentially an important species in the pharmaceutical industry as some strains produce a significant amount of antioxidants. For instance, the total yield of carotenoids from *C. sorokiniana* H-84 reaches up to 0.69% of dry matter. This was achieved in extreme conditions of elevated temperature and an increased CO_2 level up to 10%. The production of lutein in those conditions was 4.3mg/g dry weight, whereas the production of β-carotene and α-tocopherol was 0.6 and 0.1 mg/g dry weight respectively (Matsukawa et al., 2000). This property is currently of great interest as it has been shown that even a single dose of *Chlorella* extracts can have a significant positive impact on the level of serum lutein in humans (Shibata and Hayakawa, 2009), and this has been associated with a decreased risk of chronic diseases (Mares-Perlman and Erdman, 2002). This result shows that *Chlorella* species are not only rich in carotenoids, but the carotenoids are also bioavailable. Moreover, research performed on rats has revealed benefits in using dried powdered *Chlorella* as a dietary supplement. It has been demonstrated that this supplement regulates cholesterol levels and blood pressure and also has a positive influence on the life span of the tested animals (Sansawa et al., 2006).

Chlorella has a great potential to be a bioreactor for the production of pharmaceutical and industrial compounds, since it shares many fundamental metabolic pathways with higher plants. Additionally, due to its small size and simple growth requirements, the cultivation of *Chlorella* is very straightforward, reproducible and inexpensive. However, its application as bioreactor has been constrained by the lack of the routine transformation systems that would allow the genetic engineering of *Chlorella* strains for different applications. A limited number of reports exist of successful transformation of *Chlorella sp.*, and frequently these are restricted to the transient expression of foreign proteins (Jarvis and Brown, 1991, Wang et al., 2007).

Chlorella sorokiniana has been also investigated as a potential candidate for production of biofuels. Analysis of the dry mass of *Chlorella sorokiniana* shows that the species consists of on average 40% proteins, 30-38% carbohydrates and 18-22% lipids (Belkoura et al., 1997, Illman et al., 2000, Gouveia and Oliveira, 2009). However, modifying the cultivation media can change this ratio. It has been shown that changing the carbon to nitrogen ratio in media significantly increases the lipid production of *Chlorella sorokiniana* (Feng and Johns, 1991). Additionally, the possibility of changing the fatty acid (FA) composition has been demonstrated through modification of the growth

conditions. An increased aeration rate up-regulates synthesis of unsaturated dienoic and trienoic FA (Feng and Johns, 1991). Moreover, by increasing the temperature to 35^0C the ratio of polyunsaturated/saturated fatty acids was increased in favor of the polyunsaturated FAs (Belkoura et al., 2000).

Another important energy carrier currently very appealing for industry is hydrogen (H_2). Presently, the main source of hydrogen involves using fossil fuels to chemically generate the gas, therefore there is a great interest in finding alternative biological sources of H_2. Strains of *Chlorella* such as *Chlorella sorokiniana* strain Ce, *Chlorella salina* strain Mt and *Chlorella sp.* strain Pt6 isolated from the soil and foggaras's water from the Algerian Sahara, under anaerobic conditions and sulphur deprivation were capable of production and accumulation of detectable amounts of H_2 (Chader et al., 2007), and their productivity was comparable to the previously reported hydrogen-producing strains of *C. reinhardtii* (Melis et al., 2000).

1.1.4. Chlamydomonas reinhardtii

Chlamydomonas is another genus of green algae that belong to the division of *Chlorophyta*. It is an ellipsoid, soil-dwelling unicellular alga of approximately 10-14 μm in diameter. *Chlamydomonas* species are highly adaptable to various environmental conditions and are present worldwide (Harris, 1989). Their cell contains a central nucleus, large chloroplast accommodating the photosynthetic apparatus and essential metabolic pathways, multiple mitochondria and two anterior flagella embedded in basal bodies. The flagella are used by the cells for motility and during mating (Harris, 1989). There are one or more pyrenoids present in the chloroplast with starch bodies surrounding the pyrenoid. Each cell has a well-defined cell wall, with a mucilaginous layer coating the external surface of the wall of some species. Figure 1-3 presents a schematic diagram of the *Chlamydomonas* cell indicating the position of each organelle. Differences in morphological features have been the major criteria in identification of *Chlamydomonas* species. Although there are over 450 recognized *Chlamydomonas* species (Ettl, 1976), *C. reinhardtii* remains the principal laboratory species. Its cell wall normally does not contain cellulose (Harris, 2001) and numerous cell wall–defective mutants have been isolated. Cell wall components in those isolates are produced at a normal level yet they fail to assemble them in the correctly formed cell structure (Voigt et al., 1997). Cell wall-

less mutants have a practical application in genetic transformation with foreign DNA and will be discussed in chapters 5 and 6.

Vegetative cells of *C. reinhardtii* are typically haploid, with one of two genetically set mating types: *mt+* or *mt⁻*. In nitrogen-deprived conditions, cells transform into plus or minus gametes that pair along the lengths of their flagella. This results in activation of the flagellar tips and in lysis of the cell wall. Fusion of the mating gametes begins at the anterior ends of the cells and a diploid zygote is formed. The zygote is transformed to a zygospore – a non-motile dormant structure. Due to the hard surrounding wall, the zygospore is capable of enduring adverse conditions. In the presence of light and under normal nitrogen conditions, the zygospore germinates and undergoes meiosis with the subsequent release of four haploid daughter cells (Harris, 1989).

Due to its short generation time simple sexual cycle and environmental adaptability, *C. reinhardtii* has functioned as a model organism for fundamental research for decades, particularly in studies of flagella (Asleson and Lefebvre, 1998), cell wall biogenesis (Woessner and Goodenough, 1992), phototaxis (Sineshchekov et al., 2002), gametogenesis and mating cycles (Lin and Goodenough, 2007). It has also been a principal organism in essential metabolic research on lipid metabolism (Tremolieres et al., 1995, Moellering et al., 2009), starch metabolism (Buleon et al., 1997) and circadian rhythms (Mergenhagen and Mergenhagen, 1989). Specific features such as function, biosynthesis, and regulation of the photosynthetic apparatus make *C. reinhardtii* an ideal species for research on photosynthesis (Harris, 1989). Photosynthetic mutants are easy to isolate since they are capable of heterotrophic growth on media where acetate is a source of fixed carbon. Moreover, any mutations are evident instantly since *C. reinhardtii* is haploid during vegetative growth and various phenotypes can be promptly detected as single colonies on solid medium.

Although *C. reinhardtii* has been employed as a model organism for studying lipid production, such as research on cytoplasmic lipid bodies and the influence of starchless mutants on algal lipid production (Wang et al., 2009), this species does not have a practical application in biofuel production. Its growth rate and lipid productivity is significantly lower when compared to other promising algal species currently exploited (Gouveia and Oliveira, 2009).

Figure 1-3 *Chlamydomonas sp.* cell.
Schematic diagram of the *Chlamydomonas* cell based on transmission electron micrographs presenting cellular organelles and an enlarged cross section of the flagellar axoneme with colour-coded structures. Figure replicated from (Merchant et al., 2007).
The picture far left was captured using a Leica optical microscope at the Cell and Developmental Biology Department at UCL.

Nevertheless, the biotechnological aspect of biofuel research in algae is still predominantly achieved in *C. reinhardtii* since genetic engineering techniques are widely available for this species. This topic will be discussed in detail later in this chapter and chapter 5.

1.2. Chloroplast and photosynthesis

1.2.1. Function and origin of plastids

The chloroplast of *Chlorella sorokiniana* and *Chlamydomonas reinhardtii* is a single intracellular organelle of high metabolic activity where numerous metabolic processes occur, examples being: sulphur and nitrogen metabolism, biosynthesis of amino acids, nucleotids, lipids, starch, fatty acids and pigments. Moreover, this organelle can be used by a cell as storage organ for produced molecules. For that reason using the more general term 'plastid' seems to be more appropriate when describing the metabolic activity of chloroplasts.

The plastid and its membranes play a significant role in phylogenetic analysis of algae since the number of membranes is a crucial feature in defining the evolutionary position of the analysed species and is associated with with the origin of plastids. It is generally accepted that plastids are originally derived from an oxygenic photosynthetic bacterium similar to modern-day cyanobacteria that became a part of algal cells through endosymbiosis with a non-photosynthetic eukaryotic cell (Delwiche, 1999, Martin and Kowallik, 1999, Graham and Wilcox, 2000). In primary endosymbiosis the cyanobacterial membrane forms the inner plastid membrane whereas the membrane of the eukaryotic vacuola transforms into the outer plastid membrane and gene transfer occurs between the cyanobacterium and nucleus. During secondary plastid endosymbiosis a non-photosynthetic eukaryotic cell engulfes a cell containing a primary plastid. As a result, the secondary plastid contains four membranes, the nucleus of the engulfed cell is transformed to a nucleomorph and gene transfer occurs between the nucleomorph and nucleus of the host cell. Ultimately, the nucleomorph may be lost completely, as seem for several algal groups containing secondary plastids. During tertiary plastid endosymbiosis, a non-photosynthetic eukaryotic cell engulfs a cell containing a secondary plastid. In theory up to six membranes could surround the tertiary

plastid, yet in dinoflagellates there are three membrane present which suggests too many membranes is not beneficial to a photosynthetic cell (Graham and Wilcox, 2000). The mechanism of primary, secondary, and tertiary plastid endosymbiosis is presented in Figure 1-4 whereas Table 1-1 presents detailed characteristics of each of algal groups based on their plastid features.

1.2.1. Plastid genome

The chloroplast genome of *C. reinhardtii* (see Figure 1-5) is a circular 203,395 bp structure composed of 99 genes including 72 protein-coding genes. It contains approximately 34.6% of G+C and this value is higher than for *C. vulgaris* which is 31.6% (Maul et al., 2002). The *C. reinhardtii* chloroplast genome consists of two copies of a ribosomal DNA-containing inverted repeat sequence of approximately 22 kb separated by two unique regions of approximately 80 kb and 78 kb. The gene composition within the inverted repeat is characteristic of other green algae and land plants (Maul et al., 2002, Green, 2011). The genes remaining within the plastid genome are the essential genes for the core components of the photosynthetic apparatus such as PSI, PSII, the cytochrome b_6f complex, ATP synthase and Rubisco, and also components of gene expression machinery such as the bacterial type RNA polymerase, rRNAs, ribosomal units, and tRNAs (Grossman et al., 2003, Maul et al., 2002). The thylakoid membrane complexes comprise of both nucleus- and plastid-encoded units (Race et al., 1999). The rates of synthesis of plastid-encoded proteins change in response to cell development and environmental factors, and this control is regulated by nucleus-encoded factors (Barkan and Goldschmidt-Clermont, 2000).

The chloroplast genome of *Chlorella vulgaris* is a circular structure of approximately 150 kb (see Figure 1-6) and it contains no large inverted repeat frequently observed in other *Chlorella* species (Wakasugi et al., 1997). Unfortunately, the chloroplast genome of *Chlorella sorokiniana* has not been fully sequenced yet. This represents a major obstacle to the development of a chloroplast genetic transformation system that would allow us to manipulate and control the metabolism of *C. sorokiniana*. My analysis of part of the *C. sorokiniana* genome and attempts to develop a suitable chloroplast transformation method are presented in chapter 4.

Figure 1-4 Plastid origins in eukaryotes.

Figure taken from (Delwiche, 1999)

Primary endosymbiosis of free-living cyanobacteria is believed to give rise to the plastid of three algal lineages: chlorophytes (green algae), glaucophytes, and rhodophytes (red algae) approximately 1.5 billion years ago. Subsequently, secondary endosymbiosis occurring approximately 1.3 billion years ago resulted in plastid formation through engulfment of a plastid-containing eukaryote by another eukaryote. This mechanism resulted in establishment of nucleomorph containing Cryptophyta and Chlorarachniophyta from the red algal and green algal lineage respectively, together with other algal groups without an extant nucleomorph. Tertiary endosymbiosis occurred when organisms arisen from secondary endosymbiosis lineage were engulfed and integrated by Dinophyta.

Table 1-1 Algal classification based on plastid features.

http://en.wikipedia.org/wiki/Algae

Supergroup affiliation	Members	Endosymbiont	Summary
Primoplantae/ Archaeplastida	Chlorophyta Rhodophyta Glaucophyta	Cyanobacteria	These algae have primary chloroplasts, i.e. the chloroplasts are surrounded by two membranes and probably developed through a single endosymbiotic event. The chloroplasts of red algae have chlorophylls a and c (often), and phycobilins, while those of green algae have chloroplasts with chlorophyll and b. Higher plants are pigmented similarly to green algae and probably developed from them, and thus Chlorophyta is a sister taxon to the plants; sometimes they are grouped as Viridiplantae.
Excavata and Rhizaria	Chlorarachnio-phytes Euglenids	Green algae	These groups have green chloroplasts containing chlorophylls a and b. Their chloroplasts are surrounded by four and three membranes, respectively, and were probably retained from ingested green algae. Chlorarachniophytes, which belong to the phylum Cercozoa, contain a small nucleomorph, which is a relict of the algae's nucleus. Euglenids, which belong to the phylum Euglenozoa, live primarily in freshwater and have chloroplasts with only three membranes. It has been suggested that the endosymbiotic green algae were acquired through myzocytosis rather than phagocytosis.
Chromista and Alveolata	Heterokonts Haptophyta Cryptomonads Dinoflagellates	Red algae	These groups have chloroplasts containing chlorophylls a and c, and phycobilins. The shape varies from organism to organism. They may be discoid, plate-like, reticulate, cup-shaped, spiral or ribbon shaped. They have one or more pyrenoids to preserve protein and starch. The latter chlorophyll type is not known from any prokaryotes or primary chloroplasts, but genetic similarities with red algae suggest a relationship. In the first three of these groups (Chromista), the chloroplast has four membranes, retaining a nucleomorph in Cryptomonads, and they likely share a common pigmented ancestor, although other evidence casts doubt on whether the Heterokonts, Haptophyta, and Cryptomonads are in fact more closely related to each other than to other groups. The typical Dinoflagellate chloroplast has three membranes, but there is considerable diversity in chloroplasts within the group, and it appears there were a number of endosymbiotic events. The Apicomplexa, a group of closely related parasites, also have plastids called apicoplasts. Apicoplasts are not photosynthetic but appear to have a common origin with Dinoflagellate chloroplasts.

Figure 1-5 Chloroplast genome of *C. reinhardtii*
The inner circle shows *Bam*HI and *Eco*RI restriction fragments mapped according to (Rochaix, 1980) and numbered according to (Grant et al., 1980). Position 0 is shown by an orange square near the 12 o'clock position. The second concentric circle indicates seven overlapping BAC clones that span the genome. The third circle shows genes and ORFs of unknown function, including those for which disruption experiments were unsuccessful. The outer circle shows genes of known or presumed function, with sequenced or hypothesized introns shown in olive green. Genes are color coded by function, as shown at bottom.
Figure and legend reproduced from (Maul et al., 2002)

Figure 1-6 Gene map of the *C. vulgaris* C-27 chloroplast genome.
Genes shown on the inside of the circle are transcribed clockwise, and genes on the outside are transcribed counterclockwise. ORFs of ≥60 codons are included. Asterisks denote split genes. Nucleotide positions are numbered counterclockwise from the arrow (position 1) in ycf10.

Figure and legend reproduced from (Wakasugi et al., 1997)

Since plastids evolved from cyanobacterial endosymbionts, a number of similarities have been discovered among the essential components of the transcription-translation apparatus between plastids and bacterial cells (Beligni et al., 2004). There are also numerous differences between plastid and bacterial metabolism that arose during billion of years of evolution. Firstly there are substantial differences in gene complexity. Algal plastids contain a unique circular double stranded DNA (Manning et al., 1971) that is considerably reduced in size in respect to the free-living cyanobacteria (Beligni et al., 2004). In addition to gene transfer from plastid to nucleus, there are two other identified mechanisms that are responsible for this reduction: gene loss and gene substitution. Gene loss occurred as a result of a lack selective advantage to retain such genes once the cyanobacterium became an endosymbiont, for example the genes coding for the cyanobacterial cell wall (Delwiche, 1999). In the gene transfer mechanism, a chloroplast gene firstly undergoes duplication, one gene copy is transferred to nucleus, possibly via RNA-mediated mechanisms, and subsequently the chloroplast copy is deleted (Nugent and Palmer, 1991). There is limited information available on the mechanism of gene substitution, yet this process has been assigned as an important step in plastid evolution (Martin and Schnarrenberger, 1997).

Post-transcriptional steps within the plastids is regulated by nuclear-encoded factors and frequently dependent on environmental influences (Choquet and Wollman, 2002, Somanchi et al., 2005, Zerges et al., 1997). However, regardless of the listed differences, previous investigations on chloroplast regulatory regions indicated that certain promoters/5'UTRs are also functional in *E. coli* (Gatenby et al., 1988, Bateman and Purton, 2000). This trend can be explained through the presence of certain similarities within the 5'UTR regions of chloroplasts and cyanobacteria since some of the chloroplast 5' regulatory sections contain Shine-Dalgarno (SD) sequences (Drechsel and Bock, 2011). SD sequences, vital for the efficient translation initiation in bacteria, are usually GGAGG sequences placed approximately 7±2 nt from AUG. This sequence enables ribosome binding to mRNA through base-pairing to the 3' end of the 16S rRNA (Shine and Dalgarno, 1975). The presence of SD-like sequences is required within the 5'UTR region of some chloroplast genes of *C. reinhardtii* such as highly expressed *psbA* or *psbD* (Nickelsen et al., 1999) however, for the majority of chloroplast-encoded genes the presence of the SD sequences is not essential (Zerges, 2000).

1.2.2. Photosynthesis

Photosynthesis is a process of photochemical energy transduction that converts the light energy into chemical energy. The initial step of this process is located in the reaction centre complexes of Photosystem I (PSI) and Photosystem II (PSII). Both PSI and PSII are associated with thylakoid membranes and are formed of a nuclear-encoded chlorophyll a/b-binding light harvesting complex (LHC) and the plastid-encoded core complex. In PSI and PSII, the absorption of a photon by the LHC results in the rapid transfer of excitation energy to the reaction centre, where a photochemical charge separation occurs, followed by further separation of charges using an electron transfer chain (Larkum et al., 2003). Primary donors of PSI and PSII – chlorophyll P700 and P680 respectively – interact with primary acceptors in the electron transfer event.

The primary photochemical reaction of PSII is associated with the linear electron transfer from the reaction centre chlorophyll P680 to an acceptor molecule – pheophytin that reduces the first stable acceptor – plastoquinone pool, and then to the cytochrome b_6/f complex and plastocyanin. Upon the capture of light, the excited P700 passes the electrons through a series of intermediate stages and ultimately the electrons are donated to $NADP^+$ via ferredoxin and ferredoxin-$NADP^+$ reductase for the creation of the NADPH (Srirangan et al., 2011). $P680^+$ accepts electrons from a donor Z that is reduced by the water splitting system oxidising water to molecular oxygen, H^+ and electrons (Harris, 1989). In the cyclic electron transfer chain, the solar energy captured by PSI antenna molecules is transferred to P700 reaction centre and the emitted electron is captured by the primary acceptor and transferred to ferredoxin. Reduced ferredoxin returns the electron to the reaction centre via an electron transport pathway. The overall organization of the electron transfer chain is presented and described in Figure 1-7.

Both linear and cyclic electron transfers are paired with H^+ transfer from the chloroplast stroma to the thylakoid lumen, which builds a proton gradient for ATP synthesis (Srirangan et al., 2011). The generated ATP and NADPH are further used to transform atmospheric CO_2 into organic compounds by the dark reaction of photosynthesis of the Calvin cycle (Benson and Calvin, 1950).

Figure 1-7 Schematic illustration of linear (a) and cyclic (b, c) photosynthetic electron transport pathways in *C. reinhardtii.*

During linear electron flow, electrons are extracted from water at photosystem 2 (PS2) and transported to photosystem 1 (PS1) via plastoquinone (PQ), the cytochrome b6/f complex (Cyt b_6/f) and plastocyanin (PC). PS1 donates electrons to NADP+ via ferredoxin (PetF) and ferredoxin-NADP+ reductase (FNR). A proton-gradient is built up during the electron transport and used by an ATP-synthase (ATPase) to generate ATP. Both NADPH and ATP are needed for CO_2-fixation and triose-phosphate (Triose-P) generation by the Calvin cycle (Rbc = ribulose-1,5-bisphosphate carboxylase/oxygenase, Rubisco). Linear electron flow is associated with state 1 conditions, in which the light harvesting complexes (LHC) of PS2 (LHCII) are associated with PS2. (b) In *C. reinhardtii*, the antimycin A sensitive, ferredoxin dependent pathway of cyclic electron flow can be catalyzed by a supercomplex containing PS1, Cyt b_6/f, FNR and PGRL1. (c) The NAD(P)H-dehydrogenase Nda2 is supposed to reduce the PQ-pool during the antimycin A insensitive pathway of cyclic electron flow, using stromal reductants which have been generated by PS1, PetF and FNR. (b) and (c) In *C. reinhardtii*, cyclic electron transport is associated with state transitions (LHCII at PS1).

Figure and legend reproduced from (Hemschemeier and Happe, 2011)

In response to changing light conditions, two adaptation systems called rapid and long-term mechanisms are activated. The latter process involves genetic regulation of the overall chlorophyll antenna size and will be discussed in detail in chapter 3. A rapid mechanism activated in response to changing light conditions called State 1–State 2 transitions or 'state transitions' is based on rearrangement within the photosynthetic light apparatus. It was first discovered in green (Bonaventura and Myers, 1969) and red (Murata, 1969) algae and although there are significant differences between algal and plant light harvesting complexes, the concept of state transition is essentially analogous in both groups (Mullineaux and Emlyn-Jones, 2005). In principle, state transition operates as a protection mechanism initiated by changing light conditions to balance excitation of the two photosystems. In conditions that lead to over-excitation of PSII in respect to PSI, a transition to State 2 is induced and more absorbed excitation energy is transferred to PSI than to PSII. A transition to State 1 is induced when excess excitation from PSI is diverted to PSII. On a molecular level, State 1 to State 2 transition involves activation of the LHCII kinase when plastoquinone is in a reduced form as a result of PSII transfers electrons to the plastoquinone pool faster than PSI can utilize them. Similarly, a State 2 to State 1 transition involves inactivation of the LHCII kinase when plastoquinone is oxidised because PSI oxidizes the pool faster than PSII can reduce it.

1.3. Biotechnology of microalgae

1.3.1. Advantages of microalgal biotechnological applications

Important features characteristic for microalgae such as high growth rate, simplicity of cultivation together with the possibility to genetically manipulate fundamental aspects of cell metabolism (Surzycki et al., 2009) provide microalgae with the potential to revolutionize the biotechnology field. Moreover, algal photosynthetic efficiency, due to their simple cellular structure, is approximately three times higher than that of higher plants (Shimizu, 1996). Algal generation time by far exceeds the growth rates of higher plants and also the amount of accumulated proteins in algal cells exceeds the protein content of any edible tissues of higher plants (Passwater and Solomon, 1997). Yet, the most important feature enabling constant progress in algal biotechnology applications is the availability of robust and well-established transformation systems that facilitates rapid generation of transformed lines (Purton, 2006). Many green algae are considered as

GRAS organisms ("Generally Recognised As Safe"), which means the purification of the bioavailable products can be substantially reduced or eliminated altogether. Furthermore, microalgae can be cultivated in full containment, which reduces the public concern regarding the risk of environmental contamination by genetically modified algae (GMA). This section of the Introduction chapter aims to review the current situation of microalgal biotechnology.

1.3.2. Biofuels

Finding a suitable, clean, and renewable energy source to replace fossil fuels is one of the most challenging problems mankind is going to face in the near future. This subject is unsurprisingly associated with economic development and global stability and requires long-term strategic planning (Mata et al., 2010). Many options for renewable energy sources are currently being considered. In additional to the exploitation of solar energy, thermal and photovoltaic, wind and hydroelectrics, biofuels are growing more popular each year with bioethanol and biodiesel being the most common types of biofuels that are able to supplement conventional gasoline and diesel, respectively (Mata et al., 2010).

Biodiesel is a type of fuel obtained following the transesterification of fats and oils (see chapter 5 for further details). Increasing interest in biodiesel as a renewable energy source is related to its properties of being biodegradable, and the total content of solid particles and polycyclic aromatic hydrocarbons in exhaust fumes being significantly lower in respect to typical fossil-derived diesel fuel (Zajac, 2008). Additionally, biodiesel can offset the accumulation of greenhouse gas such as carbon monoxide, hydrocarbons and SO_X (Gouveia and Oliveira, 2009) and also carbon dioxide since the synthesis of plant oils used to make biodiesel involves the sequestration of CO_2 by photosynthesis.

At present, production of biodiesel is almost exclusively from vegetable oils. However, since fossil fuels are considered to be originally derived from microalgae, it is no surprise there is an increased interest in exploiting microalgae for production of various types of biofuel. Furthermore, when considering biomass and oil productivity, microalgae are shown to be the only resource of biodiesel production that can hypothetically replace fossil diesel. Table 1-2 presents potential land requirement for biodiesel production from various crop species compared to prospective microalgal productivity.

Table 1-2 Biodiesel production from various types of crop.

Replicated from (Chisti, 2007)

Crop	Oil yield (L/ha)	Land area needed (M ha) - for meeting 50% of all US transport fuel needs	% of existing US cropping area
Corn	172	1540	846
Soybean	446	594	326
Canola	1190	223	122
Jatropha	1892	140	77
Coconut	2689	99	54
Oil palm	5950	45	24
Microalgae (70% oil)	136,900	2	1.1
Microalgae (30% oil)	58,700	4.5	2.5

There are several prospective advantages of microalgal systems for biofuel production over crop-derived methods, yet there are several drawbacks associated with microalgae-derived biofuels since they require the application of novel technologies. Large-scale algal cultures have the potential to provide all year round production with a possibility to use waste or seawater for culturing, yet generating microalgal biomass is currently more expensive than growing crops. Table 1-3 lists the advantages and limitations of microalgal biofuel technology.

Although microalgae are relatively cheap and easy to cultivate in laboratory conditions, it is clear that the main obstacle for exploiting algae for biofuel production is the costly scale-up to cultivation systems such as open ponds or photobioreactors. One approach to overcome this problem is to choose a suitable fast-growing strain that it is possible to cultivate in considerable volumes in outdoor conditions. Genetic engineering approaches could then be employed to further improve the productivity of these strains. Table 1-4 presents algal species currently being considered for bioenergy applications. Two genera chosen for research on biofuel production presented in this thesis work are either easy to maintain on large scale (*Chlorella*) or have establishment genetic engineering methodologies (*Chlamydomonas*).

Table 1-3 Advantages and limitations of microalgae over plants as a source of biofuels.
(Rodolfi et al., 2009, Gouveia and Oliveira, 2009)

Advantages	**Limitations**
High oil yield per cultivated area (Table 1-2, Table 1-5)	Selection for the most productive strain is very challenging
Water requirements relatively low	Difficulties of culturing the selected strain in outdoor conditions
Possibility of use non-arable land additionally to seawater and wastewater	Limited availability of productivity data on scaled-up cultures
Potential of combining the biomass production with bio-fixation of waste CO_2	High energy input required for culture mixing, harvesting and dewatering of the biomass
Biomass after oil extraction can be used as feed, fertilizer or to produce ethanol or methane	
Biochemical composition and oil content can be easily modified and enhanced	

The second approach to reduce the costs of microalgal biofuel systems is choosing a strain characterized by high lipid content. Table 1-5 presents oil content of the selected microalgal species that could be potentially used in biofuel technologies. The yield of the biodiesel does not solely depend on oil content, but also on the biomass productivity, for instance, doubling time of *Botryococcus braunii* is approximately 4 days (Zhang et al., 2011a) and in spite of its high oil content (up to 70% of dry weight) the biomass productivity is not sufficient for industrial applications. The selection procedure for the best performing algal strain for biofuel application will be discussed further in chapter 3.

It has been common practice to use stress conditions such as deprivation of nitrogen in growth medium to increase lipid content (Rodolfi et al., 2009). The principle of this process is based on the mechanism where deficiency of nitrogen essential for protein synthesis triggers the storage of carbon accumulated in photosynthesis as triacylglycerides (TAGs) or starch (Scott et al., 2010). Figure 1-8 presents the basic overview of lipid biosynthesis by phototrophically and heterotrophically-grown cultures. The regulation of lipid biosynthesis and attempts to improve biofuel production in algae will be presented and discussed further in chapter 5.

Table 1-4 Products currently obtained from microalgae for bioenergy and the organic chemicals industry.

Adapted from (Larkum et al., 2012)

Algal species	Algal class	Product(s)	Culture technique	Advantages or drawbacks	References
Botryococcus braunii	*Chlorophyceae*	Triterpene oils	Photobio-reactor	Whole genome available	(Metzger and Largeau, 2005)
Chlorella spp	*Chlorophyceae*	Carbohydrates protein	Ponds, photobio-reactor	Widespread and adaptable	(Iwamoto, 2007)
C. reinhardtii	*Chlorophyceae*	Oils, carbohydrates, hydrogen and methane	Photobio-reactor	Transformable	(James et al., 2010)
Dunaliella salina	*Chlorophyceae*	β-Carotene	Brackish seawater ponds	Needs high salinity	(Amotz, 2004)
Nannochloropsis	*Eustigmatophyceae*	Polyunsaturated fatty acids	Seawater ponds	Lipid bodies produced under nitrogen stress	(Zittelli, 2004)
Ostreococcus tauri	*Chlorophyceae*	Oils	Photobio-reactor	Smallest known microalga, whole genome available	(Derelle et al., 2006, Shi et al., 2011)
Pavlova lutheri	*Prymnesio-phyceae*	Fatty acids, aquaculture feedstock	Photobio-reactor	Little used	(Andersen, 2004)
Arthrospira platensis (Spirulina)	Cyanobacteria	Health food	Ponds, photobio-reactor	Filamentous morphology	(Venkataraman, 1997)
Synechocystis and *Synecococcus*	Cyanobacteria	Isoprenes, oils	Photobio-reactor	Whole genomes available, transformable	(Lindberg et al., 2010)

Table 1-5 Oil content of various microalgal species.

Adapted form (Chisti, 2007, Gouveia and Oliveira, 2009)

Microalgal species	Oil content (% dry weight)
Botryococcus braunii	25–75
Chlorella sp.	28–63
Crypthecodinium cohnii	20
Cylindrotheca sp.	16–37
Dunaliella primolecta	23
Isochrysis sp.	25–33
Monallanthus salina	> 20
Nannochloris sp.	20–35
Nannochloropsis sp.	31–68
Neochloris oleoabundans	35–65
Nitzschia sp.	45–47
Phaeodactylum tricornutum	20–30
Schizochytrium sp.	50–77
Scenedesmus obliquus	35–55
Scenedesmus dimorphus	16–40
Tetraselmis sueica	15–23

Figure 1-8 Basic overview of the pathway of carbon capture and lipid biosynthesis.

Precursor fatty acids are synthesized de novo in the chloroplast, using either carbon fixed during photosynthesis, or from an exogenous supply of organic carbon; the exact nature of what enters the chloroplast is unknown in algae (dashed line). Free fatty acids are exported from the chloroplast and then converted to TAGs in the endoplasmic reticulum (ER), where they bud off into oil bodies in the cytosol. Key: (i) = acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS); (ii) = fatty acid thioesterases and acyl-CoA synthetases; (iii) = TAG biosynthesis enzymes, including acyl-CoA:diacylglycerol acyltransferase (DGAT); (iv) = oil body formation; and (v) = ADP-glucose pyrophosphorylase and starch synthase.

Figure and legend reproduced from (Scott et al., 2010)

1.3.3. Recombinant products

Although recombinant proteins produced using transgenic algae are not currently available on the market, the potential of generating therapeutic proteins on an industrial scale is increasing. Over 30 commercially important proteins have been successfully expressed in transgenic algae (Specht et al., 2010, Gong et al., 2011b) and Table 1-6 presents a list of therapeutic proteins expressed in algae in recent years. This progress was predominantly achieved in *Chlamydomonas reinhardtii* since this alga is has comparatively straightforward genetic manipulation techniques for the nuclear, chloroplast and mitochondrial genomes. The level of expression of recombinant proteins obtained varies for different therapeutics and transgenic species. In *C. reinhardtii* this amount reaches between 0.16% and 10% of total soluble protein (TSP) (Mayfield et al., 2003, Sun et al., 2003, Surzycki et al., 2009, Tran et al., 2009, Rasala et al., 2010). The expression of chloroplast-encoded recombinant proteins typically results in higher levels when compared to the nucleus-encoded proteins (Mayfield et al., 2007).

Human antiherpes large single chain antibody (HCV8-lsc) was the first human protein expressed in algae (Mayfield et al., 2003). It comprised of an IgA heavy chain protein fused to the light chain, was fully soluble, active and correctly assembled in the chloroplast. Also vaccine candidates have been produced in transgenic algae. For instance, a fully functional CTB-VP1 fusion protein (cholera toxin B (CTB) subunit fused to foot and mouth disease antigen VP1), synthesized in the chloroplast of *C. reinhardtii* gives the potential to be delivered orally (Sun et al., 2003). Moreover, various other functional therapeutics such as anti-cancer molecules (Rasala et al., 2010), anti-viral molecules (Yang et al., 2006b), cytokines, and hormones (Rasala et al., 2010) have been reported to be successfully expressed in *C. reinhardtii*.

Table 1-6 Recombinant proteins recently expressed in algae.

Modified from (Specht et al., 2010)

Protein name	Protein type	Comments	Ref.
HSV8-lsc	Human antiherpes large single chain antibody	First mammalian protein expressed	(Mayfield et al., 2003)
CTB- VP1	Cholera toxin B (CTB) subunit fused to foot and mouth disease VP1 (FMDV VP1)	CTB is a potent mucosal adjuvant; FMDV VP1 fusion to CTB improves the vaccine against foot and mouth disease	(Sun et al., 2003)
HSV8-scFv	Human antiherpes large single chain fragment variable antibody	Classic single-chain antibody	(Mayfield and Franklin, 2005)
hMT-2	Human metallothionein-2	Resistance to UV-B exposure	(Zhang et al., 2006)
hTRAIL	Human Tumor necrosis factor-related Apoptosis-Inducing Ligand	Anti-viral and anti-tumour properties	(Yang et al., 2006a)
M-SAA	Bovine mammary-associated serum amyloid	prophylaxis of bacterial and viral infections	(Manuell et al., 2007)
CSFV-E2	Classical swine fever virus structural protein E2	Vaccine candidate	(He et al., 2007)
hGAD65	Human glutamic acid decarboxylase 65	Marker and treatment of type 1 diabetes	(Wang et al., 2008)
ARS2-crEpo-his_6	Human erythropoietin fused to extracellular enzyme arylsulfatase gene and a C-terminal hexahistidine-tag	Secretion	(Eichler-Stahlberg et al., 2009)
83K7C	Full-length IgG1 human monoclonal antibody against anthrax protective antigen 83		(Tran et al., 2009)
IgG1	Human and mouse IgG1 antibodies		(Tran et al., 2009)
CTB-D2	D2 fibronectin-binding domain of Staphylococcus aureus fused with the cholera toxin B subunit	Oral vaccine candidate	(Dreesen et al., 2010)
EPO	Erythropoietin	Signal peptide	(Rasala et al., 2010)
10NF3, 14NF3	Domains 10 and 14 of human fibronectin	Important protein in cell adhesion, migration, growth and differentiation	(Rasala et al., 2010)
M-SAA-IFNβ1	Interferon β	Treatment of multiple sclerosis (MS)	(Rasala et al., 2010)
Proinsulin		Hormone regulating blood sugar level, used in treatment type I diabetes	(Rasala et al., 2010)
VEGF	Human vascular endothelial growth factor isoform 121	Protein important in treatment of pulmonary emphysem, erectile dysfunction and depression	(Rasala et al., 2010)
HMGB1	High mobility group protein B1	Important protein in wound healing and anti-cancer therapies	(Rasala et al., 2010)

1.3.4. Bioremediation

In the natural environment, chemical and biological contaminants have been continuously increasing due to various applications of agriculture, urbanization and industrialization. Since microalgae are known to play an important role in the natural environment, there is growing research interest in their practical application to combat various environmental issues such as the greenhouse effect and waste treatments (Ghasemi et al., 2011). With increasing knowledge of their physiology, metabolism and environmental adaptation, microalgae have the potential to be applied in diverse projects. For instance, they can be used as pollution control agents (Perales-Vela et al., 2006) or biosensors for detection of toxic compounds (Zamaleeva et al., 2011). There are two main principles behind applications of living cells as pollution control agents; firstly is to convert toxic chemicals into non-toxic compounds and secondly to transform them into products of an economic value (Burton, 2001).

Microalgae are recognized for their ability as efficient microorganisms in the accumulation and degrading of numerous kinds of environmental pollutants such as pesticides (Megharaj et al., 1994), crude oil (Sorkhoh et al., 1992), phenolic compounds (Ellis, 1977) and xenobiotics (Megharaj et al., 1987). Moreover, microalgal strains, particularly *Chlorella* species, are primarily used in a tertiary waste treatment (Valderrama et al., 2002).

1.3.5. Feed source for aquaculture

Aquaculture, defined as farming of aquatic organisms, contributed to 43% of the total yield of wild fisheries in 2007 (Bostock et al., 2010). Furthermore, microalgal production as a feed for aquaculture reached 1000 tonnes in 1999, and is expected to expand in the near future (Spolaore et al., 2006). The main aspect associated with microalgal applications for aquaculture is related to their nutritional content (polyunsaturated fatty acid, protein and vitamin content) since they are an important part of the balanced diet of molluscs and peneid shrimp or indirectly as food. Table 1-7 presents the most common species and their applications used in aquaculture.

Table 1-7 Commercial algal cultures and their applications.

Replicated from (Hemaiswarya et al., 2011)

Algae	Morphology	Application
Nannochloropsis	Small green algae	Growing rotifers and in fin fish hatcheries, used in reef tanks for feeding corals and other filter feeders, very high EPA level
Pavlova	Small golden-brown flagellate, very difficult to grow so it is not produced by many hatcheries	Used to increase the DHA/EPA levels in broodstock, oysters, clams, mussels and scallops; it is popular with cold water fish hatcheries (cod) for enriching rotifers
Isochrysis	Small golden-brown flagellate	Enrichment of zooplankton such as *Artemia*, used in shellfish hatcheries and used in some shrimp hatcheries, good size for feeding brine shrimp and copepods, oysters, clams, mussels, and scallops
Tetraselmis	Large green flagellate	Excellent feed for larval shrimps and contains natural amino acids that stimulate feeding in marine animals, used in conjunction with *Nannochloropsis* for producing rotifers, good size for feeding brine shrimp, standard feed for oysters, clams, mussels, and scallops, excellent feed for increasing growth rates and fighting zoea syndrome
Thalassiosira weissflogii	Large diatom	Used in the shrimp and shellfish larviculture, considered by several hatcheries to be the single best alga for larval shrimps, also good for feeding copepods and brine shrimps, post-set (200 l and larger) oysters, clams, mussels, and scallops for broodstock conditioning
Dunaliella	Small green flagellate	Used to increase vitamin levels in some shrimp hatcheries and also for the coloration
Chaetoceros	Diatom	Used to increase vitamin levels in some shrimp hatcheries

1.4. Genetic tools and techniques for transformation of microalgae

1.4.1. Transformation techniques

A prerequisite for genetic engineering of any algal species is the establishment of reliable and reproducible transformation systems, ideally for both the nuclear and chloroplast genomes. To date, over 22 algal species have been successfully transformed using either nuclear or chloroplast transformation methods (Gong et al., 2011b).

Currently the most efficient method for the delivery of foreign DNA into the algal cell is microparticle bombardment using the biolistic particle gun method (Heiser, 1992) with a helium-powered device. Microcarrier particles – either tungsten or gold – are coated with a plasmid containing a foreign gene and fired at an algal lawn spread on a nutrient agar plate. The DNA coated particles are accelerated to a very high velocity thus they are able to pass through the algal cells, leaving the DNA inside. Microparticle bombardment is particularly effective for DNA delivery into the chloroplast where the DNA must be traverse multiple membranes. Once inside the organelle the cloned DNA integrates into the chloroplast genome via homologous recombination between corresponding homologous sequences of cpDNA and the plasmid.

For nuclear transformation of *C. reinhardtii* agitation of a cell/DNA suspension with ~0.4 mm glass beads has proved the most popular since it was first applied (Kindle, 1990). The method is simple with high rates of transformed cells and stability of transformants in the absence of selection is high (Kindle, 1990). In contrast, the electroporation method requires a complex apparatus that uses electric pulses to deliver the foreign DNA into the cell. This method involves several optimisation steps and using a cell wall-less strain is required. However, it has the highest rate of nuclear transformation for *C. reinhardtii* out of the three transformation systems described above (Coll, 2006). Other methods that have proved to be suitable for nuclear transformation of microalgae involve agitation with silicon carbide whiskers (Dunahay, 1993, Te et al., 1998), sonication in the presence of polyethylene glycol (Jarvis and Brown, 1991) or co-cultivation with the plant pathogen *Agrobacterium tumefaciens* (Kumar et al., 2004b). The list of available methods for algal nuclear transformation and examples of successful events is presented in Table 1-8.

Table 1-8 Available methods of nuclear transformation and examples of successful transformation events performed on microalgal species.

Method	Species	Reference
Microprojectile bombardment	*Chlamydomonas reinhardtii*	(Debuchy et al., 1989)
	Dunaliella salina	(Feng et al., 2008)
	Chlorella sorokiniana	(Dawson et al., 1997)
	Haematococcus pluvialis	(Teng et al., 2002)
	Volvox carteri	(Schiedlmeier et al., 1994)
	Gonium pectorale	(Lerche and Hallmann, 2009)
	Cyclotella cryptyca	(Dunahay et al., 1995)
	Navicula saprophila	(Dunahay et al., 1995)
	Phaeodactylum tricornutum	(Apt et al., 1996)
	Cylindrotheca fusiformes	(Fischer et al., 1999)
	Thalassiosira pseudonana	(Poulsen et al., 2006)
	Closterium peracerosum	(Abe et al., 2011)
	Chaetoceros sp.	(Miyagawa-Yamaguchi et al., 2011)
Electroporation	*Chlamydomonas reinhardtii*	(Shimogawara et al., 1998)
	Dunaliella salina	(Lu et al., 2011)
	Chlorella sp	(Wang et al., 2007)
	Cyanidoschyzon merolae	(Minoda et al., 2004)
	Nannochloropsis sp.	(Kilian et al., 2011)
Glass beads	*Chlamydomonas reinhardtii*	(Inoue et al., 1997)
	Platymonas subcordiformis	(Cui et al., 2012)
	Dunaliella salina	(Feng et al., 2008)
Silicon carbide fibers	*Chlamydomonas reinhardtii*	(Dunahay, 1993)
	Symbiodinium microadriaticum	(Te et al., 1998)
	Amphidinium sp.	(Te et al., 1998)
Agrobacterium tumefaciens	*Chlamydomonas reinhardtii*	(Kumar et al., 2004b)
	Haematococcus pluvialis	(Kathiresan et al., 2009)
Polyethylene glycol	*Chlorella ellipsoidea*	(Jarvis and Brown, 1991)

1.4.2. Chloroplast genetic engineering

The first transgenic chloroplast was created using a photosynthetic mutant of *C. reinhardtii* containing a large deletion in the chloroplast *atpB* gene. This mutant was bombarded with the wild type version of the *atpB* gene cloned into a plasmid, and selection was based on the restoration of photosynthesis and therefore an ability to grow on a minimal medium (Boynton et al., 1988).

Selection of transformants using the mechanism of restoration of photosynthesis was also used in the presented work. The Purton lab has created a suitable recipient line containing a large deletion in another gene essential in photosynthesis – *psbH*, what resulted in photosynthetic deficiency of the recipient cell line. Additionally, the recipient line contains the *aadA* marker that confers the antibiotic resistance. Figure 1-9 presents the mechanism of the transgene integration within chloroplast genome developed in the Purton lab where the integration of the gene of interest (GOI) occurs via homologous recombination of the flanking sequences within the chloroplast genome and the transformation vector and is associated with restoration of the fully functional *psbH* gene. Chloroplast genetic engineering will be further discussed in chapter 4, chapter 5 and chapter 6.

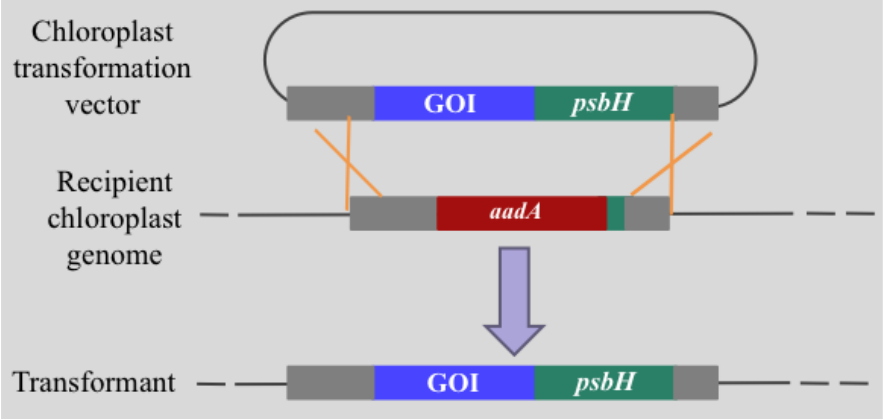


Figure 1-9 Schematic diagram of the transgene (GOI) insertion via homologous recombination during chloroplast transformation.

The recipient cell line comprises of the *aadA* marker incorporated in the genome and contains a truncated *psbH* gene (green box). The chloroplast transformation vector contains the transgene cassette (GOI) and the *psbH* gene. Homologous recombination occurs between flanking sequences (grey boxes) present in the chloroplast expression vector and the recipient chloroplast genome what leads to creation of the transformant line, where the *aadA* marker is replaced by the GOI and the *psbH* gene.

1.4.3. Nuclear genetic engineering

The first reliable report documenting successful nuclear transformation in *C. reinhardtii* was performed on an arginine-requiring mutant (*arg7*) deficient in argininosuccinate lyase (ASL) activity. The mutant was rescued to arginine prototrophy by microparticle bombardment with a plasmid carrying the wild-type ASL gene (Debuchy et al., 1989). Nuclear transformation (Figure 1-10) generally occurs as a random insertion of the transgene within the nuclear genome (Gumpel and Purton, 1994). This random integration has been exploited as a method of insertional mutagenesis allowing the cloning and sequencing of unknown disrupted nuclear genes from mutants with interesting phenotypes (Lumbreras and Purton, 1998). Despite the very low level of homologous recombination (HR) in the *C. reinhardtii* nucleus, there are reports of efficient HR in the nucleus of other algal species including *Nannochloropsis sp*. and *Cyaniodischyzon merolae* (Gumpel et al., 1994, Kilian et al., 2011, Minoda et al., 2004). An additional advantage of these nuclear transformation methods is the ability to transform nuclear genome with large DNA fragments such as 50 kB cosmid clones (Purton and Rochaix, 1994). Nuclear genetic engineering method will be further discussed in chapter 4.

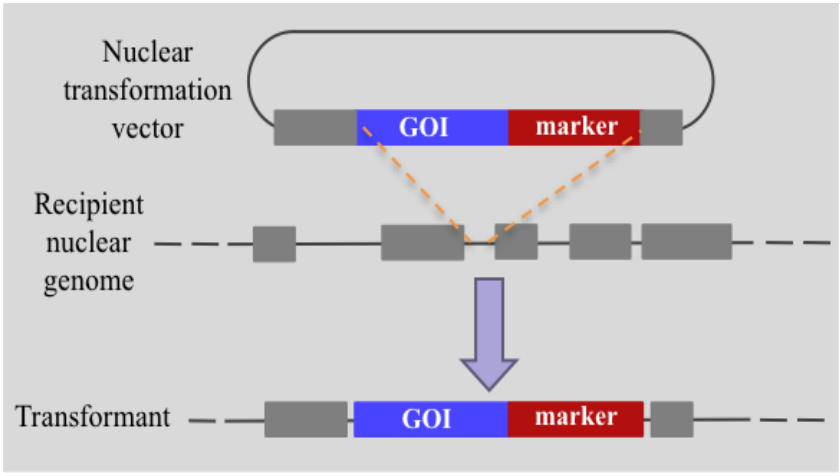


Figure 1-10 Schematic diagram of the random transgene (GOI) insertion during the nuclear transformation.

1.4.4. Chloroplast versus nuclear transformation

There are several advantages of chloroplast transformation over nuclear transformation. Firstly, the chloroplast genome is present in up to a hundred copies therefore a foreign gene within the chloroplast genome is present at a significantly higher copy number than one inserted into the nuclear genome. Moreover, chloroplasts are reported to be able to accumulate high level of foreign proteins (Miyagawa et al., 2001), without interference with the chloroplast metabolism. In addition, foreign genes are integrated into the chloroplast via homologous recombination (Heifetz, 2000) whereas nuclear insertion in *C. reinhardtii* occurs predominantly via random integration (Debuchy et al., 1989). This process enables a direct gene insertion into the known position within the genome, hence the number of chloroplast transformants that need to be screen for a successful transgenic line is significantly fewer than for nuclear transformant lines. Chloroplast transgenes are not subjected to gene silencing and multiple transgenes can be expressed as a single operon. In addition, chloroplast genes are maternally inherited from the mating-type *minus* parent and this allows containment of transgenes as is also seen in higher plants (Bateman and Purton, 2000, Purton, 2006, Verma and Daniell, 2007). Table 1-9 presents the evaluation of the chloroplast and nuclear transformation methods.

Table 1-9 Comparison of the chloroplast and nuclear transformation methods.

	Nuclear transformation	Chloroplast transformation
Number of plastome copies	Number of chromosomes is specific for each species, usually diploid; May result in multiple integration of transgenes	10–100 copies of plastome resulting in multiple copies of transgene within the transformed cell
Efficiency of the transgene expression	Controlled by gene regulation, protein accumulation can be limited	Multiple transgene copies result in higher transcription level and increased protein accumulation
Arrangement and transcription of genes	Each transgene is independently inserted into genome	Transgenes can be arranged in operons introduced into the genome during one transformation event
Gene silencing	Reported on transcriptional and post-transcriptional level	Not reported
Position effect	Random insertion of transgenes results in inconsistent expression level of transgene between transformants	No position effect reported; consistent expression level of transgene between transformants
Glycosylation	Analogous to the pattern observed for higher plants	Absent
Gene containment	Easily transferred to other species	Genes can be transferred via maternal inheritance
Transgenic lines	Variable gene expression	Uniform gene expression
Type of inheritance of the transgene	Mendelian	Maternal
Organelle of the transgene expression	Cytosol, chloroplast; can also be secreted	Chloroplast
Toxic and foreign proteins	The accumulation of toxic proteins in the cytosol can be lethal	The lethal effect is minimized thanks to the chloroplast compartmentation

1.5. Objectives of this work

The major aim of this work was to develop novel platforms in microalgal biotechnology for both high and low value products. Since transformation methods have been successfully developed for *Chlamydomonas reinhardtii* in the Purton lab, the research was originally designed for this algal species with the aim to transfer this technology to a "superior alga" characterized by significantly faster growth rate than *C. reinhardtii.*

The research described in this thesis addressed the following goals:

1. To identify and characterize a very fast growing algal strain and to optimise its productivity.
2. To develop suitable chloroplast and nuclear transformation methodologies for the identified microalga.
3. To investigate the possibility of chloroplast metabolic engineering for alkane production in *C. reinhardtii*
4. To produce high value recombinant proteins such as an HPV vaccine component and the anti-HIV microbicide candidate cyanovirin-N in the *C. reinhardtii* chloroplast.

CHAPTER 2 MATERIALS AND METHODS

2.1 Strains and culture conditions

2.1.1 *Chlorella sp*.

Chlorella vulgaris was obtained from the Culture Collection of Algae and Protozoa [Oban, Scotland; www.ccap.ac.uk] and *Chlorella sorokiniana* UTEX1230 was obtained from the UTEX Culture Collection of Algae [The University of Texas at Austin, USA], whereas *Chlorella sorokiniana* H-1983 and H-1986 strains were provided by the Culture Collection of Algae of Charles University in Prague [Czech Republic].

Additional strains used in this work included:

1. Six mutants of *C. sorokiniana* UTEX1230 containing truncated chlorophyll antenna (TAM1, TAM2, TAM3, TAM4, TAM5, and TAM6)
2. Four photosynthesis-deficient mutants of *C. sorokiniana* UTEX1230 (PSM315, PSM941, PSM1345, and PSM1645)
3. Seven cycloheximide-resistant mutants of *C. sorokiniana* UTEX1230

The algae were maintained on Tris-acetate phosphate (TAP) medium (Table 2-2). However for some applications the following media were used: *Euglena gracilis*: Jaworski's Medium (EG:JM) prepared by mixing EG and JM media in 1:1 proportions as detailed in Table 2-1 or high salt minimal (HSM, Table 2-2). The 3N-BBM+V (Bold Basal Medium with 3-fold Nitrogen and Vitamins) medium (Starr and Zeikus, 1993) was tested during the optimization of the growth of *C. sorokiniana* UTEX1230 in minimal medium. In addition, for experiments with TAM isolates where an organic carbon source was required, HSM growth medium was supplemented with 0.2% (w/v) glucose.

Liquid cultures were grown under continuous illumination at 25^0C or 35^0C depending on the application with rotary agitation at 100 rpm and an average light intensity of 25-50 µmol m^{-2} s^{-1} [Innova 4430, New Brunswick Scientific], whereas 1.5% agar plates were kept under continuous illumination at 18-23^0C and average light intensity of 50 µmol m^{-2} s^{-1}.

Table 2-1 *Euglena gracilis* (EG) and Jaworski's medium (JM) growth media recipes.
To prepare EG medium add 1ml of each stock solution to 991ml ddH_2O; Adapted from CCAP Culture collection main website (www.ccap.ac.uk).

***Euglena gracilis* medium (EG)**	**Stock solution♯**	**per 200 ml**
$Ca(NO_3)_2.4HO$	1	4.0 g
KH_2PO_4	2	2.48 g
$MgSO_4.7H_2O$	3	10.0 g
NaHCO3	4	3.18 g
EDTAFeNa	5	0.45 g
$EDTANa_2$	5	0.45 g
H_3BO_3	6	0.496 g
$MnCl_2.4H_2O$	6	0.278 g
$(NH_4)6Mo_7O_{24}.4H_2O$	6	0.20 g
Cyanocobalamin	7	0.008 g
Thiamine HCl	7	0.008 g
Biotin	7	0.008 g
$NaNO_3$	8	16.0 g
$Na_2HPO_4.12H_2O$	9	7.2 g

Jaworski's Medium (JM)	**Per 1 litre**
Sodium acetate trihydrate	1 g
"Lab-Lemco" powder (Oxoid L29) *	1 g
Tryptone (Oxoid L42) *	2 g
Yeast extract (Oxoid L21) *	2 g
$CaCl_2$	90 μM

* Unipath Ltd, Basingstoke, UK

Table 2-2 Tris-acetate phosphate (TAP and high salt minimal (HSM) growth media recipes.
Adapted from (Harris, 1989)
*4 x Beijerinck salts: 0.3 M NH_4Cl, 14 mM $CaCl_2 \cdot 2H_2O$ and 16 mM $MgSO_4 \cdot 7H_2O$
*1 M (K)PO_4, pH 7: 1 M K_2HPO_4 titrated to pH 7.0 with 1M KH_2PO_4
*2 x PO_4 for HSM: 80 mM K_2HPO_4 and 50 mM KH_2 PO_4, adjusted to pH 6.9 with KOH
*Trace elements: 180 Mm H_3BO_3, 77 mM $ZnSO_4 \cdot 7H_2O$, 26 mM $MnCl_2 \cdot 4H_2O$, 18 mM $FeSO_4 \cdot 7H_2O$, 7 mM $CoCl_2 \cdot 6$ H_2O, 6 mM $CuSO_4 \cdot 5H_2O$, 0.9 mM $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$.

For 1 litre	Tris-acetate phosphate (TAP)	High salt minimal (HSM)
H_2O	975 ml	925 ml
Tris	2.42 g	-
*4 x Beijerinck salts	25 ml	25 ml
*1 M (K)PO_4, pH 7	1 ml	-
*2 x PO_4 for HSM	-	50 ml
*Trace elements	1 ml	1 ml
Glacial acetic acid	to pH 7.0 (approximately 1 ml)	-

2.1.2 *Chlamydomonas reinhardtii*

Chlamydomonas reinhardtii wild type CC-1021 strain was obtained from the Chlamydomonas Resource Center (www.chlamy.org) and was derived from the original 137C isolate (Harris 1989). All strains of *C. reinhardtii* used in this work, including the recipient strains for chloroplast transformation are listed in the Table 2-3. The *C. reinhardtii* strains were cultured in either TAP or HSM growth media depending on the application. Liquid cultures were grown under continuous illumination at 25^0C, 100 rpm rotary agitation and the average light intensity 25-50 µmol m^{-2} s^{-1}, whereas 1.5% agar plates were kept at 18-23^0C and the average light intensity of 50 µmol m^{-2} s^{-1}.

Table 2-3 *Chlamydomonas reinhardtii* strains used in this work

C. reinhardtii strain	Application	Comments
WT_{12}		Wild Type
BST-SAME (mt+)	Chloroplast transformation recipient	BST-SAME is a PSII-deficient strain created by disruption of the chloroplast-encoded *psbH* gene with the *aadA* cassette. The *aadA* cassette was inserted into the *Bst*XI site in the same orientation as *psbH* gene. The created strain is not only photosynthetically impaired but also spectinomycin (Spc) and streptomycin (Str) resistant. Selection of transformants occurs via restoration of photosynthesis on minimal medium.
CW15.J3 (mt-)	Cell wall-deficient cell line	CW15.J3 is a wild type cell wall-less strain (Davies and Plaskitt, 1971). It was used as recipient cell line in Chapter 4 to test functionality of the *C. sorokiniana psbA* promoter and 5'UTR region. Additionally, this strain was a starting point for creation of TN72 recipient cell line (Ninlayarn, 2012).
TN72 (mt+)	Chloroplast transformation cell wall-deficient recipient	The TN72 strain, similarly to BST-SAME contains the *psbH* gene disrupted with the *aadA* cassette. Therefore it is also a PSII-deficient Spc + Str-resistant strain. The detailed steps of the TN72 strain creation are presented in (Ninlayarn, 2012).
KRC1001-11A	Chloroplast transformation recipient	KRC1001-11A PSI-deficient strain is bearing mutations in *psaA* exon 3 (*psaA-3*) (Santabarbara et al., 2010)

2.1.3 *Escherichia coli*

E. coli DH5α strain used in cloning was obtained from CLONTECH Laboratories, Inc. [genotype: F-, (φ80*lacZ*Δ*M15*), Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR17*(r_k-, m_k+), *supE44*, *thi-1*, *gyrA96*, *relA*, λ] whereas in the Site Directed Mutagenesis (SDM) application *E. coli* XL1-Blue strain was used [genotype: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* ($F' proAB lacI^qZ\Delta M15$ Tn*10* (Tet^r))]. Both strains were grown on Luria-Bertani (LB) liquid growth medium containing 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract and 0.17 M NaCl at 37°C in a shaking incubator [Stuart Scientific Incubator S.I.60, UK with built-in shaker IKA-VIBRAX-VXR, IKA® Labortechnik, Germany] overnight at 170 rpm and also on LB plates where the LB medium was supplemented with 2 % (w/v) nutrient agar at 37^0C in a static incubator [LAB. Companion, Fisher Scientific].

2.2 Growth rate measurement

All starter cultures of *Chlorella* strains were grown in 15 ml of TAP medium in 50 ml conical flasks. All experiments, unless stated, were performed in 100 ml of appropriate medium in 250 ml conical flasks containing HSM (High Salt Medium) or JM (Jaworski Medium) minimal media or TAP (Tris Acetate-Phosphate) medium containing sodium acetate as carbon source under constant agitation and constant illumination with the average light intensity of 50 µmol m^{-2} s^{-1}.

Growth rates of each strain were measured in three ways: 1) by recording optical density 2) by measuring cell count or 3) by recording biomass production. The specific growth rate was calculated using the equation $\mu = \mathrm{Ln}(N_2/N_1)/(t_2 - t_1)$ with N_1 and N_2 being either dry weight or optical density (750 nm) at the beginning (t_1) and at the end (t_2) of the exponential growth phase.

2.2.1 Optical density (OD) measurement

Optical density of each culture was recorded using a Unicam UV/Vis Spectrometer [Thermo Electron Corporation, USA] at 750 nm. Samples from the mid-log phase onwards were diluted five fold with the adequate medium. All cultures were started with an OD of approximately 0.04 and cultured until they reached stationary phase.

2.2.2 Cell count measurement

Algal cells were counted using a haemocytometer. For *Chlamydomonas reinhardtii*, 10 µl of iodine tincture (19.7 mM iodine in 95% (v/v) ethanol) were added to 1 ml of cells from a liquid culture in stationary phase to inhibit motility and the cell suspension was then diluted 10 fold. *Chlorella* sp. cell cultures were diluted 5 or 20 fold during the early and mid-log phase respectively and up to 100 fold during the stationary phase before applying to the haemocytometer.

2.2.3 Biomass production measurement

Biomass production was measured by recording the dry weight where the pelleted cells were washed with 0.1 M phosphate buffer (pH 6.9) and freeze-dried in pre-weighed

plastic conical tubes [BD Falcon^{TM}, USA] using an Edwards Freeze Dryer Modulyo Pirani 10 at -20^{0}C for 24 hours or until the weight no longer decreased.

2.2.4 Growth analysis using 'Spot tests'

Approximately 7 µl of a tested culture in mid-log phase was applied on either HSM or TAP agar plates depending on the application, and allowed to dry. The plates were grown under continuous illumination and average light intensity of 50 µmol m^{-2} s^{-1} for 5-10 days depending on the tested strain.

To measure the influence of the reduced oxygen level on algal growth, the spot test analysis was performed using the anaerobic system BBL GasPak Pouch System [Becton Dickinson, Marylands, USA] according to the manufacturer's instructions.

2.3 Analysis of photosynthesis-deficient mutants of *C. sorokiniana*

2.3.1 Isolation of photosynthesis-deficient mutants

Chlorella sorokiniana UTEX1230 strain was grown for 7 days in minimal medium (HSM) until it reached the stationary phase. A flask containing TAP medium and 1 mM 5-fluorodeoxyuridine (FdUrd) was inoculated with 5 x 10^7 cells of the previous culture (Spreitzer and Mets, 1981) and a flask containing TAP only was used as a control. The cultures were grown in the dark under constant agitation for 10 days until they reached stationary phase. The FdUrd-treated cells were washed in 0.02 M potassium phosphate buffer (pH 6.9), re-suspended in 0.1 M phosphate buffer (pH 6.9) to 2 x 10^8 cells/ml. Afterwards the re-suspended cultures were subjected to ethyl methanesulfonate (EMS) treatment with either 0.135 M or 0.270 M solutions of EMS (Loppes, 1968) and incubated in the dark for 2 hours (Spreitzer and Mets, 1981). Thereafter the cells were washed twice in 0.02 M potassium phosphate buffer (pH 6.9). The survival rate after EMS treatment measured on TAP plates (Spreitzer and Mets, 1981) was approximately 10% and the cells were plated on TAP + 10 mM metronidazole and maintained in dim light at the average light intensity of 5 µmol m^{-2} s^{-1} (Galloway and Mets, 1989). Single colonies appeared after approximately 8 weeks. Approximately 3000 colonies were isolated and tested for photosynthetic activity and sensitivity to light in a spot test on

HSM and TAP plates. The spot tests were subsequently subjected to the chlorophyll fluorescence analysis.

2.3.2 77 K (low temperature) fluorescence measurement

Liquid cultures were cultured in HSM until they reached early log phase and an OD of approximately 0.2 at 750 nm. 200 µl samples were injected into a sample holder containing two flat round pieces of glass with a diameter of 14 mm taped up by a metal ring and subsequently frozen in liquid nitrogen. The samples were excited at 435 nm and low temperature (77 K) and the fluorescence emission spectra were detected between 600 and 780 nm using a Jobin-Yvon Flouromax-3 instrument as described previously (Rees et al., 1992, Ruban et al., 1992, Johnson and Ruban, 2009).

2.3.3 UV mutagenesis

10 ml of mid-log cultures (approximately 10^7 cells/ml) of *Chlorella sorokiniana* UTEX1230 were subjected to UV irradiation (6W UV bulb at 254 nm, culture was positioned approximately 8 cm from the bulb) for 19 minutes. Under these conditions a survival rate of approximately 10% was achieved. The cells were left to recover in the dark for 2 hours to prevent photo-reactivation followed by plating 10 µl of 100 fold diluted UV-treated and non-treated cultures on 15 TAP plates per each condition. Single colonies appeared after 7-10 days; approximately 500 UV-treated colonies were isolated with sterile toothpicks, re-streaked on separate TAP plates and subsequently subjected to the chlorophyll fluorescence analysis and tested for chlorophyll content.

2.3.4 Relative chlorophyll fluorescence analysis

Mutant and wild type strains were spotted onto solid media one week before the analysis. *In vivo* chlorophyll fluorescence was measured and recorded using a FluorCam 700MF [Photon Systems Instrument] with the settings presented in Table 2.4. The isolates were first dark-adapted for 10 minutes after which time a F_0 measurement was taken. Directly afterwards the isolates were exposed to a 800 ms flash of saturating white light that allowed the determination of F_m measurements and the created fluorescence images were captured by the digital video camera of the fluorimeter (Polle et al., 2003). The mximum

PSII efficiency can be calculated from the equation $(F_m–F_o)/F_m$. The described method was suitable for testing the photosynthetic activity of both types of created mutants: the chlorophyll antenna mutants and photosynthesis-deficient mutants.

Table 2-4 FluorCam 700MF settings for photosynthetic efficiency measurements.
F_o: minimal fluorescence, F_m: maximal fluorescence

Global Parameters	
Sensitivity (%)	80
Electronic Shutter	1
Visual Frame (s)	2.36
Fo Measurement	
Fo Measurement Duration (ms)	2
Every n^{th} Frame	10
Fm Measurement	
Pulse Duration (ms)	800
Pulse Intensity (%)	100

2.4 Analysis of the chlorophyll antenna of *C. sorokiniana* UTEX1230 strains

2.4.1 Preparation of thylakoids

Cultures at late-log phase (approximately 6 x 10^9 cells) were pelleted in a centrifuge [Megafuge 1.0R Heraeus, Kendro Laboratory Products, Germany] for 4 min at 2000 x *g*. The pellet was re-suspended in 45 ml of GB (grinding buffer) containing 0.35 M sorbitol, 50 mM Tricine pH 7.9, 10 mM NaCl, 5 mM MgCl, 0.5% (w/v) skimmed milk powder and 1x protease inhibitor cocktail (Benzimidine 1mM, PMSF 1mM, Aminocaproic acid 2 mM). The cell suspension was homogenised twice using a cell disruptor [Constant Cell Disruption Systems, Constant Systems Ltd, UK] at 1.48 KPSI at 4^0C. Unruptured cells and cell debris were pelleted at 4^0C by centrifugation for 5 min at 1500 x *g* [Centrifuge 5415R, Eppendorf]. Isolated thylakoids remaining in the supernatant were transferred to new tubes and centrifuged at 4^0C for 20 min at 30 000 x *g* [Sorvall RC28S, DuPONT]. The pellet was washed with RB (resuspension buffer) containing 10 mM NaCl, 5 mM MgCl, 0.5% (w/v) skimmed milk powder, protease inhibitor cocktail) and centrifuged at 30 000 x *g* for 30 min at 4^0C. Subsequently, the pellet containing purified thylakoids was resuspended in 800 µl of B4 buffer containing 0.4 M sorbitol, 10 mM Hepes KOH (pH 7.5), 15 mM NaCl and 5 mM $MgCl_2$. The isolated thylakoid fractions were frozen in liquid nitrogen and stored at -80^0C.

2.4.2 Pigment analysis

Pigments were extracted from purified thylakoids with 80% acetone (v/v), and subsequently separated and quantified by HPLC using a reverse-phase C18 column (Synergi 4u hydro-RP 80A 250 x 4.60 mm, Phenomenex) according to the modified protocol described previously (Gilmore and Yamamoto, 1991). Acetone extracts were separated by chromatography at 1.5 ml min^{-1} starting with an aqueous mixture of acetonitrile/methanol/ 0.1 M TrisHCl buffer (pH 7.9) (72:8:3) for 1 min followed by a 15 min linear gradient of methanol/hexane (4:1).

For the determination of the chlorophyll a/b ratio of isolated thylakoids, 10 µl of purified extracted thylakoids was re-suspended in 200 µl of 90% (v/v) acetone buffered with Na_2CO_3 and 10 µl of ddH_2O and incubated on ice for 5 minutes. The solution was

centrifuged for 10 min at maximum speed in order to remove the cell debris and the supernatant was used for further analysis. For the chlorophyll a/b ratio measurement of the whole cell extracts, 2 ml of late-log culture (an OD at 750 nm of approximately 0.8) were pelleted in a microfuge for 5 min at maximum speed. 100 µl of NN-dimethyl formamide was added to the pellet and incubated on ice for 30 min in order to extract the pigments without rupturing the cells. The solution was centrifuged for 15 min at maximum speed. Afterwards 12 µl of the supernatant was added to 500 µl of 90% (v/v) acetone buffered with Na_2CO_3 and 50 µl of ddH_2O.

The spectra of the acetone extracts of thylakoids and whole cell extracts were recorded in the visible range from 350 to 750 nm with an SLM-Aminco DW 2000 spectrophotometer [SLM Instruments] and compared to the library of spectra of pure pigments in 80% (v/v) acetone.

2.4.3 Pigment extraction for total chlorophyll measurement

1 ml of a mid-log culture (OD at 750 nm of approximately 0.7) was pelleted in a microcentrifuge [Biofuge pico, Heraeus Instruments, Germany] at maximum speed for 1 min. 50 µl of dimethyl formamide [Sigma] was added to the pellet and sonicated in a water sonicator [Kerry Ultrasonics, UK] containing ice-cold water for 20 min. The extracts were incubated on ice overnight. Afterwards, 950 µl of 85% (v/v) acetone was added, the solution was thoroughly mixed and the debris was pelleted in the microcentrifuge for 1 min at maximum speed. OD was measured using a UNICAM UV/VIS Spectrometer at 663.6 nm and 646.6 nm and the total chlorophyll content was calculated according to the equations previously provided (Porra et al., 1989).

2.4.4 Gradient Native Gel Preparation

Large gradient polyacrylamide gels were freshly cast. 4% and 12% acrylamide solutions (Table 2-5) were prepared and polymerisation was activated by adding 10% (w/v) AMPS and TEMED solutions.

Table 2-5 Recipes for 4% and 12% acrylamide solutions for the gradient separation gel preparation and 3.5% stacking gel preparation.

	4%	**12%**	**Stacking gel 3.5%**
Acrylamide/bis (48/1.5)	1.3 ml	3.9 ml	0.7 ml
Tris-Gly buffer [120 mM Tris, 480 mM Glycine]	1.6 ml	1.6 ml	1 ml
Glycerol (50% v/v)	3 ml	3 ml	2 ml
dd H_2O	Up to 16 ml	Up to 16 ml	Up to 10 ml
Ammonium persulphate (AMPS 10% w/v)	90 µl	90 µl	40 µl
Tetramethylethylenediamine (TEMED)	35 µl	35 µl	10 µl

The activated solutions were placed in the two chambers of a gradient maker with a small magnet in the chamber where mixing of both acrylamide solutions took place. The gradient maker and the mechanism of mixing of 4% and 12% acrylamide solutions are presented in Figure 2-1. The stacking gel contains 3.5% acrylamide stock in Tris-Gly buffer at a pH of 6.8. Approximately two hours later the separation gel set and the stacking gel solution was poured on top. The gel was further polymerised for 24 hours at 4^0C in the dark. Gels were run for 18 h at 130 V at 4°C in a basic buffer [containing 120 mM Tris and 960 mM glycine]: internal buffer (+) contained a ten times diluted basic buffer, whereas the external buffer (-) contained ten times diluted basic buffer plus 0.1% of Deriphat-160.

Figure 2-1 Schematic diagram of a gradient maker applied to create linearly increasing gradient gels for native gel analysis.

Adapted from Springer Images

12% acrylamide solution is gradually diluted with 4% acrylamide solution before the mixed solution is poured into a chamber made of two glass plates separated by a 1 mm spacer.

2.4.5 Preparation of samples for native gel

All the preparation work was performed on ice and in dim light.

The chlorophyll concentrations from thylakoids were calculated according to the method described earlier (Porra et al., 1989). 35 µg of chlorophyll per sample was prepared in duplicate and centrifuged in the microfuge at 4^0C for 15 min at maximum speed. The pellet was resuspended in 45 µl of RS (resuspension solution) containing 50 mM BisTris-HCl (pH 7.0), and 40% (w/v) glycerol. Subsequently, 45 µl of the solubilisation solution (RS supplemented with 10% (v/v) of detergent Triton X-100) was added and samples were vortexed for 1 min at the maximum setting. The samples were incubated for 10 min on ice in the dark. Afterwards, un-solubilised debris was pelleted in the microfuge at 4^0C for 15 min at maximum speed. 2x sample buffer containing 20% (v/v) glycerol, and 0.02% (w/v) bromophenol blue was added to the collected supernatant and loaded onto a previously prepared and completely polymerised native gel.

2.4.6 Protein gel electrophoresis and electrophoretic transfer

Protein gel electrophoresis was performed using sodium dodecyl sulphate polyacrylamide gels and the Mini-PROTEAN® System [BioRad]. The gels were prepared according to the Laemmli's recipe (Laemmli, 1970) as presented in Table 2-6. The 15% resolving gel was left to polymerise for 1 hour at room temperature and afterwards a stacking gel was applied. Polymerised gels were assembled into the Mini-PROTEAN® System [BioRad], and this was filled with Tris-glycine-SDS electrophoresis buffer containing 0.025 M Tris, 0.192 M glycine and 0.1% (w/v) SDS, pH 8.3.

For immunoblotting the thylakoid samples corresponding to 0.05 µg, 0.1 µg and 0.25 µg of chlorophyll were mixed with 2x sample buffer containing 20% (v/v) glycerol, 2% SDS and 0.02% (w/v) bromophenol blue and each sample was loaded to a marked well on the gel. Prestained Molecular Weight Marker [Sigma] was used to estimate the size of the proteins. Samples were run for 3 hours at room temperature at 120 V using a PowerPac 300 [BioRad].

Table 2-6 Laemmli gel recipe for SDS-PAGE (Laemmli, 1970).

Reagents	Volume
15% resolving gel	**6 ml**
Acrylamide/bisacrylamide 40% stock, 37:1 [Sigma, Dorset, UK]	2.25 ml
Resolving gel buffer (8 x stock) 3 M Tris-HCl, pH 8.8	0.75 ml
10% SDS	60 µl
ddH_2O	2.7 ml
10% ammonium persulfate (AMPS) [Sigma, Dorset, UK]	250 µl
N,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED) [Sigma, Dorset, UK]	2.5 µl
3.75% stacking gel	**3.5 ml**
Acrylamide/bisacrylamide 40% stock, 37:1 [Sigma, Dorset, UK]	0.33 ml
Stacking gel buffer (4 x stock) 0.5 M Tris-HCl, pH 6.8	0.875 ml
10% SDS	35 µl
ddH_2O	2.1 ml
10% ammonium persulfate (AMPS) [Sigma, Dorset, UK]	175 µl
N,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED) [Sigma, Dorset, UK]	2.5 µl

Following separation on the polyacrylamide gel, proteins were transferred to Hybond™ ECL nitrocellulose membrane [Amersham GE Healthcare]. 12 pieces of 3MM Whatman paper were cut to the same size as the gel and the nitrocellulose membrane and all items were immersed for 30 min in Towbin buffer containing 25 mM Tris, 192 mM glycine and 20% (v/v) methanol. A "sandwich" was assembled on the semi-dry transfer cell in the following order: 6 pieces of 3MM Whatman paper, Hybond™ ECL nitrocellulose membrane, gel, and 6 pieces of 3MM Whatman paper. Protein transfer from the polyacrylamide gel to the membrane was performed using the semi-dry transfer cell TRANS-BLOT®SD [BioRad] at a constant voltage of 18 V [Fisons FEC 570 Powerpac].

2.4.7 Immunodetection

After completing SDS-PAGE and electrophoretic transfer, the membrane was incubated overnight at 4^0C with blocking solution containing PBS (0.137 M NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8.1 mM Na_2HPO_4, pH 7.2), 0.2 % (v/v) Tween 20 and 5 % (w/v) skimmed powdered milk. Subsequently, the membrane was incubated for 2 hours at room temperature with the primary antibody diluted in blocking solution, washed three times in blocking solution and incubated for 1 hour at room temperature with the secondary antibody diluted in blocking solution. Table 2-7 lists the primary and secondary antibodies and their dilutions used in this application.

Table 2-7 Antibody and dilution factors used in immunoblotting for the chlorophyll antenna measurement.

Antibody	Source	Type	Dilution factor
Primary antibody			
α-PsaA (PSI-A) (~70kDa)	Agrisera [AS06172]	Rabbit, polyclonal serum	1:2000
α- CP47 (~47kDa)	Bassi Group, Verona	Rabbit, recombinant	1:1000
α- LHCII (~24kDa)	Bassi Group, Verona	Rabbit, recombinant	1:1000
Secondary antibody			
Anti-rabbit IgG	Sigma [A3687]	Goat, alkaline phosphatase	1:15000

After incubation with the secondary antibody the membrane was washed twice with blocking solution and once with PBS. Freshly prepared developing buffer was made by adding 66 µl of NBT solution (50 mg/ml of nitrotetrazolium blue chloride [Sigma] dissolved in 70% dimethylformamide) and 33 µl of BCIP solution (50 mg/ml of 5-bromo-4-chloro-3-inodyl phosphate disodium salt [Sigma]) to 10 ml of developing solution (100 mM Tris-HCl (pH 9.5), 100 mM NaCl, 5 mM $MgCl_2$). The membrane was incubated for approximately 10-20 min until expected bands reached required intensity. Afterwards the developed membrane was washed 4 times with ddH_2O. The signal amplitude was quantified (n = 3) using GelPro 3.2 software [Bio-Rad].

2.4.8 *In vivo* fluorescence and NPQ measurement

Fluorescence induction kinetics to determine PSII functional antenna size in dark-adapted algal cultures was measured and recorded using a home-built time-resolved camera fluorimeter. Variable fluorescence was induced with a green light of 15 µmol photons m^{-2} s^{-1}, on thylakoids (10 µg chlorophyll/ml) in a measuring buffer containing 10 mM HEPES, pH 7.8, 5 mM $MgCl_2$, 30 µM DCMU, and 50 µM nigericin. The reciprocal of time corresponding to two-thirds of the fluorescence rise ($T_{2/3}$) was considered as a measure of the PSII functional antenna size (Malkin et al., 1981). NPQ (non-photochemical quenching) was measured through chlorophyll fluorescence of late-log algal cultures (OD at 750 nm of approximately 0.8) that were either adapted to high light (500 µmol photons/m^2/s) or low light (50 µmol m^{-2} s^{-1}) conditions at room temperature using a PAM 101 fluorimeter (Heinz-Walz) (Andersson et al., 2001).

2.5 DNA manipulation

2.5.1 MICS [Molecular Identification of *Chlorella sorokiniana* strains] analysis on non-denaturing polyacrylamide gels

8% polyacrylamide gels were prepared according to the recipe: 7.5 ml of sterile glycerol, 6 ml of 40% acrylamide:bis (19:1), 3 ml of 10 x TBE (1 M Tris, 0.9 M Boric Acid, 0.01 M EDTA), 0.3 ml of 10% (w/v) ammonium persulphate (AMPS), 18 µl of Tetramethylethylenediamine (TEMED) solution up to 30 ml total volume. The gels were left for polymerisation for 2 hours and afterwards all sample wells were washed with

running buffer – 1 x TBE. All prepared DNA samples were mixed with 6 x DNA Loading Dye & SDS Solution [Fermentas] and loaded into the sample wells. The gels were run for 90 min at 60 V [Gibco power supply], stained for 20 min in an ethidium bromide bath (1 μg/ml) and subsequently visualised in a UV illuminator [UVP Gel Documentation System].

2.5.2 Total genomic DNA isolation from *C. sorokiniana* for Southern blot analysis

5 ml of *Chlorella sorokiniana* cultures from early stationary phase (OD at 750 nm of approximately 1) was harvested by centrifugation at 9,000 x *g* for 10 min. The formed pellet was then resuspended in 0.5 ml CTAB buffer (2% hexadecyltrimethylammonium bromide [Sigma], 0.1 M Tris-HCl pH 8 [Sigma], 1.4 M NaCl [Sigma], 10 mM EDTA [Sigma], 2% β-mercaptoethanol [Sigma], RNaseA 100 μg/ml [Sigma]) and incubated for 1 hour at 65^0C. The mixture was extracted with 0.5 ml of phenol:chloroform:isoamyl alcohol 25:24:1 [Sigma] by centrifugation at 9,000 x *g* for 5 min and the aqueous phase was then extracted with two volumes of absolute ethanol and 0.3 M sodium acetate. The DNA was pelleted by centrifugation for 30 min at 9,000 x *g* and the formed pellet was washed in 70% ethanol and dissolved in ddH_2O.

2.5.3 Algal genomic DNA isolation for PCR analysis

Total genomic DNA of different algal species was extracted by mixing a cell suspension, absolute ethanol and 5% Chelex resin [BioRad] in the proportion 1:1:10 respectively and boiled for 5 min. The debris and resin from the mixture was then pelleted in a microcentrifuge [Biofuge pico, Heraeus Instruments] at maximum speed for 2 min and the supernatant was used for further analysis.

2.5.4 Construction of chloroplast expression vectors

2.5.4.1 Expression vectors

Vectors pUC9 and pUC18 were obtained from the Purton collection. Plasmid pUC-atpX-AAD was kindly provided by Professor Michel Goldschmidt-Clermont from Université

de Genève. Vectors pMK-RQ-*FatB1* and pMK-*L1* conferring kanamycin resistance gene and also pMA-CVN conferring ampicillin resistance gene were obtained from GeneArt [Regensburg, Germany]. All three plasmids contained synthetic genes *FatB1*, *L1* or CVN codon optimised for the chloroplast genome of *Chlamydomonas reinhardtii*. Further details of DNA sequences and cloning strategy are presented in the appendices and the result chapters 5 and 6, respectively.

The pASapI expression vector was designed and created by Dr Saul Purton. It is a pUC8-based vector designed for the targeted incorporation of transgenes into a neutral locus between *trnE2* and *psbH* within the *C. reinhardtii* chloroplast. Isolation of transformants is established through restoration of photosynthetic activity of the recipient strain that has a partially deleted *psbH* gene. A "gene expression cassette" is constructed with the promoter/5'UTR sequence isolated from the *atpA* gene, the coding sequence of a foreign gene, and the 3'UTR of *rbcL*. The pASapI vector was subsequently altered by replacement of the promoter/5'UTR sequence of *atpA* with the promoter/5'UTR of the following genes: *chlL* (pCSap2), *psbA* (pPSapI), or *psaA* (pSSapI). Further details of the sequence and cloning strategy are presented in the subchapter 5.3.2.1 and in Ninlayarn (2012).

2.5.4.2 Restriction enzyme digestion

The restriction enzymes used in recombinant DNA work were supplied by New England BioLabs and were applied according to the manufacturers' instructions, using the appropriate buffer supplied by NEB.

2.5.4.3 Ligation

For ligation reactions, linearised vector molecules, insert DNA molecules, T4 DNA ligase buffer [New England Biolabs] and T4 DNA ligase [New England Biolabs] were added according to the manufacturer's instructions to give the total volume of 10 µl. The ratio of insert:vector DNA was approximately 6:1. Ligation reactions were incubated for 2 hours at room temperature and the ligation reaction was used directly for transformation of the competent DH5α strain of *E. coli*.

2.5.5 Preparation of competent *E. coli*

The DH5α strain of *E. coli* was re-streaked from a frozen glycerol stock onto an LB plate and incubated overnight at 37^0C. A single colony was chosen to inoculate 10 ml of fresh LB medium that was cultured overnight at 37^0C. 1 ml of the overnight culture was subsequently used to inoculate 100 ml of fresh LB medium that was cultured for 2.5 hours at 37^0C. Afterwards, the 100 ml culture was kept on ice for 15 min and pelleted in pre-cooled 30 ml tubes [Sterilin Ltd, Fisher Scientific] at 9000 x *g* for 5 min. 40 ml of ice-cold 50 mM $CaCl_2$ was added to the cell pellet and the re-suspended cells were kept on ice for 30 min. The cell suspension was pelleted at 9000 x *g* for 5 min and the cell pellet was re-suspended in fresh ice-cold 50 mM $CaCl_2$ up to 8 ml. The solution was ultimately mixed with 3.5 ml of 50% (v/v) glycerol, frozen as aliquots in liquid nitrogen and stored at -80^0C.

2.5.6 Transformation of *E. coli*

100 µl of competent DH5α were thawed and added to a 15 ml conical tube [BD Biosciences] that was pre-cooled on ice for 5 min prior to application. Approximately 10 ng of plasmid DNA or 10 µl of the ligation reaction was added to the thawed competent DH5α cells and the mixture was incubated on ice for 30 min. Cell suspension incubated with DNA was subjected to a heat shock at 42^0C for 1 min and 1 ml of fresh LB medium was added. The cell suspension was incubated at 37^0C for 1 hour in a shaking incubator [Stuart Scientific Incubator S.I.60, UK with a built-in shaker IKA-VIBRAX-VXR, IKA® Labortechnik, Germany] at 170 rpm. Afterwards, 30 µl of the culture was plated on selective LB plates and incubated overnight at 37^0C. The selective LB plates contained either 100 µg/ml of ampicillin [Sigma], 50 µg/ml of kanamycin [Sigma] or 100 µg/ml spectinomycin [Sigma] depending on the antibiotic resistance gene present in the plasmid. Single colonies were picked and used to inoculate 5 ml of LB containing the suitable antibiotic and were incubated overnight at 37^0C with shaking at 170 rpm.

2.5.7 Plasmid DNA isolation

Plasmid DNA was isolated from overnight *E. coli* cultures using a QIAprep Spin Miniprep Kit [Qiagen] or QIAprep Spin Midiprep Kit [Qiagen] according to the manufacturer's instructions.

2.5.8 Polymerase chain reaction (PCR) and DNA sequencing

PCR amplifications were performed according to the manufacturer's instructions and protocols using the Phusion® High-Fidelity DNA Polymerase [New England Biolabs]. DNA sequencing was performed by the Scientific Support Services of The Wolfson Institute for Biomedical Research at UCL.

The details of the primers used in PCR amplification and sequencing can be found in the appendices and result chapters 3, 4, 5 and 6.

2.5.9 Site Directed Mutagenesis

Site Directed Mutagenesis (SDM) was performed according to the protocol attached to the QuikChange® Site-Directed Mutagenesis Kit [Stratagene] with some modifications. In order to create an *Eco*RV restriction site between *psbB* and *clpP* genes within the amplified 3.14 kb fragment of chloroplast DNA of *C. sorokiniana* UTEX1230 (see Figure 4-3) a set of primers was designed: RV.SDM.F (CGTGTTCTTAATTTTAGGATATCACATTTTTTTTTACC) and RV.SDM.R (GGTAAAAAAAAATGTGATATCCTAAAATTAAGAACACG). Table 2-8 presents the applied temperature parameters for the PCR amplification using RV.SDM.F and RV.SDM.R primers. Subsequently, each PCR reaction was subjected to digestion with *Dpn*I restriction enzyme in order to destroy the parental DNA template.

Table 2-8 Temperature parameters applied for the PCR amplification during the site directed mutagenesis application.

pUC.Cs.C3.1 expression vector reaction:	Control reaction:
95^0C 30s 95^0C 30s | 55^0C 1 min | x 12 68^0C 3 min 10s |	95^0C 30s 95^0C 30s | 55^0C 1 min | x 12 68^0C 5 min |

2.5.10 Agarose gel DNA electrophoresis

DNA fragments were separated on 1% (w/v) agarose gels made with 1 x TAE buffer (40 mM Tris, 1mM sodium EDTA, 17.5 mM glacial acetic acid). 6x DNA Loading Dye & SDS Solution [Fermentas] was added to DNA samples. O'GeneRuler™ 1 kb Plus DNA Ladder, ready-to-use, 75-20,000 bp [Fermentas] was used to estimate the size of the DNA fragments. Gels were submerged in 1 x TAE buffer in electrophoresis tanks [Sigma]. The gels were run for 90 min at 80 V [Gibco power supply], stained for 20 min in an ethidium bromide bath (1 μg/ml) and subsequently visualised in a UV illuminator [UVP Gel Documentation System].

2.5.11 Chloroplast transformation

2.5.11.1 Glass bead-mediated transformation method

Glass bead-mediated transformation was performed according to the previously described method (Kindle et al., 1991) with minor modifications. The recipient strain was cultured in TAP medium until it reached mid-log phase and was pelleted for 4 min at 2000 x *g*. The *C. reinhardtii* cells were harvested when the cell density reached 2 x 10^6 cells/ml and *C. sorokiniana* at a cell density of 1.2 x 10^7 cells/ml. The cell pellet was re-suspended in HSM medium. For *C. reinhardtii* strains, the final cell density was 9 x 10^7 cells/ml whereas for *C. sorokiniana* this value was 4.5 x 10^8 cells/ml. 0.3 ml of each cell suspension was vortexed at the maximum speed [Vortex Genie II, Fisher] for 15 seconds with 300 mg of glass beads and 3 μg of the suitable chloroplast expression plasmid in a glass 5 ml conical tube. Each sample was mixed with HSM medium containing 0.5% (w/v) agar that had been melted and cooled to 40^0C. The cell suspension was spread on the appropriate selective medium. Plates were stored overnight under dim light (5 μmol m^{-2} s^{-1}) at 25^0C and on the next day transferred to continuous illumination conditions at 25^0C and an average light intensity of 50 μmol m^{-2} s^{-1}.

2.5.11.2 Biolistic DNA delivery method

Microcarrier particles (1 µm particles of either tungsten or gold) were coated with plasmid DNA according to the protocols designed for *Chlamydomonas* (Choquet et al., 1992, Remacle et al., 2006, Svarovsky et al., 2008, Sanford et al., 1993). 6 mg of the microcarrier particles were washed three times in 70% (v/v) ethanol and re-suspended in 25% (v/v) glycerol mixed with 35% (v/v) ethanol [Sigma]. 3 mg of the tungsten (*C. reinhardtii*) or gold (*C. sorokiniana*) particles were re-suspended in 25% (v/v) glycerol and vortexed with 1 µg of the plasmid DNA in the presence of 50 µl of 2.5 M $CaCl_2$ [Sigma] and 20 µl of 0.1 M spermidine trihydrochloride [Sigma]. The coated microparticles were washed with 70% ethanol and re-suspended in 48 µl of absolute ethanol. The mixture was then split between six samples and applied onto the centre of the sterilised macrocarriers. The coated microparticles were fired into the algal lawn using the PDS-1000/He Biolistic Particle Delivery System [Bio-Rad, USA] as it is presented in Figure 2-2 at a pressure of 1100 psi using 1100 rupture discs [BioRad].

The applied settings, including the apparatus setting, vacuum generation, shooting, and release steps were carried out accordingly to the manufacturer's instruction and the guidelines given in Sanford et al. (1993).

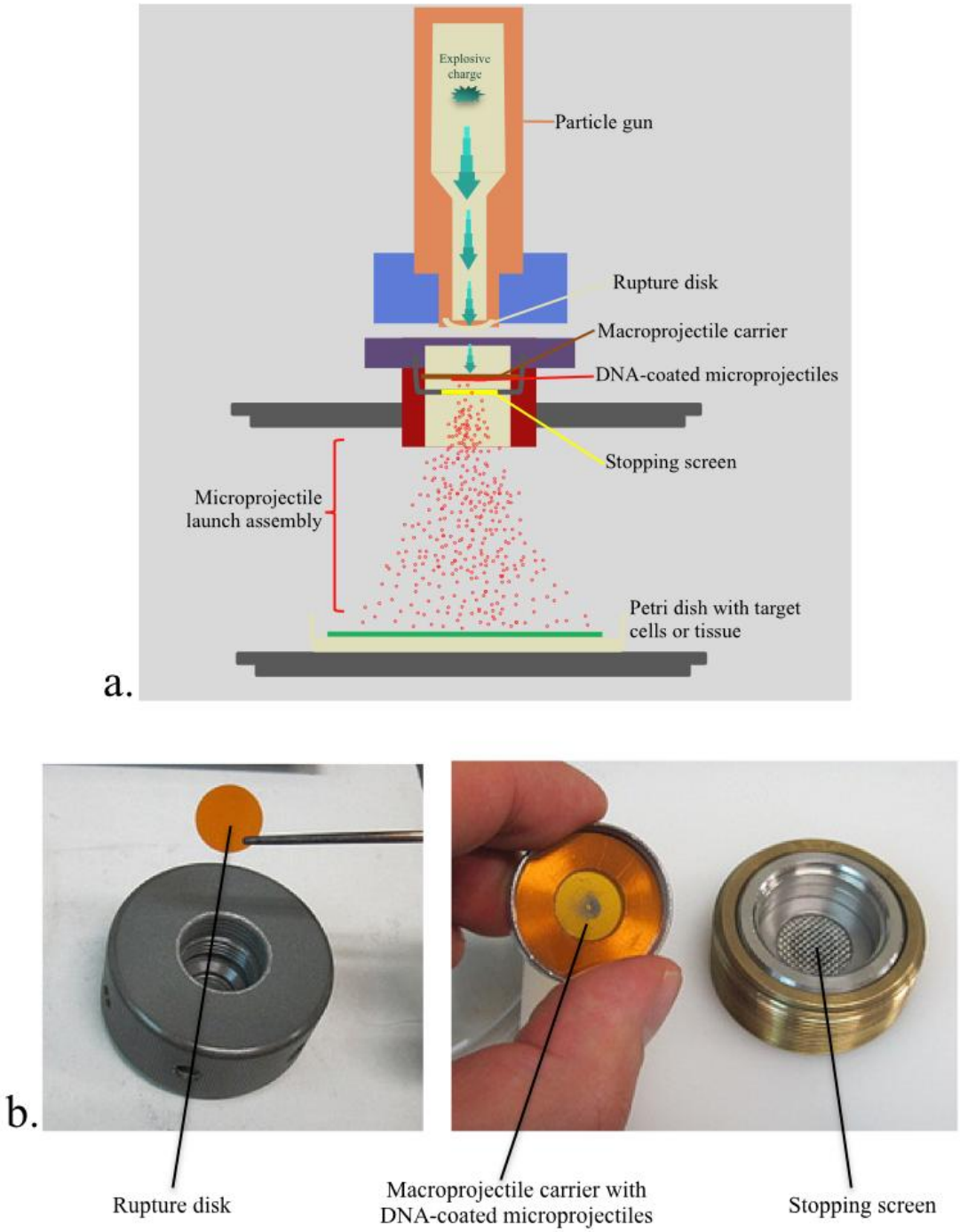


Figure 2-2 Biolistic particle delivery system.

a. Diagram presenting the principle of the biolistic bombardment method.
Adapted from: http://depts.washington.edu/genomelb/BiolisticBombardment&MiscV3.html
The optimisation process to increase the transformation efficiency involves adjustment of three parameters: the distance between the macroprojectile carrier and the target cell lawn, the helium pressure, and the particle size.
b. Assembly of the consumables used in the biolistic transformation method

2.5.12 Selection and analysis of homoplasmic transformants

Each colony representing a putative transformant was subjected to three re-streaking cycles on fresh selective plates to ensure homoplasmy (transgene is incorporated into every copy of the chloroplast genome in each transformed line). Afterwards, each transgenic line was subjected to PCR analysis to verify the integration of the transgenic DNA into the chloroplast genome. The details of this analysis are presented in chapter 4, 5, and 6.

2.5.12.1 Total RNA extraction

10 ml of algal cultures in mid-log phase (OD at 750 nm of approximately 0.6) were pelleted at 4^0C for 4 min at 2000 x *g*. The method was performed according to the TRI REAGENT® [Sigma] product information with some modifications. 1 ml of ice-cold TRI REAGENT® was added to each sample and the cells were lysed by pipetting. 0.2 ml chloroform was added and the solution was vortexed for 15 seconds. Afterwards the solution was centrifuged in a microcentrifuge [Centrifuge 5415R, Eppendorf] at maximum speed for 15 minutes at 4^0C and the upper aqueous phase was harvested into a fresh tube. RNA was precipitated with 0.5 ml isopropanol [Sigma] and centrifuged in the microcentrifuge at the maximum speed for 15 minutes at 4^0C. RNA precipitate was washed with 1 ml of 75% (v/v) ethanol [Sigma], air-dried and re-suspended in 87.5 µl of DNase and RNase free water. Afterwards the RNA solution was subjected to DNase treatment using the RNase-Free DNase Set [Qiagen] according to the manufacturer's instructions and protocol.

2.5.12.2 Complementary DNA (cDNA) synthesis and reverse transcription PCR (RT-PCR)

cDNA synthesis was performed using the First Strand cDNA Synthesis Kit [Fermentas, Thermo Scientific] according to the manufacturer's instructions and protocol using an oligo(dT)18 primer. The first strand of cDNA was directly used as a template in the reverse transcription PCR amplification using the specific primers listed in Chapter 5.

2.5.12.3 Preparation of whole cell protein extracts for western analysis

20 ml of *C. reinhardtii* was harvested after the cultures reached an OD at 750 nm of approximately 0.6. This was done by centrifugation at 4500 x *g* for 10 min. The pellet was then re-suspended in 0.6 ml of solution A containing 0.8 M TrisHCl pH 8.3, 0.2 M sorbitol, 1% β-mercaptoethanol). The samples were frozen in liquid nitrogen and stored at -80^0C.

Prior to protein electrophoresis, 100 µl of the crude protein extract was mixed with 10 µl of 10% (w/v) SDS and boiled for 1 min in order to denature the proteins. Subsequently, all cell debris was pelleted in the microfuge for 2 min at maximum speed. 80 µl of each sample containing solubilised proteins was loaded into a designated well on the SDS-polyacrylamide gel. PageRuler^{TM} prestained protein marker [Fermentas, Thermo Scientific] was used to estimate the size of the proteins. Samples were run for 3 hours at room temperature at 120 V using a PowerPac 300 [BioRad].

2.5.12.4 Chlorophyll removal from the protein extracts using an acetone precipitation method

Four times the sample volume of ice-cold acetone was added to each protein sample, vortexed for 15 s at the maximum speed and incubated for 60 minutes at -20°C. Subsequently, each sample was centrifuged at 4°C for 10 minutes at 13 000-15 000 x *g* and the protein pellet was briefly air-dried and dissolved in the appropriate volume of 1x SDS-loading buffer containing 0.3M Tris HCl pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 0.125% (w/v) bromophenol blue**.**

2.5.12.5 Western blot analysis

After completing SDS-PAGE and electrophoretic transfer (see section 2.4.6 for details), the membrane was subjected to western blot analysis using either the ECL-based system or using the Odyssey®infrared imaging system [Licor-Bioscience]. Both applications have identical protocol requirements.

The membrane was blocked overnight in blocking buffer containing TBS (20 mM Tris base, 137 mM NaCl, 1M HCl pH 7.4), 0.1% Tween-20, 5% skimmed milk at 4°C under

continuous agitation on orbital shaker [Orbital shaker SSM1, Stuart®]. Subsequently, the membrane was incubated with the primary antibody diluted in blocking buffer for 2 hours at room temperature under continuous agitation. The details of the antibodies used are presented in Table 2-9.

Following incubation with the primary antibody, the membrane was washed three times in TBS-T buffer containing TBS (20 mM Tris base, 137 mM NaCl, 1 M HCl pH 7.4), 0.1% Tween-20 and the membrane was incubated with the appropriate secondary antibody (see Table 2-9) diluted in blocking buffer for 1 hour at room temperature under continuous agitation. Afterwards, the membrane was washed three times in TBS-T and once in TBS buffer. Washed membranes incubated with DyLightTM 800-conjugated secondary antibody was ready to be analysed and quantified on the Odyssey®infrared imaging system according to the manufacturer's instruction. Membranes incubated with ECL™ Peroxidase labelled secondary antibody were subjected to development with the ECL detection kit [SuperSignal West Pico Chemiluminescencent substrate, Pierce] according to the manufacturer's instructions. After activation of the chemiluminescence signal on the secondary antibody, a sheet of Hyperfilm ECL [GE Healthcare] was exposed to the membrane in complete darkness and the film was developed using an automatic film processor.

Table 2-9 Primary and secondary antibodies and their dilution factors used for western blot analysis

Antibody	Source	Dilution factor
Primary antibody		
αHA	Rabbit polyclonal, Sigma-Aldrich	1:2000
α-HPV16 L1 [CamVir 1]	Mouse monoclonal, Abcam	1:2000
α-CVN	Polyclonal rabbit serum, kindly provided by Prof Julian K-C. Ma of St. George's University of London	1:1000
Secondary antibody		
ECL™ Peroxidase labelled anti-rabbit IgG, (from donkey)	GE Healthcare	1:5000
ECL™ Peroxidase labelled anti-mouse IgG, (from donkey)	GE Healthcare	1:5000
Goat Anti-Rabbit IgG (H+L) DyLight™ 800 conjugated	Thermo Scientific	1:15000
Goat Anti-Mouse IgG (H+L) DyLight™ 800 conjugated	Thermo Scientific	1:15000

2.5.12.6 Southern blot analysis

Chlorella sorokiniana genomic DNA was digested with the appropriate restriction enzyme and the DNA fragments were separated on a 1 % (w/v) agarose gel. Afterwards the gel was placed in denaturing solution (1.5 M NaCl and 0.5 M NaOH) for 30 min and then washed briefly with distilled water. The gel was then incubated twice for 20 min in the neutralizing solution (1 M Tris, pH 7.4, 1.5 M NaCl). The DNA from the gel was transferred overnight to an uncharged nylon membrane (Hybond-N, GE Healthcare) according to the protocol (Sambrook and Russell, 2001). Subsequently, DNA was fixed to the membrane by baking at 80^0C for 1 hour. DNA labelling and detection using the DIG High Prime DNA Labelling and Detection Starter Kit II [Roche Diagnostics GmbH] was done according to the manufacturer's instructions.

2.6 Hydrocarbon analysis

2.6.1 Hydrocarbon extraction

2.6.1.1 Supernatant extracts from algal cells

After centrifuging the cells, 5 ml of the supernatant was added to 5 ml of ethyl acetate and the mixture was vortexed for 1 minute at maximum speed. The sample was distributed into several 1.5 ml tubes and centrifuged in a microfuge at 4^0C for 1 minute at maximum speed. Top phase was collected and analysed by gas chromatography (GC).

2.6.1.2 Cell extracts

Approximately 20 mg of wet cell pellet was re-suspended in 1 ml of Tris HCl pH 7.4 containing 0.01% Triton100 and the mixture was sonicated for 10 pulses of 10 seconds each with 10 seconds on-ice incubation between each pulse. Afterwards, 1 ml ethyl acetate was added and the mixture was vortexed for 1 minute at the maximum speed. The debris of each sample was pelleted following centrifugation in a microfuge at 4^0C for 1 minute at maximum speed. The top phase was collected and analysed on the GC.

2.6.2 Gas Chromatography analysis

Gas Chromatography (GC) analysis of the extracted hydrocarbons was performed on a 6850 Network GC System [Agilent Technologies, US] equipped with a flame ionization detector (FID). The samples were analysed using an Omegawax 250 (30 m long; 0.25 mm internal diameter, 0.25 μm film thickness) capillary column using helium as the carrier gas flowing at a rate of 30 cm s^{-1}. A 7683B series Injector [Agilent Technologies] was used to inject 1 μl of each sample in split mode at a 100:1 ratio. The temperature of the injector and the detector were maintained at 250^0C and 260^0C respectively. The temperature oven was held at 50^0C for 1 min, ramped to 150^0C at 10^0C min^{-1} followed by a linear increase of 2^0C min^{-1} up to 260^0C with a final hold for 4 min. The retention times of product peaks were compared to the pooled standards purchased from Sigma listed in Chapter 5.

CHAPTER 3. EVALUATION AND GROWTH OPTIMIZATION OF *CHLORELLA SOROKINIANA* AS A SOURCE OF BIOFUEL

3.1 Introduction

The World's fossil oil consumption has increased drastically over the past few decades as the transport sector relies mainly on oil and gas. Based on the observation of the oil consumption it has been forecasted that a peak in global conventional oil production is inevitable by the year 2030 (Okullo and Reynes, 2011). Therefore, it is essential to look for a renewable alternative to fossil fuel. Current renewable oil production is dominated by oilseed crops and meets only a fraction of present needs and it also takes the valuable land space from crop plants. Consequently, an idea proposed in the late 1960s (Oswald and Golueke, 1960) to exploit algal cells as factories for biofuel production has an increased number of followers in recent years (Chisti, 2008, Brennan and Owende, 2010). *Botryococcus braunii* has been of great interest to microbiologists for years researching the possibility of exploiting algae for biofuel (Brown et al., 1969). This green colonial microalga is capable of producing considerable amounts of lipids, but more importantly also long-chain hydrocarbons accounting for over 60% of the dry weight (Metzger and Largeau, 2005). The growth rate of this species however, is so slow that it has become clear it can never be exploited on a commercial scale. Therefore, the new focus is currently on other *Chlorophyta* that are fast growing, and have a high lipid content, such as *Dunaliella*, *Chlamydomonas*, *Nannochlori*s or *Chlorella*.

There are several advantages of using algae as a source for biofuel when compared to plants. Firstly, the oil productivity of microalgae is approximately 15-300 times higher than in plants (Schenk et al., 2008) as the harvesting cycle is between 1-10 days depending on species whereas for plants this number is 1-2 times a year. Such fast growth rates open the possibility for continuous harvest of algal cultures and significantly increased oil yields over shorter periods of time. The maintenance of algal cultures requires considerably less water than terrestrial crops (Dismukes et al., 2008) moreover, some species are able to tolerate a certain level of salinity (Stengel et al., 2011) and this ability could reduce substantially the running cost of algal culture since fresh water can be the limiting factor in the areas most suitable for algal growth.

Large-scale cultivation can be achieved on algal farms established on non-arable land and through biofixation of CO_2 they can greatly reduce greenhouse gas emission (Searchinger et al., 2008). In addition to biofuel production, algal farms can be employed simultaneously in wastewater treatment (Woertz et al., 2009) as the organic waste can potentially be the source of nutrients such as N and P. Also the ability to biofixate CO_2 from industrial flue gas can increase the positive impact on the environment. Furthermore, algal cultures can be also a potential source of valuable products such as nutritional supplements, proteins, antioxidants, and pigments (Spolaore et al., 2006), an enrichment source for animal feed or fertilizer (Anemaet et al., 2010) and also as a substrate in bio-ethanol and biogas production (Collet et al., 2011).

In nature algae store the energy in the form of starch or lipids and the interest currently is focused on microalgal strains accumulating lipids, particularly non-polar triacylglycerols (TAGs) since they are the main substrate for biodiesel production. During the oil to biodiesel conversion, TAGs are subjected to a transesterification process to produce fatty acid (FA) methyl esters (FAMEs) (Sheehan, 1998).

Figure 3-1 presents this schematic reaction.

The total lipid content varies from strain to strain and it can be between 1-75% of the dry weight (Rodolfi et al., 2009). It is however, on average approximately 40% of the dry weight. This value is usually achieved by nitrogen depletion in the growth medium – a well known lipid-induction method since the middle of last century (Spoehr and Milner, 1949). It has also been reported that increased light intensity may not only induce the accumulation of TAGs but also the proportion of polyunsaturated fatty acids (PUFA) to monounsaturated and saturated FA decreases significantly (Hu et al., 2008a). Since the polyunsaturated oils are particularly susceptible to oxidation whilst in storage, the high irradiance during algal growth could potentially increase the quality of algal oils.

Figure 3-1 Production of biodiesel from storage lipids extracted from plants or algae.

Adapted from (Scott et al., 2010)

Schematic diagram of the esterification reaction of triaclylglycerides for fatty acid methyl ester production.

The lipid content is one of the major factors when choosing the right algal strain for biofuel production. However, the growth and biomass production rates are equally important. Reports from 1950-60s (Sorokin and Myers, 1953b, Sorokin and Krauss, 1959, Sorokin and Krauss, 1962, Sorokin, 1967) drew our attention to a thermotolerant alga, *Chlorella sorokiniana*, whose doubling time can reach 2.5 hours in optimal temperature and light conditions. Based on the recent reports it has been calculated that potential lipid productivity of this strain could reach 100 mg/L/day (Griffiths and Harrison, 2009).

On the other hand, optimal light utilization by the photosynthetic apparatus of algal cells is essential for obtaining the maximum biomass. It has been estimated that photosynthetic efficiency reaches only 8-10% (Hambourger et al., 2009) and several factors of the light-dependent process contribute towards this inefficiency (Figure 3-2). Firstly it has been estimated that only 50% of the energy that reaches the Earth and falls on a plant is captured by the light harvesting complexes (LHC) of the chlorophyll antenna as the absorption range of the photosynthetically active radiation falls between 400-700 nm. (Kruse et al., 2005). Secondly, plants and algal cultures maintained under high irradiance receive more solar energy than they are able to use and the excess of the captured energy is dissipated as heat or fluorescence (Muller et al., 2001). Additionally, a process of self-shading within dense algal cultures contributes towards the inefficiency of solar light utilisation as algae are mostly maintained in mass culture and they form a thick layer of cells where the surface layer receives excess light but the light penetration to the deeper layer is restricted (Polle et al., 2003). Therefore, interest in optimising the light-dependent reactions has recently increased (Polle et al., 2002, Wu et al., 2011a, Work et al., 2011) and will be also discussed and evaluated later in this chapter.

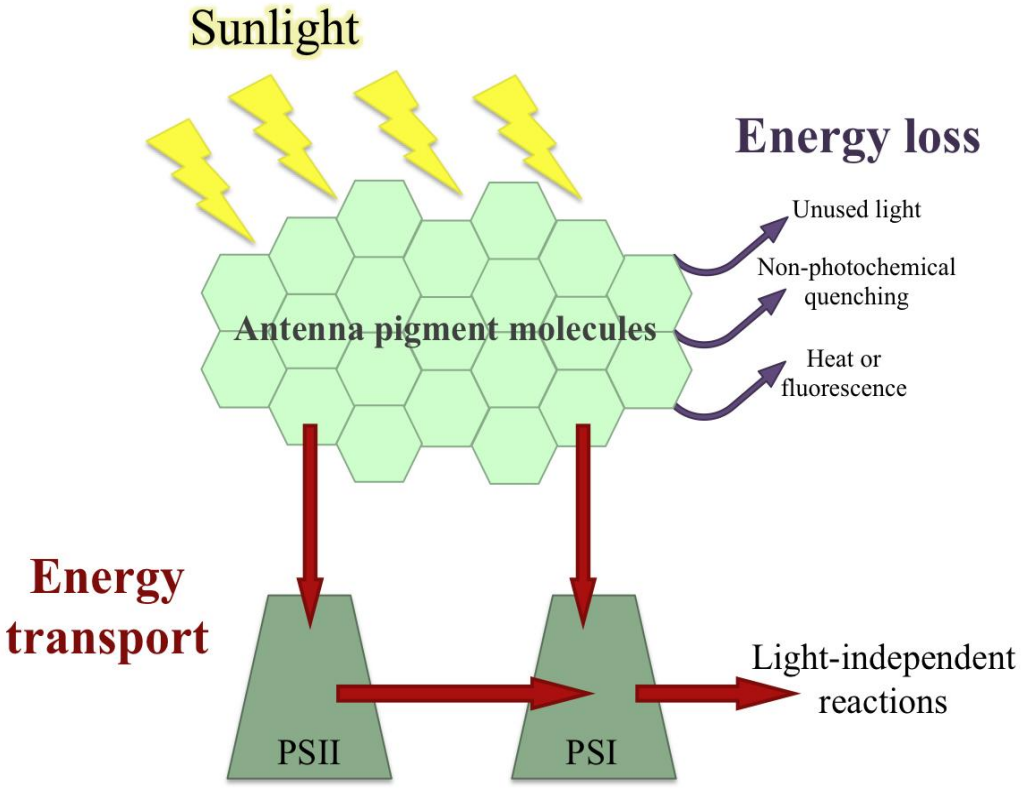


Figure 3-2 Solar energy balance captured by antenna molecules.
Within the thylakoid membranes of the algal chloroplasts the solar energy captured by the chlorophyll antenna molecules is transferred as excitation energy to the reaction centres of the PSII and PSI. The captured energy can be subsequently converted to the chemical energy and used in light-independent reactions of photosynthesis or can be emitted as heat or fluorescence or in a form of non-photochemical quenching.

3.2 Background and aims of the project

An essential factor in algal biofuel research is strain selection and the key consideration is the ability to accumulate a substantial amount of lipids. Several algal species have been selected in recent years and assessed for their potential for oil production (Wilhelm and Jakob, 2011, Griffiths and Harrison, 2009, Graham et al., 2012, Chen et al., 2012). Furthermore, the selected strain should have a competitive advantage over competing species with high growth rate and increased biomass production. Additionally, this strain should tolerate extreme conditions such as elevated and low temperatures and changing light conditions.

In this work, the optimal growth conditions for biofuel production were established for one of the fastest growing microalgal strains. The growth rates of four *Chlorella* strains were measured under various settings including macronutrients, temperature and light conditions in order to establish the limiting features in their lipid production. Since fresh water availability is one of the critical factors of algal growth, the possibility for using seawater for algal cultivation were also assessed. Additionally, in order to maximise the light utilization during photosynthesis, chlorophyll antenna mutants of the fast growing *Chlorella sorokiniana* strain were created and their adaptability to limited light conditions was evaluated.

3.3 Results and Discussion

3.3.1 Development of molecular tools for *Chlorella sorokiniana* strain identification – MICS method [Molecular Identification of *Chlorella sorokiniana* strains]

Currently the algal collection in the Purton lab consists of three strains of *Chlorella sorokiniana* – UTEX1230, H–1983 and H–1986 (see Chapter 2 for reference). Taxonomic identification of *Chlorella* species is considerably difficult as all species from this genus are characterised by an asexual reproductive cycle combined with the lack of obvious morphological features. Therefore the identification of the *Chlorella* strains is very challenging especially if the culture is maintained in open systems where contamination with unwanted strains can easily occur. The traditional identification methods involve taking into consideration morphological, physiological and biochemical features. The morphological identification method is not truly reliable as the cell size and shape tends to change depending on the environmental factors (Agustí, 1991, Parsons and Takahashi, 1973) whereas physiological and biochemical methods are time consuming and require remarkable precision (Kessler, 1978). A more efficient identification involves a PCR-based method to determine the differences in conserved and variable regions within the chloroplast-encoded 16S rDNA sequence (Burja et al., 2001, Wu et al., 2001).

Three single nucleotide differences were discovered in the 480 bp region of the 16S sequence of *Chlorella sorokiniana* strains UTEX1230, H-1983 and H-1986 and amplified using primers CS1 and CS2 designed by Wu et al. (2001) as seen in Figure 3-3.a. The observed differences are sufficiently significant to be useful in designing a new molecular technique for *C. sorokiniana* strain differentiation: MICS (Molecular Identification of *Chlorella sorokiniana* strains). This method involves digestion of the amplified 480 bp fragment of the 16S rDNA gene of the *C. sorokiniana* strain with *Apo*I and *Hpy*188I restriction enzymes and the expected DNA fragment pattern is presented in Figure 3-3.b. H-1983 strain is characterised by the presence of a 240 bp fragment when its 16S rDNA PCR product is digested with *Hpy*188I whereas both UTEX1230 and H-1986 strains lack this band but possess 160 bp and 85 bp bands. 150 bp, 60 bp and 22 bp bands are typical for all three *C. sorokiniana* strains. Digestion of 16S rDNA with *Apo*I results in presence of 130 bp and 60 bp bands in all three tested strains. A 285 bp fragment is characteristic

for H-1986 whereas 200 bp and 85 bp fragments are specific for UTEX1230 and H-1983 strains.

The digested DNA fragments are then separated on a polyacrylamide gel. Each *Chlorella sorokiniana* strain is characterised by a different pattern of the digested DNA fragments and the results are presented in Figure 3-3.c. The MICS method turned out to be useful in further tests of *Chlorella sorokiniana* strains, namely the competition assay. As stated earlier, contamination with unwanted strains can easily occur in open algal cultivation systems and if this happened it would be convenient to test which *C. sorokiniana* strain withstands the competition with other strains. To assess which *C. sorokiniana* strain is the most robust, media were inoculated with 5×10^7 cells of each strain and maintained at 25^0C, 32^0C and 39^0C for either 7 or 12 days. 5ml of each culture was then harvested, genomic DNA extracted, and PCR of 16S rDNA using the CS1 and CS2 primers was performed.

Figure 3.4 shows the pattern of the 16S bands after digestion with *Apo*I and *Hpy*188I. The UTEX1230 strain appears to be present in 7-day cultures maintained at all temperature ranges, however, it disappears in 12-day cultures. The H-1986 strain is present in both 7 and 12-day cultures in all temperature ranges and becomes dominant in older cultures. The H-1983 strain grows well at 25^0C and 32^0C, yet 32^0C appears to be the maximal temperature for its growth. 7-day cultures at 39^0C contain both the UTEX1230 and H-1986 strains and longer maintenance at this temperature results in the generation of a H-1986 monoculture.

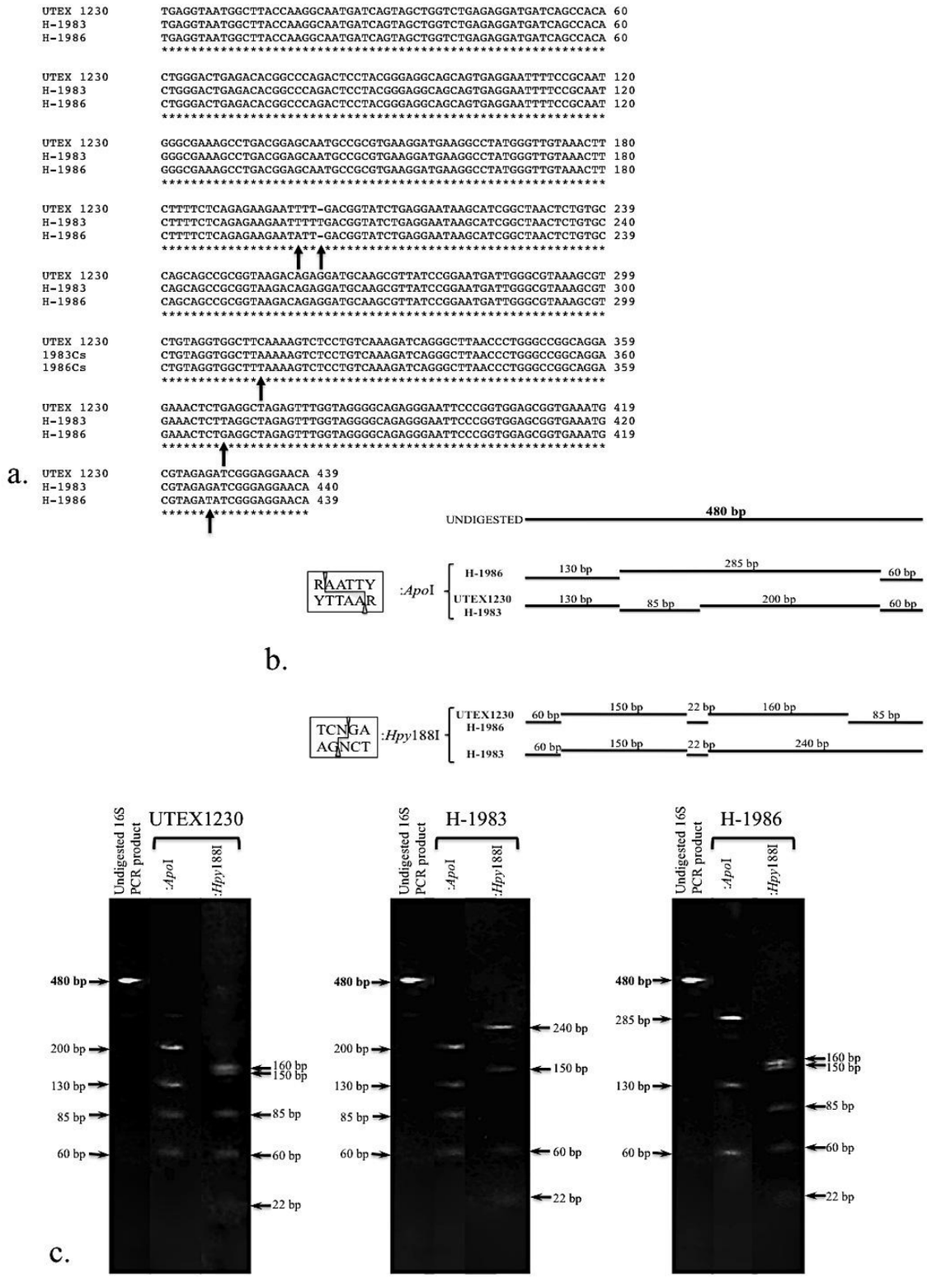


Figure 3-3 The mechanism of *C. sorokiniana* strain identification using the MICS method [Molecular Identification of *Chlorella sorokiniana* strains]

a. Alignment of the sequenced 480 bp 16S PCR product of the three *C. sorokiniana* strains: UTEX1230, H-1983 and H-1986; arrows show the differences in sequence between each strain
b. Predicted DNA fragment sizes resulted in digestion of 16S PCR product of *C. sorokiniana* strains UTEX1230, H-1983 and H-1986 with *Apo*I and *Hpy*188I restriction enzymes
c. Polyacrylamide gel electrophoresis of the 480 bp fragment of 16S PCR products of the three *C. sorokiniana* digested with *Apo*I and *Hpy*188I restriction enzymes.

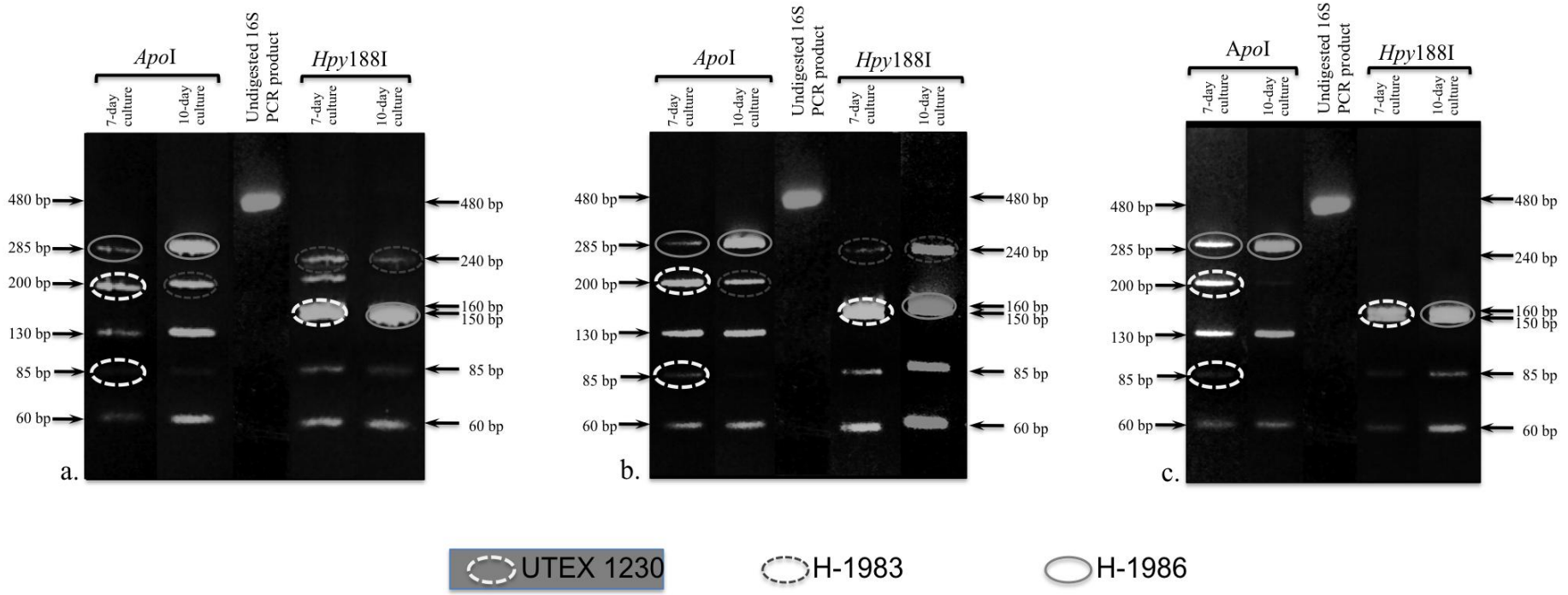


Figure 3-4 Polyacrylamide gel electrophoresis of the 16S PCR product digested with *Apo*I and *Hpy*188I

PCR using the CS1 and CS2 primers was performed on DNA isolated from 7 and 12-day old cultures inoculated with all three *C. sorokiniana* strains. The bands characteristic for each strain are circled. Conditions a), b), c) represent temperatures 25^0C, 32^0C, 39^0C respectively.

Several other identification methods for microalgae have been developed in recent years, for instance an immunofluorescence assay (Rodgers et al., 1995), a resonance Raman method (Brahma et al., 1983), flow cytometry (Balfoort et al., 1992) or bright field microscopy (Pech-Pacheco et al., 2001). Nevertheless, all of the listed methods require specialised and very expensive equipment. The created MICS method takes another step from the established PCR-based method and allows for a quick and reliable scan of the culture for *C. sorokiniana* strains, yet it does not require any additional equipment.

3.3.2 Comparison of the growth rates and biomass production of four *Chlorella* strains

Research by Constantine Sorokin performed over fifty years ago (Sorokin and Krauss, 1959, Sorokin and Myers, 1953b) reported fast-growing *Chlorella sorokiniana* strains tolerating elevated temperatures up to 40^0C. These reports suggest that this species could potentially be used on a commercial scale because of its very high biomass productivity. Three *C. sorokiniana* strains (UTEX1230, H-1983 and H-1986) were evaluated for growth and biomass production. *Chlorella vulgaris* was also included in this analysis since the vast majority of reports assessing the *Chlorella* genus in terms of growth were based on this species (Heredia-Arroyo et al., 2011, Collet et al., 2011, Kumar et al., 2010, Powell et al., 2009, Liang et al., 2009).

3.3.2.1 Optimal temperature conditions for growth of *C. sorokiniana* strain

In order to select the best performing algal strain, cultures of these four *Chlorella* strains were grown under a wide temperature range [16.5^0C, 20^0C, 25^0C, 30^0C, 35^0C, 39^0C, 42^0C] until they reached the stationary phase. The presented growth rate [μ, h^{-1}] values were calculated based on the measurements during the logarithmic growth phase, cultured in TAP medium. The μ value for all four *Chlorella* strains at each temperature condition is shown in Figure 3-5. These values were the highest for UTEX1230 in all temperature conditions up to 39^0C with 35^0C being optimal, where the μ, h^{-1} reached 0.2.

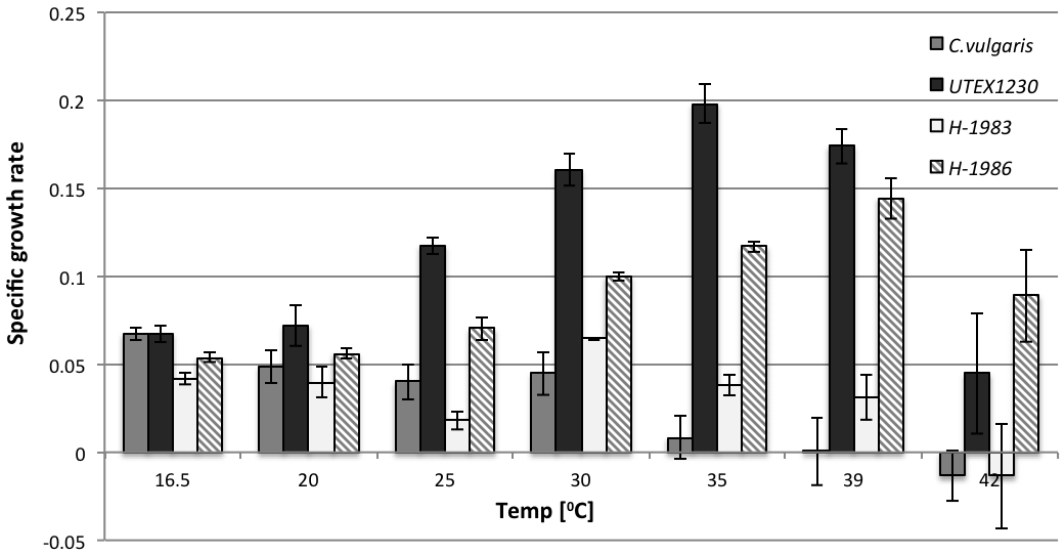


Figure 3-5 Specific growth rates of four *Chlorella* strains.

The specific growth rates [μ, h^{-1}] was calculated for *C. vulgaris* and three strains of *C. sorokiniana*: UTEX1230, H-1983 and H-1986 at seven temperature points 16.5^{0}C, 20^{0}C, 25^{0}C, 30^{0}C, 35^{0}C, 39^{0}C, 42^{0}C.

Error bars show ±SD (n=3).

The UTEX1230 strain was able to reach the stationary phase within less than 20 hours at 35^0C – this was established to be the optimal growth temperature. It took approximately 21 hours at 39^0C to reach the stationary phase and 29 hours at 30^0C. The specific growth rates [μ, h^{-1}] below 30^0C and above 40^0C decreased significantly when compared to the rates at the optimal temperature, reaching on average 0.117 at 25^0C, 0.072 at 20^0C and 0.067 at 16.5^0C. 42^0C appears to be the upper temperature limit, as the specific growth rate was noticeably reduced to approximately 0.045. Figure 3-6 presents the compiled growth characteristics at the temperature range 16.5^0C – 42^0C.

The H-1986 strain appears to have an extended upper temperature limit with 39^0C being the optimal temperature for its growth, nevertheless, the specific growth rate at this temperature it is still lower for H-1986 than for UTEX1230 (μ, h^{-1} of approximately 0.14 and 0.175 respectively). At 42^0C H-1986 strain grows faster than UTEX1230 where the μ value reached 0.089 and 0.045 respectively. The third of the tested strains – H-1983 grows much slower than any other *Chlorella* strain tested in this experiment. Its growth is not affected by increasing temperature, as the μ values are comparable in the temperature range 16.5^0C-39^0C. *Chlorella vulgaris* does not appear to endure temperatures above 35^0C and its growth rate at 16.5^0C was comparable to the growth of UTEX1230 strain, with a specific growth rate value of approximately 0.07 for both strains. In conclusion, the optimal growth temperature of H-1986 strain (39^0C) is higher than that of UTEX1230 (35^0C) whereas for H-1983 and *C. vulgaris* this value is lower, reaching 20^0C and 16.5^0C respectively.

The standard temperature used when establishing the optimal parameters for *Chlorella sp.* growth is typically 25^0C however, it is often close to the upper limit for the growth of most *Chlorella* strains that on occasions can be raised to 30^0C (Kessler, 1985).

A search for thermotolerant strains for culturing in open pond systems and in photobioreactors is financially justified since one of the biggest maintenance cost involves cooling the cultures (Fishman et al., 2010) and this cost is based on the energy consumption and water demand. Several *Chlorella* strains not only struggle to tolerate temperatures above 25^0C but also show abnormal pigmentation when cultured in higher temperatures. For instance, *C. zofingiensis* becomes orange as a result of secondary carotenoid formation and *C. protothecoides* becomes white (Kessler, 1985).

Various optimal temperature conditions for *Chlorella sorokiniana* growth have been reported. The most common temperature chosen as optimal for *Chlorella sorokiniana* is 39^0C (Sorokin and Krauss, 1962, Sorokin, 1971, Eyster, 1978), 38^0C (Heath, 1979), 40^0C (de-Bashan et al., 2008) whereas other records recommend 37^0C (Cuaresma et al., 2009). When establishing the optimal temperature conditions for the species, the strain selection plays a significant role nonetheless, it seems the combination of light, medium and contaminants is equally important (de-Bashan et al., 2008).

C. reinhardtii has been the most popular single-celled algal model organism for decades (Harris, 2001) and it has recently been recognized as a candidate host for biofuel production (Morowvat et al., 2010). Additionally, there has been increased interest in *C. vulgaris* as a potential source of lipids for biofuel production (Widjaja et al., 2009). Therefore, the productivity of the three microalgal species was compared. The average cell size differs between each strain with *C. sorokiniana* having a cell diameter of approximately 2-4 μm, *C. vulgaris* 3-5 μm and *C. reinhardtii* 10-12 μm. In order to determine whether the cell size might influence the biomass production of *Chlorella vulgaris*, *Chlorella sorokiniana* and *Chlamydomonas reinhardtii*, the continuous cultures were grown in optimal temperature conditions for each strain and their productivity is presented in Figure 3-7. *C. vulgaris* and *C. reinhardtii* strains were cultured at 25^0C whereas *C. sorokiniana* UTEX1230 at 35^0C. As a result, the UTEX1230 strain shows the highest productivity with the biomass yield reaching 0.8g of dry weight/24 hours whereas this value was 4 fold and 10 fold lower for *C. vulgaris* and *C. reinhardtii* respectively.

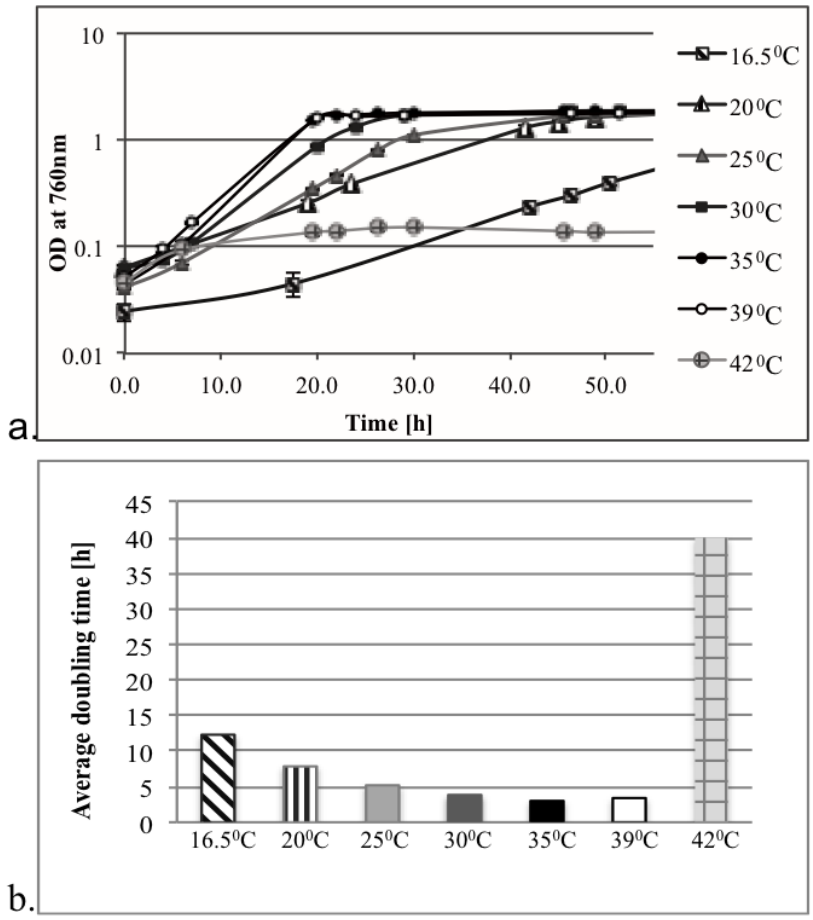


Figure 3-6 Growth rates measurement of *C. sorokiniana* UTEX1230 between 16.5^{0}C and 42^{0}C.

a. Growth curves representing the growth rates at 7 temperature points within 16.5^{0}C and 42^{0}C. Symbols and error bars show means ±SD (n=3)
b. Average doubling time calculated for each temperature point

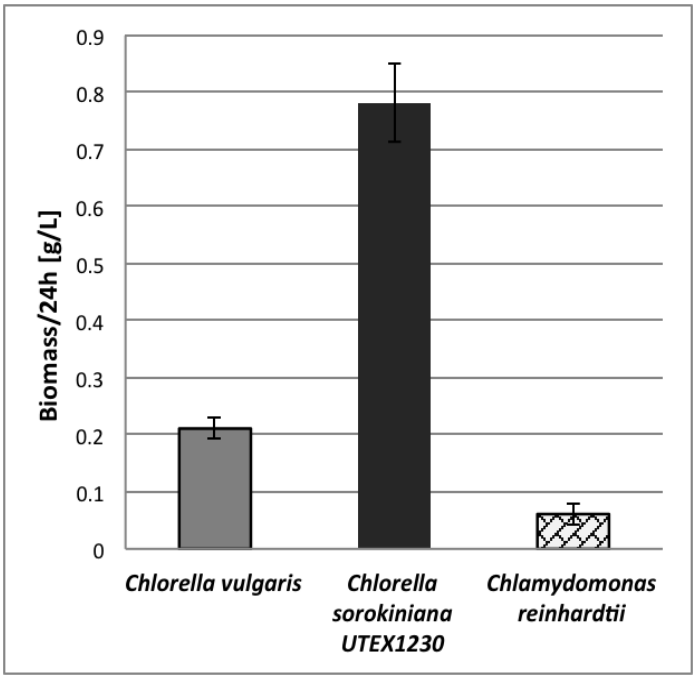


Figure 3-7 Biomass production of *C. vulgaris, C. sorokiniana*UTEX1230 and *C. reinhardtii*.

The biomass was calculated over 24 hours from the beginning of the log phase of growth. Error bars show ±SD (n=3)

3.3.2.2 Medium requirements for optimal growth of *C. sorokiniana* UTEX1230

3.3.2.2.1 Optimal minimal medium

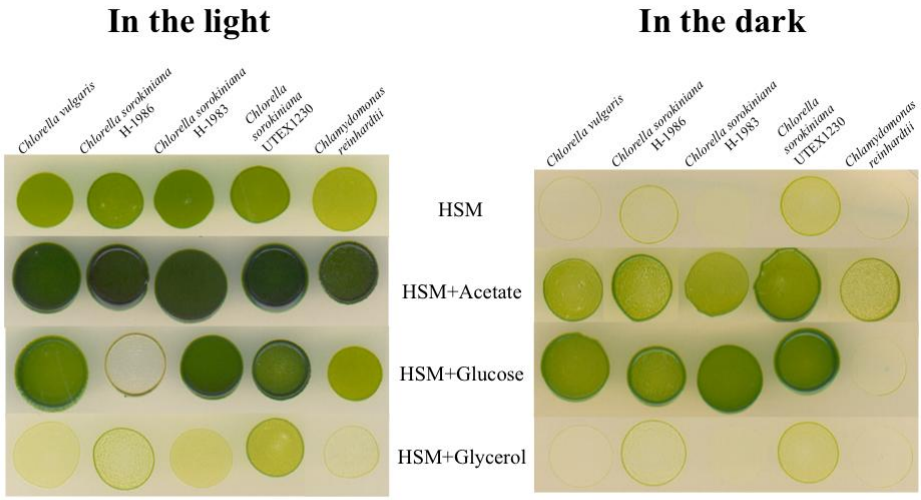
The mineral composition of the growth medium is vital for achieving the highest growth rates of any organism. Three minimal growth media recommended in the literature were tested: 3N-BBM+V, HSM and JM in the same temperature [35^0C] and light intensity conditions [50 µmol m^{-2} s^{-1}]. Specific growth rate was shown to be the highest in HSM [µ, h^{-1}= 0.075], whereas 3N-BBM+V appears to be the least optimal for *C. sorokiniana* growth [µ, h^{-1}= 0.02].

Bacterial contamination frequently stimulates algal growth (Mouget et al., 1995). Therefore, the effects of lack of sterilization step on the growth of *C. sorokiniana* was tested. It was determined that there is no difference in growth rates between the cultures maintained in autoclaved and un-autoclaved media and that was true for both minimal and organic carbon–enriched conditions (data not shown). Bacterial contamination is a common incident in large-scale cultures, particularly in open pond systems; therefore showing that sterile conditions are not required for the *C. sorokiniana* growth is a big step towards establishing the suitable conditions for the scaled-up cultures.

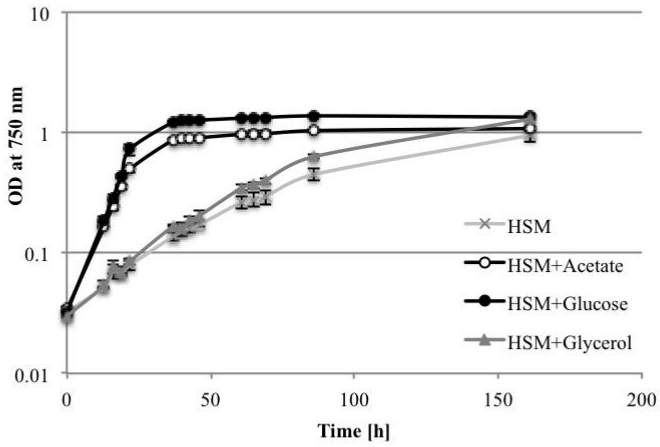
3.3.2.2.2 Carbon source

The increasing interest in culturing algae for biofuel production also raised interest in other algal applications such as production of high value products (e.g. recombinant proteins) where algal cultures would utilise the organic carbon in closed containers in dark. The quality and yield of the product of interest in this case are more important than increased maintenance costs. Therefore, the growth rates of *C. sorokiniana* UTEX1230 in minimal medium (HSM) were measured and compared to growth in either TAP or HSM with 3.67mM of sodium acetate, glucose or glycerol added (Figure 3-8 a and b). Figure 3-8.a presents the growth tests where all four *Chlorella* strains and *C. reinhardtii* were included.

a.



b.



c.

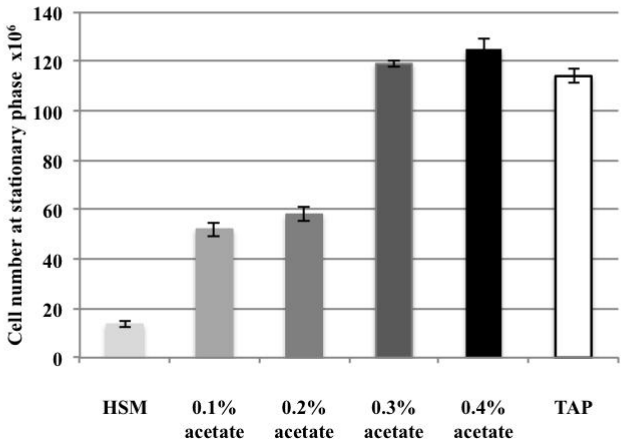


Figure 3-8 Analysis of algal growth on various types of carbon source.

a. Spot tests of *C. vulgaris*, *C. reinhardtii* and *C. sorokiniana* strains: H-1986, H-1983, UTEX1230 on minimal medium and with 3.67 mM of sodium acetate, glucose or glycerol added and maintained under constant illumination or in dark.
b. Growth curves of UTEX1230 strain in minimal medium and with 3.67 mM of sodium acetate, glucose or glycerol added. Symbols and error bars show means ±SD (n=3)
c. Cell density at the stationary phase in minimal medium HSM and with 0.1%, 0.2%, 0.3% or 0.4% of sodium acetate added. Error bars show ±SD (n=3)

Experiments a,b.c were performed at constant illumination and average light intensity 50 µmol m^{-2} s^{-1}.
HSM – minimal medium.

Supplementing algal growth using fixed carbon in the form of cheap sugars might aid biofuel production. Hence, attempts were made to determine if *C. sorokiniana* is efficient in obtaining energy from this type of carbon source. All tested algal strains are able to utilise glucose as seen in Figure 3-8.a where *C. reinhardtii* served as negative control since this strain lacks the hexose symporter protein that enables glucose uptake from the growth medium (Doebbe et al., 2007). *C. sorokiniana* H-1986 growth is clearly inhibited when maintained on medium containing glucose in light, whereas it can efficiently utilise glucose when kept in dark. It has been shown in several microalgal strains that light can inhibit glucose uptake and growth in mixotrophic conditions whereas this process was unaffected in heterotrophic conditions (Nicolas et al., 1980, Lalucat et al., 1984, Kamiya and Kowallik, 1987). The most promising strain in terms of growth – *C. sorokiniana* UTEX1230 does not have any of the above-mentioned limitations as it grows happily on both acetate and glucose in light and dark. Consequently, the influence of both carbon sources on growth of UTEX1230 were examined. Figure 3-8.b presents growth curves of *C. sorokiniana* UTEX1230 grown in the light on all three types of organic carbon: sodium acetate, glucose and glycerol compared to growth in minimal medium (HSM). Adding glycerol to the minimal medium did not affect growth rates of UTEX1230. Growth rates of UTEX1230 cultures containing acetate and glucose were almost identical, however, cell density was higher in cultures containing glucose. On the other hand, cell density of the UTEX1230 culture in the stationary phase was increasing with increasing concentration of the organic carbon in the growth medium (Figure 3-8.c). Growth rate and cell number at the stationary phase was measured in cultures maintained in HSM (negative control) and HSM supplemented with various concentrations of sodium acetate: 0.1%, 0.2%, 0.3% or 0.4%. Growth rates of every culture supplemented with any concentration of acetate were almost identical (data not shown). Nevertheless, there was a direct correlation between the acetate concentration and cell density at the stationary phase (Figure 3-8.c). Also, the acetate concentration is directly proportional to the time when cultures reach the stationary phase (data not shown). In conclusion, the concentration of the organic carbon in growth medium not only influences the longevity of the culture but also the productivity of the algal culture. *C. sorokiniana* UTEX1230 strain – as will be shown and discussed later in this chapter – suppresses photosynthetic activity in the presence of any available organic carbon source.

3.3.2.3 Effect of irradiance on growth of *C. sorokiniana* UTEX1230

Light is the one of the most important limiting factors during algal growth, particularly during phototrophic growth. The effect of irradiance on growth rate is presented in Figure 3-9. Mixotrophic growth (in TAP medium under constant illumination) is more suitable for the *Chlorella sorokiniana* UTEX1230 strain than photo- (in HSM medium under constant illumination) or heterotrophic (in TAP medium in the dark), since the highest growth rates were observed when the carbon source is supplied in the form of sodium acetate (TAP medium).

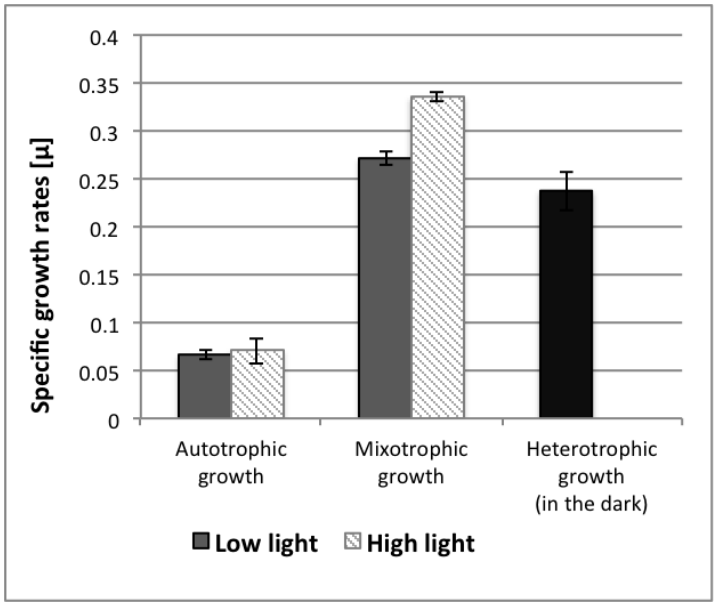


Figure 3-9 Influence of the irradiance intensity on the growth rates of *C. sorokiniana* UTEX1230.

The phototrophic (in HSM medium) and mixotrophic (in TAP medium) cultures were maintained at two light conditions: low (approximately 50 µmol m^{-2} s^{-1}) and high (approximately 120 µmol m^{-2} s^{-1}). Heterotrophic growth (in TAP medium) was performed in the dark.
Error bars show ±SD (n=3)

A substantial increase in UTEX1230 growth rate was observed when the organic carbon was present in the growth medium compared to phototrophic growth when all cultures were illuminated with 50 µmol m^{-2} s^{-1}. There was however, no significant increase in growth rate when the mixotrophic culture was illuminated with 50 µmol m^{-2} s^{-1} compared to the culture maintained in the dark suggesting that growth under these conditions is principally due to utilisation of the fixed carbon. On the contrary, light intensity appears to have a high impact on mixotrophic growth of the UTEX1230 strain where increasing light intensity during the exponential phase from 50 µmol m^{-2} s^{-1} to 120 µmol m^{-2} s^{-1} resulted in a rise of the specific growth rate of approximately 1.24 fold. This increase in growth rate related to elevated irradiance was not observed during the phototrophic growth on minimal medium.

Depending on the natural habitation, every algal strain has their own optimal light requirements and light saturation values. Culture density and overshadowing play a significant role and that is predominantly the reason why algae do not reach the optimal growth levels. The higher the light intensity, the higher the chance that the light will penetrate deeper within the culture, yet excessive illumination may result in photoinhibition. Hence establishing the maximum value of irradiation is equally important. In general, green algae tend to tolerate higher irradiance than red algae (Figueroa et al., 1997, Arnold and Murray, 1980). *C. sorokiniana* has been reported to endure very high illumination (Cuaresma et al., 2009). It was discovered however, in laboratory conditions light intensity of 300 µmol m^{-2} s^{-1} substantially inhibited growth rate of UTEX1230 strain (data not shown).

3.3.3 Tolerance to salinity of *C. sorokiniana* UTEX1230

The possibility of using seawater to partially offset the freshwater requirement of culturing *Chlorella sorokiniana* was investigated. It has been reported that increased salinity can affect the growth of freshwater algae (Reynoso and de Gamboa, 1982). Since the salinity tolerance is specific to each species, the growth of the *Chlorella sorokiniana* UTEX1230 strain was measured in a wide range of sodium chloride concentrations. The results are presented in Figure 3-10. The cultures were maintained in TAP medium supplemented with 0.5%, 1% 1.75% or 3.5% NaCl to mimic the different level of salinity that is observed in various environments.

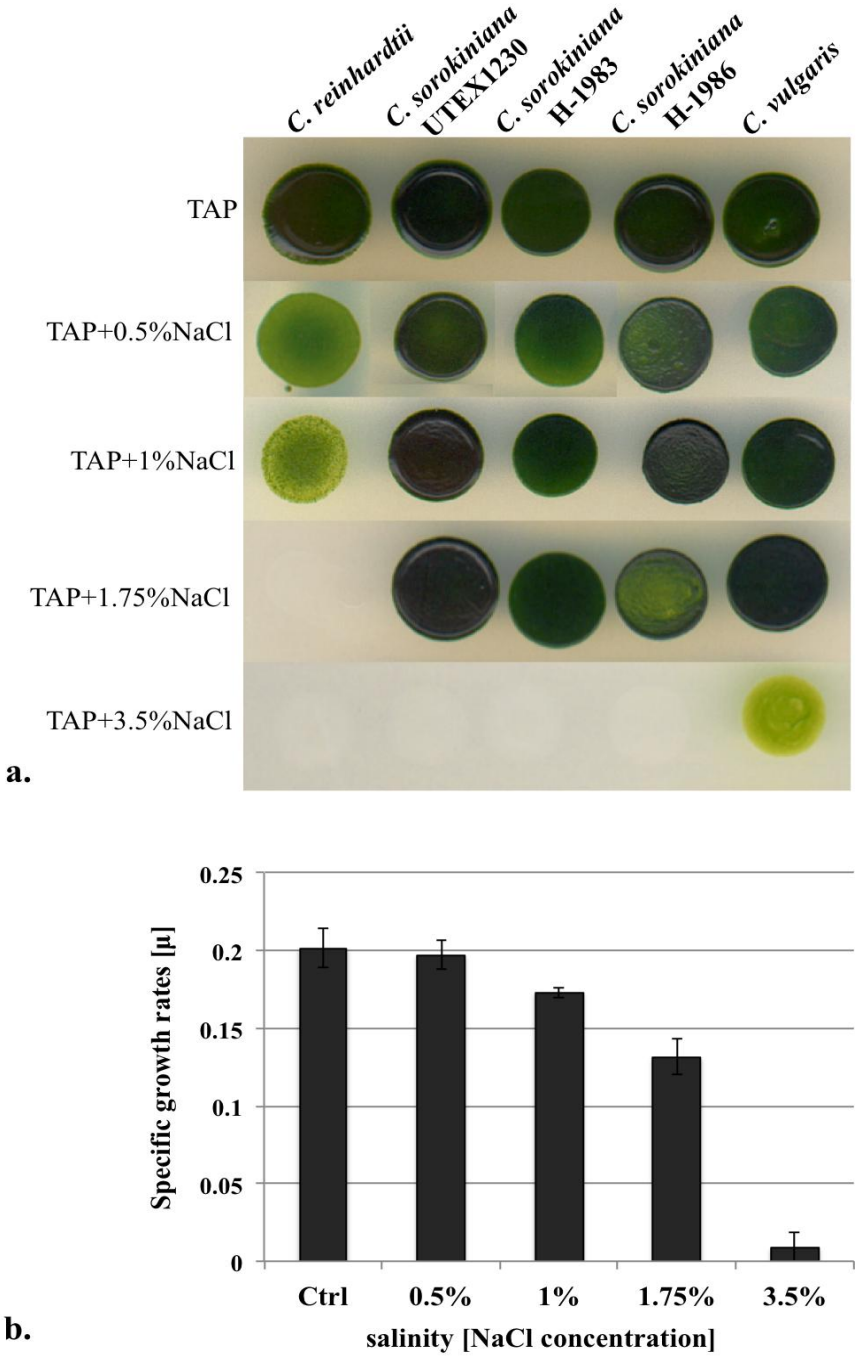


Figure 3-10 Effect of salinity on algal growth rates

a. Growth test of five microalgal strains (*C. reinhardtii, C. sorokiniana* strains: UTEX1230, H-1983, H-1986 and *C. vulgaris*) evaluating tolerance to four sodium chloride concentrations: 0.5%, 1%, 1.75% or 3.5%.

b. Specific growth rate of UTEX1230 strain calculated for each salinity condition. The experiment was performed in TAP medium with added NaCl stated concentrations. Error bars show ±SD (n=3).

However, the other types of salts also present in seawater such as magnesium were not included. Figure 3-10.a presents a spot test examining the influence of salinity on algal growth. Brackish water (up to 1% NaCl) appears not to influence growth of any of the *C. sorokiniana* strains, since there is no noticeable difference in growth rate in brackish water when compared to control (TAP). Conversely, growth of *C. reinhardtii* is substantially inhibited even at low NaCl concentrations. Figure 3-10.a indicate the salinity tolerance among *C. sorokiniana* strains is on comparable levels while *C. vulgaris* is characterised by the highest resistance to enhanced salinity. Specific growth rates were calculated for *C. sorokiniana* UTEX1230 in different salinity levels as presented in Figure 3-10.b. 0.5% NaCl does affect the growth of UTEX1230 whereas cultures containing 1% NaCl showed a decrease in growth of approximately 12% when compared to control, however this decline was not statistically significant. UTEX1230 cultured in a medium containing 1.75% NaCl displayed approximately 35% slower growth rate than control.

Using seawater for biofuel production from algae is not a new idea (Wagener, 1983, Wagener, 1981), yet it was limited to the maintenance of marine algae. The experiment presented in Figure 3-10 was designed to test the possibility of supplementing a freshwater culture of algae with seawater to a tolerable concentration that does not affect algal productivity. Increased salinity tolerance could also be useful when considering maintenance of algal cultures in open pond systems on land with increased salinity or in areas where fresh water is in limited supply. Additionally, increased salinity acts as a stressful environmental condition capable of inducing lipid production in green algae (Lei et al., 2012).

3.3.4 Growth of *C. sorokiniana* UTEX1230 at reduced oxygen levels

One of the major problems that occur during the maintenance of dense algal cultures in open pond systems is a very low oxygen concentration. This experiment was designed to test the influence of decreased oxygen level on *Chlorella sorokiniana* growth using a GasPak Pouch System to create a micro-anaerobic environment. The results are presented in Figure 3-11. This system decreases significantly the oxygen level down to 1%, which mimics the potential oxygen conditions that may occur in the lower layers of the bulk culture.

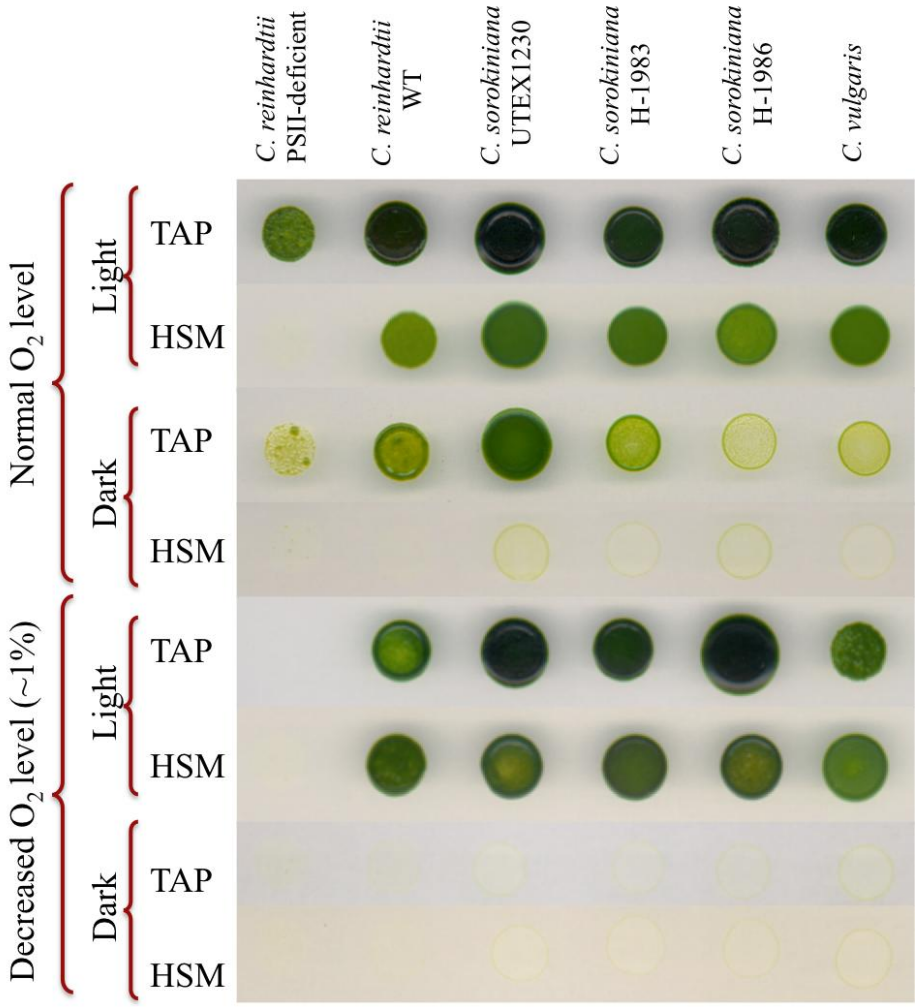


Figure 3-11 Algal growth at reduced oxygen level conditions

Spot tests of *C. vulgaris*, *C. reinhardtii* wild type (WT) and a photosynthetically-deficient strain and three *C. sorokiniana* strains: H-1986, H-1983, UTEX1230 on minimal medium (HSM) and with acetate as a carbon source (TAP). Cultures were grown in the dark and constant illumination at an average light intensity 50 µmol m^{-2} s^{-1} at normal and decreased oxygen levels.

Figure 3-11 presents the data obtained from a growth test of two strains of *Chlamydomonas reinhardtii* – wild type and photosynthetically deficient (PSII-), three strains of *Chlorella sorokiniana* (UTEX1230, H-1983 and H-1986) and *Chlorella vulgaris* in significantly reduced oxygen level conditions. In this experiment HSM and TAP plates were placed in either aerobic or anaerobic conditions, under continuous illumination or in dark. Lack of oxygen appears to have no influence on growth of all algal species maintained in light except for the PSII- strain where the mutant is dependent on O_2 for heterotrophic growth, since no difference between the control and plates in anaerobic chamber was observed. On the other hand – as expected – no growth was observed in decreased oxygen levels in the dark on both plates – with or without organic carbon. However, a decreased oxygen level is not the only issue in culturing algae on a bigger scale – increased O_2 in closed photobioreactors is extremely poisonous for algae. To avoid any of the extreme situations related to oxygen level culture in both systems – open ponds and photobioreactors are constantly mixed, which also enables the cultures to avoid algal sedimentation but also to allow the cells uniform access to light or nutrients.

3.3.5 Attempts to improve the photosynthetic efficiency of *C. sorokiniana* UTEX1230 by reducing the chlorophyll antenna size.

3.3.5.1 Background

The light-harvesting antenna consists of a structure of nucleus-encoded light harvesting proteins binding chlorophylls a and b and also xanthophylls (de Bianchi et al., 2011). Light harvesting during photosynthesis is the primary activity performed by the chlorophyll antenna that has the capacity to regulate the photon yield in order to avoid photodamage. This process is very efficient in moderate light. It is however, disturbed when organisms are exposed to stressful environmental conditions that can induce excitation of PSII by chlorophyll singlet excited state ($^1Chl^*$). This leads to the increased probability of forming triplet chlorophyll ($^3Chl^*$), which results in production of single oxygen ($^1O_2$) molecules (Melis, 1999, de Bianchi et al., 2011). $^1O_2$ accompany the formation of reactive oxygen species (ROS) therefore the photoprotective mechanisms within cells are associated with either detoxification of ROS or preventing their formation (de Bianchi et al., 2011, Demmig-Adams and Adams III, 1992). The latter process can be achieved through non-photochemical quenching (NPQ) that converts the excited states

into heat (Phillip et al., 1996) or by quenching $^{3}Chl^{*}$ or neutralising ROS (de Bianchi et al., 2011).

In the natural environment where light is frequently limiting and high competition for light occurs, algae contain large light harvesting chlorophyll antenna complexes (LHC) in connection with their photosystems. This enables maximal accumulation of solar energy (Polle et al., 2002, Kosourov et al., 2011). However in high light conditions this type of LHC structure results in the decreased efficiency of photosynthesis per chlorophyll molecule (Sukenik et al., 1987) as the uptake of the light photons is greater than the photosynthetic capability of the algal cell (Polle et al., 2002). The excess of absorbed photons not utilised in photosynthesis is dissipated in a form of heat or fluorescence and this wastage can reach up to 80% (Polle et al., 2002). Subsequently, in a bulk culture algal cells in the top layer absorb the excess of light energy that cannot be utilised and at the same time overshadow the cells from the lower layers, particularly in dense cultures and this leads to an overall low light utilisation efficiency (Sukenik et al., 1987, Kosourov et al., 2011).

In recent years there has been an increased interest in maximisation of the light delivery to all cells and captured light utilisation within algal cultures through the generation of mutants containing light harvesting chlorophyll antenna smaller that wild type (Kosourov et al., 2011, Polle et al., 2002, Stephenson et al., 2011, Wu et al., 2011a). It has been demonstrated that the truncated light harvesting antenna complex (*tla*) mutants of *Chlamydomonas reinhardtii* are characterised by an increased cell density compared to wild type (Polle et al., 2003) but also the productivity of these mutants is improved (Kosourov et al., 2011). Truncated chlorophyll antenna *tla* mutants of *C. reinhardtii* described in the literature were created through DNA insertion and chemical mutagenesis (Polle et al., 2002). Since the ultimate goal of this project is to obtain an improved *C. sorokiniana* strain for biofuel production in open pond systems, the aim of this experiment was to find an alternative way of obtaining isolates with truncated chlorophyll antenna. UV mutagenesis was therefore applied to increase the chance of mutation within the LHC region resulting in reduced antenna size.

The experiments presented in this part were performed in collaboration with the group of Professor Roberto Bassi of the University of Verona, Italy.

3.3.5.2 Chlorophyll fluorescence analysis of six isolated TAMs

As mentioned in the introduction of this chapter, light absorbed by photosynthetic cells can be utilised during photosynthesis, dissipated as heat or re-emitted in a form of fluorescence. All three processes occur in competition, where increase in efficiency of one process results in decrease of the remaining two. Therefore, by recording the yield of fluorescence we can obtain information on the efficiency of photochemistry and heat dissipation.

The isolated truncated chlorophyll antenna mutants are expected to exhibit a low chlorophyll fluorescence phenotype as the photon over-absorption and dissipation of the absorbed energy will be significantly lower compared to the wild type strain with standard-size antenna (Polle et al., 2003). In this project, UV mutagenesis was performed on wild type *Chlorella sorokiniana* UTEX1230 in order to induce a mutation affecting the chlorophyll antenna size. Approximately 3000 colonies were picked and screened for fluorescence emission. The results of the 6 most promising isolates are presented in Figure 3-12.

All six mutant colonies (TAM 1-6) were additionally screened for photoautotrophic, mixotrophic and heterotrophic growth ability and all TAM isolates grow very well in each of the trophic conditions demonstrating that they are not affected in photosynthesis itself, but they reveal a fluorescence yield lower than wild type. In TAM 2 and TAM4 the relative chlorophyll fluorescence was the lowest out of all 6 lines when cultured with or without an organic carbon source (Figure 3-12). Additionally, TAM2 shows the lowest chlorophyll fluorescence out of all tested strains cultured in the dark in TAP medium. A lower than WT chlorophyll fluorescence yield can be a reflection of stunted antenna, therefore further screening analysis of the six putative antenna mutants TAM 1-6 was accomplished.

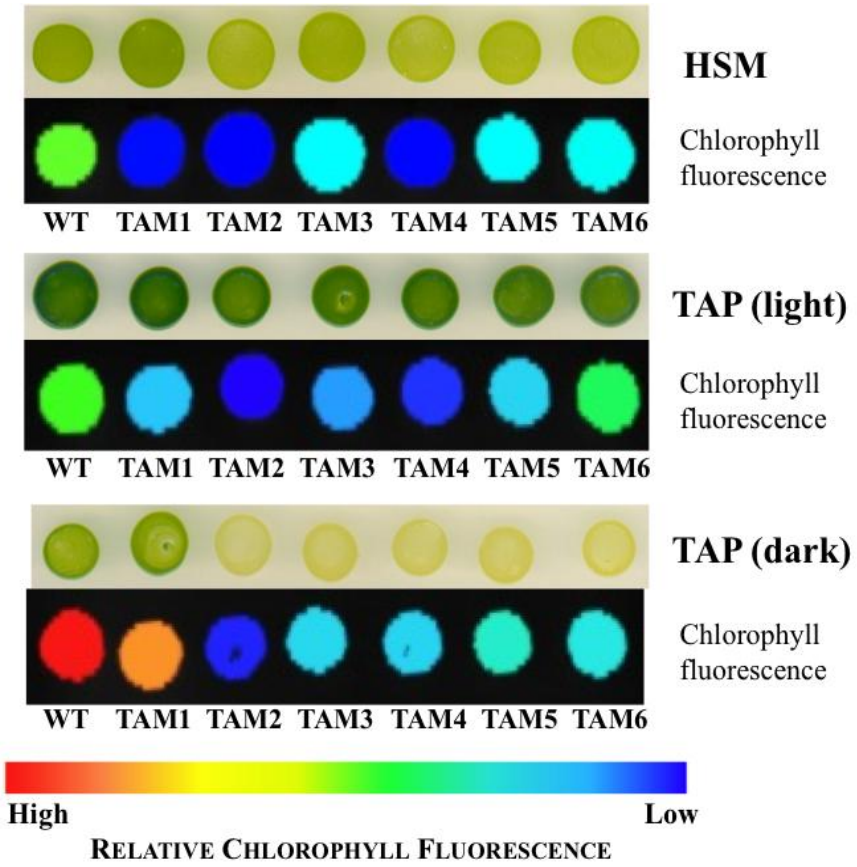


Figure 3-12 Growth and fluorescence analysis of six putative truncated chlorophyll antenna mutants (TAM) and wild type (WT) of *Chlorella sorokiniana* UTEX1230.

Spot tests of each TAM isolate and WT were maintained in light (HSM – phototrophically, TAP – mixotrophically) or in the dark (TAP – heterotrophically) for 7 days. Afterwards the chlorophyll fluorescence analysis was performed for each condition. Spot tests were performed in the dark or under constant illumination and average light intensity of 50 µmol m^{-2} s^{-1}.

Interestingly, the TAM 2 – TAM 6 isolates display a yellow-in-the-dark phenotype (Figure 3-12) that reverts back to green phenotype within hours after exposure to light (data not shown). One of the crucial steps in the chlorophyll biosynthesis pathway is reduction of protochlorophyllide to chlorophyllide and there are two mechanisms controlling this process in most photosynthetic eucaryotes (excluding gymnosperms): light-dependent process that is catalysed by NADPH: protochlorophyllide oxidoreductase (POR) or light-independent process that is required in chlorophyll formation in the dark (Rochaix et al., 2004). In *Chlamydomonas reinhardtii* products of chloroplast genes *chlL*, *chlN* or *chlB* and minimum seven out of *y-1* to *y-10* nuclear loci are required in the light-independent process of the protochlorophyllide reduction and it has been reported that mutations within any of those genes result in the yellow-in-the-dark phenotype (Cahoon and Timko, 2000, Zhang, 2007). This phenotype has been frequently present in wild population of *C. reinhardtii* due to the genetic instability of the *y-1* locus (Harris, 1989), and it is easily induced with various chemicals such as streptomycin (Gross and Dugger, 1969). UV treatment is a random type of mutagenesis hence the isolated TAM cell lines additionally expressing various levels of chlorophyll fluorescence can also demonstrate further morphological changes.

3.3.5.3 Chlorophyll and carotenoid content analysis of the isolated TAMs

The primary indicator of the functional chlorophyll composition is the chlorophyll a/b ratio as chlorophyll b is present solely in the LHC antenna structure whereas chlorophyll a is in the PSI and PSII reaction centres in addition to the antenna systems (Lichtenthaler and Buschmann, 2001). Additionally to chlorophylls, light harvesting complexes contain another type of pigment – carotenoids, that play an important role during the photooxidative stress and are particularly effective against ROS and $^1O_2$ toxicity (Ramel et al., 2012). Therefore, the total concentration of carotenoids and chlorophyll a and b was measured for the six isolated putative chlorophyll antenna mutants (TAM 1-6) and the wild type of *C. sorokiniana* UTEX1230. The correlation between total chlorophylls and carotenoids and the chlorophyll a/b ratio is presented in Figure 3-13.

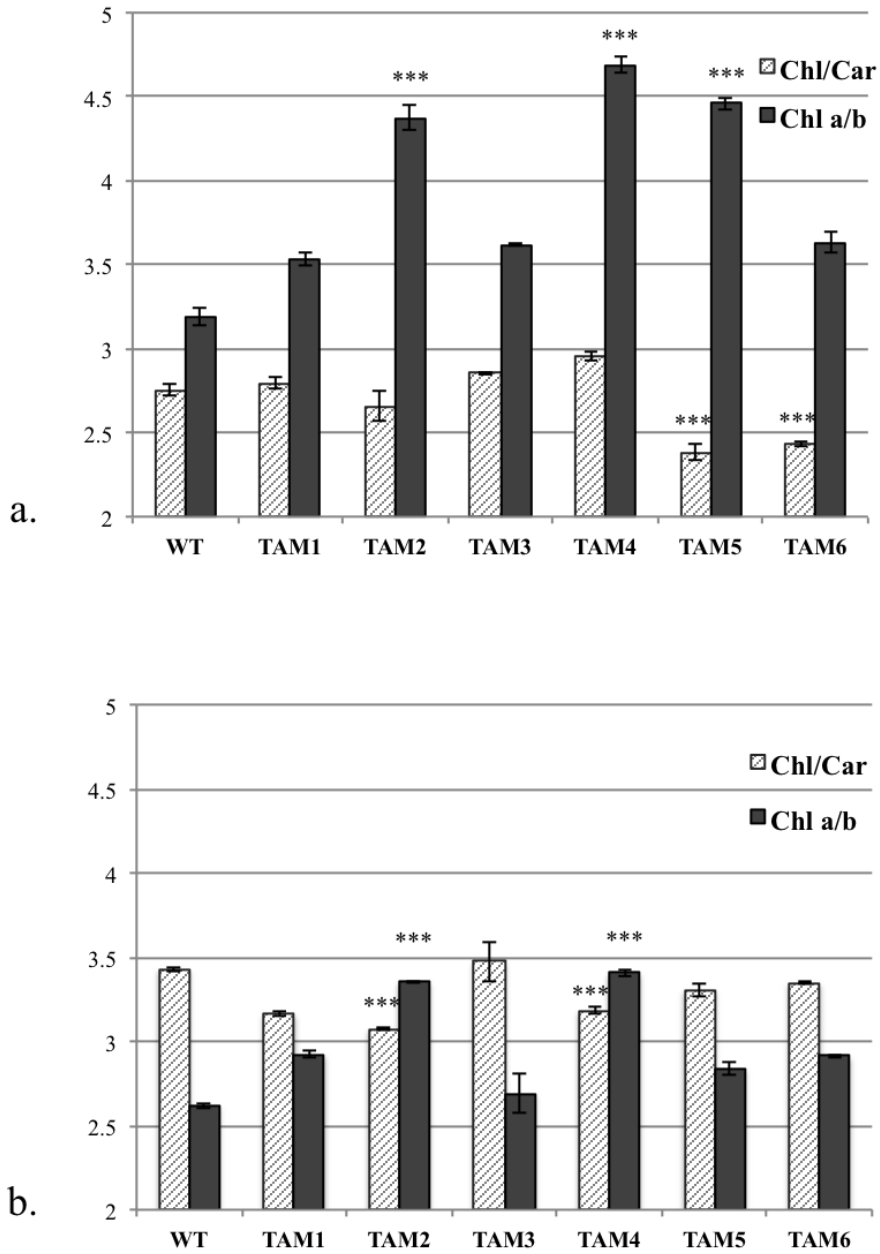


Figure 3-13 Chlorophyll and carotenoid ratio analysis of the putative antenna mutants of *C. sorokiniana* UTEX1230 (TAM 1-6) and wild type (WT).
Chlorophyll a/b ratio (Chl a/b) and total chlorophyll to carotenoid content (Chl/Car) of the TAM mutants calculated for the isolated thylakoids from cultures maintained in rich medium (a), and (b) in minimal medium. Error bars show ±SD (n=3). Significantly different values calculated using Student's *t* test of the TAM mutants compared to WT where extremely significant values were based on p>0.001 and were marked with three asterisks.

The association between pigments within the photosystems was measured for two growth conditions – mixo- and phototrophic in order to test how these influence the relationship between chlorophylls and carotenoids. The WT chlorophyll a/b ratio is elevated in the cultures grown mixotrophically compared to the phototrophic cultures suggesting an adaptive process in which light-harvesting is down-regulated when an organic carbon source is available. TAM2 and TAM4 have a significantly higher chlorophyll a/b ratio when maintained phototrophically and mixotrophically compared to wild type that indicates the possibility that both isolates have a reduced antenna size. This result confirms the preliminary data presented in Figure 3-12. Interestingly the chlorophyll a/b ratio was also significantly elevated in TAM5 cell line compared to wild type however, this was observed only in mixotrophic conditions.

Chlorophyll a/b ratio is linked to photoadaptation in plants (Apel and Kloppstech, 1978) and green algae (Hess and Tolbert, 1967) where light intensity is inversely proportional to the chlorophyll a/b ratio. It has been observed that under limited light conditions the chlorophyll antenna size increases in order to capture the maximal amount of photons and this is seen as a relative increase of chlorophyll b to a (Dale and Causton, 1992). Additional environmental conditions other than light are reported to be associated with changes in chlorophylls proportion. For instance, decreasing temperature leads to a decrease in the chlorophyll a/b ratio in plants (Riesselmann and Piechulla, 1990) but also nutrient deprivation such as iron deficiency results in the reduced chlorophyll a/b ratio (Varsano et al., 2006).

While the chlorophyll a/b ratio increased, the total chlorophyll to carotenoid ratio greatly decreased in the mixotrophically maintained cultures with added acetate compared to the phototrophic cultures (Figure 3-13). This phenomenon is most likely related to the abundance of the photosynthetic apparatus and since organic carbon is present in the environment there is no need for the cells to activate the maximum number of photosystems, hence the total chlorophyll level is reduced (Figure 3-14). In mixotrophic conditions TAM5 and TAM6 have significantly lower total chlorophyll to carotenoid ratio however, in phototrophic conditions this ratio was significantly reduced in TAM2 and TAM4 isolates compared to wild type.

Carotenoids in plant and algal cells play several roles mainly in light absorption and photoprotection but also in energy transfer within the light harvesting complexes (Liu et al., 2004). Figure 3-13 presents the total chlorophyll to carotenoid level in rich (a) and

minimal (b) medium under constant illumination. When grown mixotrophically only TAM5 and TAM6 have this ratio significantly decreased, which indicates that in these isolates the carotenoid level is increased when compared to the wild type. Nevertheless, in phototrophic conditions the total chlorophyll to carotenoid level is significantly reduced only in TAM2 and TAM4. The carotene profile of phototrophic cultures was measured in all TAM isolates and wild type, and the presented data in Figure 3-15 will be discussed later.

The total chlorophyll yield per cell is an additional indicator for the antenna size as up to 50% of the total chlorophyll is linked to the light harvesting complexes (Kan and Thornber, 1976). Figure 3-14 presents the total chlorophyll concentration per cell of the strains TAM 1-6 and WT cultured in minimal (Figure 3-14.a) and rich (Figure 3-14.b) medium maintained under constant illumination. When cultured mixotrophically, all TAM isolates showed a decreased chlorophyll level per cell and this difference was significant for TAM 2-6: from approximately 2 pg/cell to 0.075 pg/cell for TAM2 and TAM3, 0.1 pg/cell for TAM4, 0.5 pg/cell for TAM5 and 0.6 pg/cell for TAM6. In phototrophic cultures the total chlorophyll level was significantly lower for TAM2 and TAM4 only and this value was decreased from 0.5 pg/cell to 0.25 pg/cell and 0.33 pg/cell, respectively. The total chlorophyll concentration changed depending on the type of growth (Figure 3-14), as the average chlorophyll concentration for WT was 0.5 pg/cell in phototrophic cultures, whereas it was substantially decreased to 0.2 pg/cell in mixotrophic cultures. This result confirms the earlier finding where it has been demonstrated that in *Chlorella protothecoides* heterotrophic growth leads to decreased light harvesting complex size (Sasidharan and Gnanam, 1990).

Previous research established that environmental conditions are also able to affect chlorophyll yield in cells. For instance, chlorophyll levels in algal cells decline substantially under the nitrogen deprivation (Hortensteiner et al., 2000). Furthermore, decreasing temperature has been reported to have a positive influence on total chlorophyll concentration in plants (Piechulla and Riesselmann, 1990). Additionally, changing illumination has an impact on this pigment as increasing light intensity leads to chlorophyll degradation (Lindahl et al., 1995). Also, photoinhibition results in a lower chlorophyll level compared to normal conditions (Harper et al., 2004).

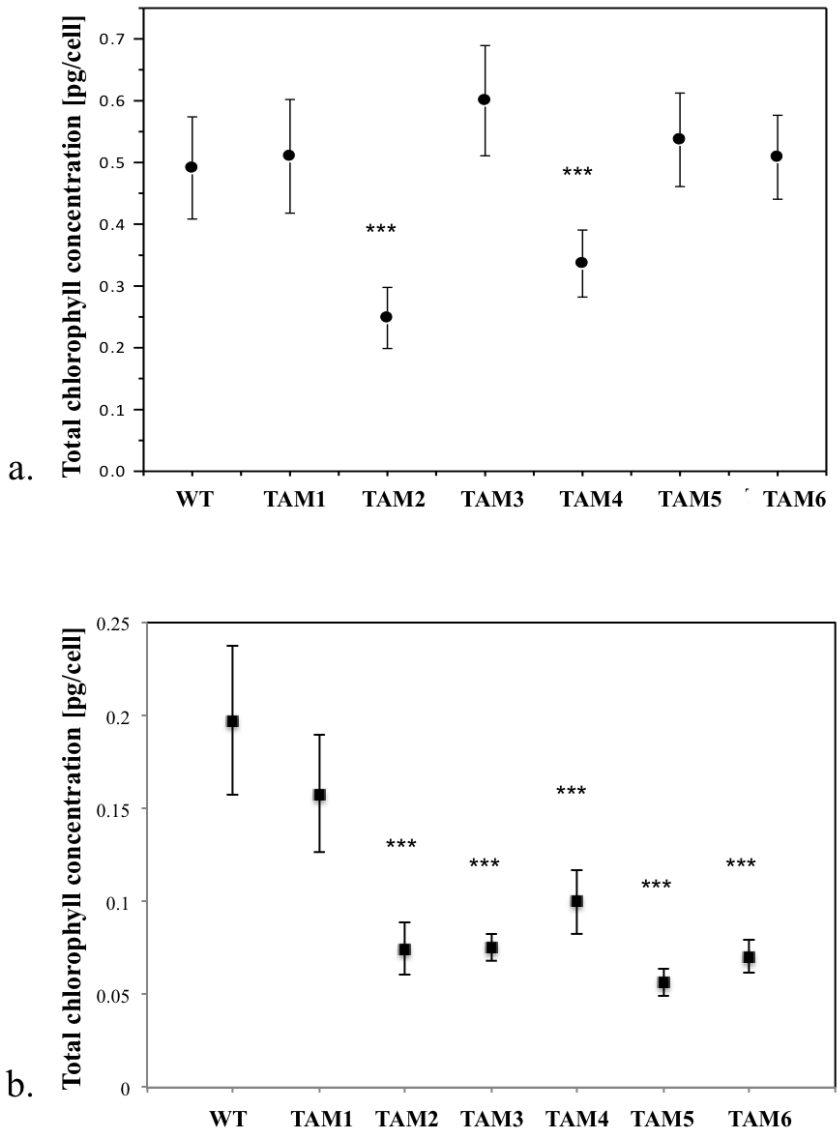


Figure 3-14 Total chlorophyll content of the truncated chlorophyll antenna mutants (TAM 1-6).

Total chlorophyll content was measured on thylakoids isolated from the TAM mutants and wild type (WT) strain of *Chlorella sorokiniana* UTEX1230 cultured in minimal (a), and rich medium (b). Symbols and error bars show means ±SD (n>3). Significantly different values calculated using Student's *t* test of the TAM mutants compared to WT were based on p>0.001 and were marked with three asterisks. Cultures were maintained under constant illumination and average light intensity of 100 µmol m^{-2} s^{-1}.

The mechanism of decreasing chlorophyll yield in extreme environmental conditions is a form of photoadaptation, since the smaller total yield of the pigments within the light harvesting complex results in a smaller number of photons captured, and decreased probability of induction of photoinhibition. It is known antenna size and carotenoid composition in plants changes with light acclimation therefore, the aim was to test how decreased antenna size affects carotenoid layout. The major types of carotenoids in photosynthetic cells are either derivatives of α-carotene (for instance lutein that is the most abundant type of carotenoids) or β-carotene (for instance zeaxanthin and xanthophyll cycle pool) (Demmig-Adams and Adams III, 2002). Lutein plays an important role in the LHCII formation (Hirschberg, 2001), whereas zeaxanthin and the xanthophyll cycle are associated with thermal energy dissipation (Demmig-Adams and Adams III, 2002). During the xanthophyll cycle, zeaxanthin is converted to violaxanthin when light needs to be used efficiently during photosynthesis, whereas during excess light conditions this process is reversed (Demmig-Adams et al., 1996). Carotenoid composition in thylakoids is variable and is subjected to constant changes depending on the environmental conditions such as light intensity or temperature conditions (Adams III et al., 1993) however, the question I wanted to address was how reduction in size of chlorophyll antenna affects the carotenoid profile of *C. sorokiniana* UTEX1230. The carotenoid composition of TAM isolates was therefore measured and compared to WT and the results are presented in Figure 3-15.

The average concentration of neoxanthin in WT but also for TAM3 and TAM5 was 6.6 mol/100 mol chlorophylls. This amount somewhat decreased in TAM1 and TAM6 and was the lowest in TAM2 and TAM4 down to just under 6 mol/100 mol chlorophylls (Figure 3-15.a). Neoxanthin levels in plants represents approximately 14% of total carotenoids in low light conditions although it is not essential in LHC formation and its function in photooxidative stress can easily be replaced by violaxanthin (Dall'Osto et al., 2007). Conversely, violaxanthin concentration in the isolates that demonstrated decreased levels of neoxanthin, TAM2 and TAM4, was increased approximately 4- and 4.5-fold– respectively when compared to wild type, as seen in Figure 3-15.b. TAM5 and TAM6 showed approximately 2-fold rise in violaxanthin levels, whereas for TAM1 and TAM3 violaxanthin levels increased approximately 25%.

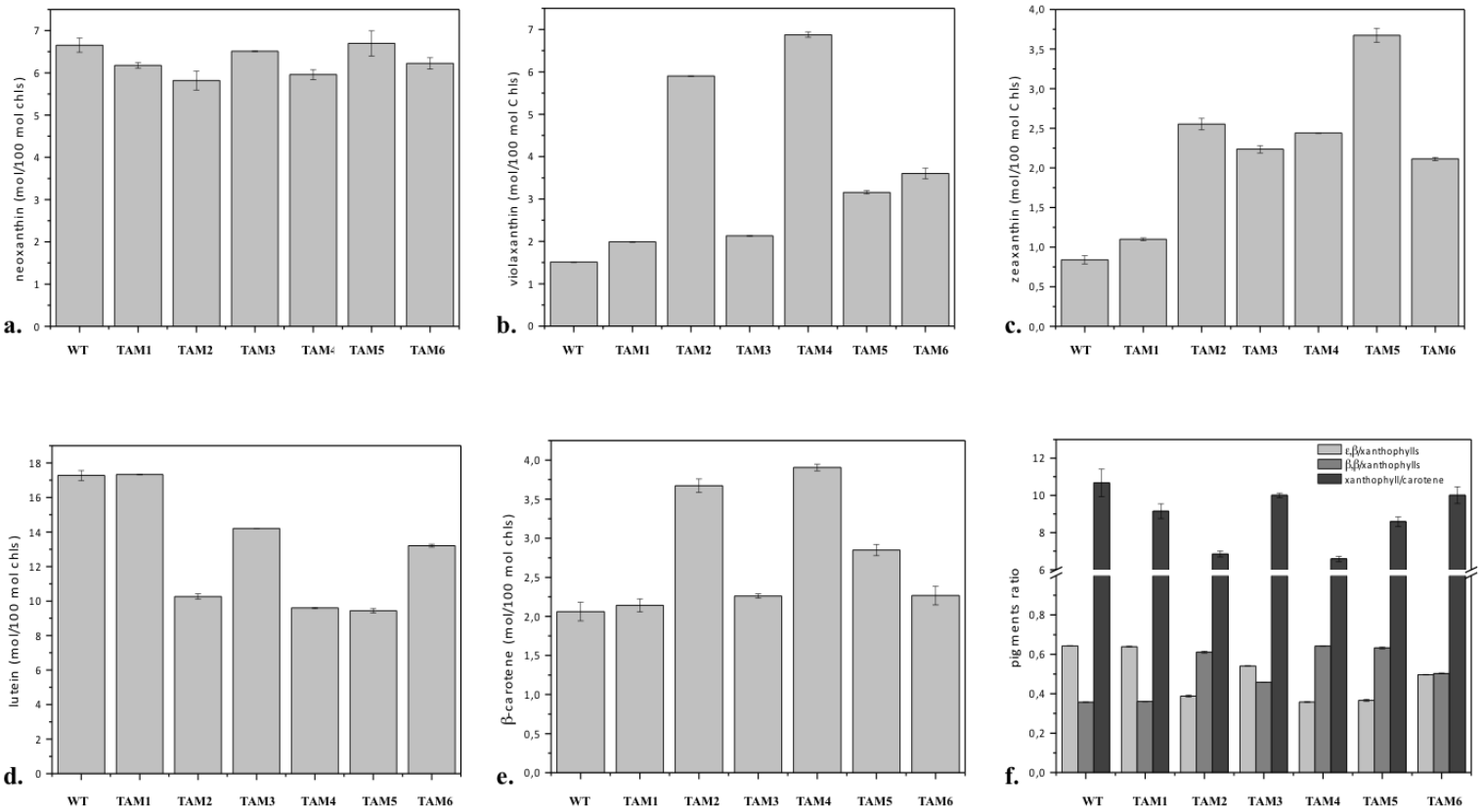


Figure 3-15 Pigment analysis of the chlorophyll antenna mutants (TAM 1-6) and wild type (WT) of *C. sorokinianana* UTEX1230 cultured in minimal medium under constant illumination and average light intensity of 100 μmol m^{-2} s^{-1}. Columns and error bars show means ±SD (n=3) .

a-e. Concentration of neoxanthin, violaxanthin, zeaxanthin, lutein and β carotene [mol] respectively per 100 mol of total chlorophyll concentration.
f. Pigment ratios of xanthophylls:

ε,β xanthophylls [α-carotene+lutein]/total xanthophylls, **β,β xanthophylls** [β-carotene + zeaxanthin]/total xanthophylls and **total xanthophylls/β carotene**

As discussed earlier, non-dissipating violaxanthin plays a role in the xanthophyll cycle and is rapidly converted to the dissipating zeaxanthin during the light stress and both pigments are ordinarily in balance yet, surprisingly, zeaxanthin levels are also substantially increased in almost all TAM isolates compared to WT, except for TAM1. In TAM2 and TAM3, a 3-fold increase in zeaxanthin concentration was observed, in TAM3 and TAM6 zeaxanthin was increased approximately 2.5-fold, whereas in TAM5 the level of this pigment raised over 4.5 fold when compared to WT. Since the protective response of zeaxanthin to increased light is always faster than any other pigment (Carnicas et al., 1999), the results suggest TAM isolates are better equipped than wild type for rapid regulation to changing light conditions.

Both pigments of the xanthophyll cycle are accountable for physiological control of the light harvesting system where violaxanthin is responsible for disaggregation, whereas zeaxanthin for aggregation of proteins within LHC (Ruban et al., 1997). Zeaxanthin level tends to be directly proportional to the intensity of light and in lichens is higher in the shaded isolates than those grown in direct sun (Adams III et al., 1993). Furthermore, in plants increased zeaxanthin levels are associated with areas of very low illumination, where photosynthesis and photorespiration are inhibited (Demmig-Adams et al., 1989). Moreover, increased zeaxanthin levels results in enhanced heat tolerance in plants (Davison et al., 2002). As an antioxidant, zeaxanthin plays a dual role as this pigment, additionally to down-regulating the chlorophyll singlet excited state, increases the efficiency of quenching the chlorophyll triplets (Betterle et al., 2010).

Lutein concentration is not altered only in TAM1 when compared to wild type, whereas in TAM2, TAM4 and TAM5 it is decreased almost 2-fold (Figure 3-15.d). Lutein accounts for almost half of the total carotenoids in shaded plants and it decreases to just over one-third in plants exposed to high irradiance (Demmig-Adams and Adams III, 1992). Similarly to zeaxanthin, lutein plays a role in excess energy dissipation where excess energy accepted from chlorophyll is scattered as heat (Josue and Frank, 2002).

On the contrary to lutein concentration, β-carotene was increased almost 2-fold in TAM2 and TAM4 compared to WT, in TAM5 this rise was of 25%, whereas in TAM1, TAM3 and TAM6 β-carotene levels did not change at all compared to wild type (Figure 3-15.e). β-carotene represents approximately 10% of the total carotenoids of the shaded plants and its concentration increases to over 20% in plants exposed to high irradiance (Demmig-Adams and Adams III, 1992). In addition, β-carotene in plants is linked to accumulating

excess singlet oxygen however this process happens mostly in PSII reaction centres rather than chlorophyll antenna (Ramel et al., 2012). These results suggest that in TAM2 and TAM4 the protective structure of the light harvesting complexes and reaction centres to high irradiance is more advanced than in wild type and the rest of the TAM isolates.

The total β,β xanthophylls carotenoid pool size mainly represented by β-carotene and zeaxanthin is reported to be increased with light (Bailey et al., 2004) and this is observed in TAM2, TAM4 and TAM5 compared to WT, whereas the total ε,β xanthophyll level diminished (Figure 3-15.f), which confirms the findings that lutein content declines with increased irradiance (Bailey et al., 2004, Demmig-Adams and Adams III, 1992). It appears the carotenoid composition of the most promising antenna mutants TAM2, TAM4, but also TAM5 to a certain extent, resemble the phenotype of the cell lines adapted to high light conditions whereas WT and the remaining TAMs show the carotenoid profile of strains not affected by high illumination.

3.3.5.4 Analysis of the chlorophyll antenna size in TAMs and wild type (WT)

In order to further evaluate the biochemical size of the chlorophyll antenna, western analysis was performed. Figure 3-16.b presents the results of immunoblotting the PSII components – LHCII that is a marker of the external antenna protein complex and CP43 that is the marker of the inner antenna. The biochemical size of the chlorophyll antenna was calculated based on the semi-quantitative LHCII/CP43 ratio of the immunotitration of thylakoid proteins normalised against the chlorophyll content (Figure 3-16.a). The chlorophyll antennae were significantly reduced in TAM2 and TAM4, which correlates with the experiments reported earlier in this chapter, yet TAM5 and TAM6 surprisingly showed reduction in the chlorophyll antenna, whereas the LHC of TAM1 and TAM3 – as expected – is of WT dimensions.

Acclimation to changing light conditions alters the accumulation of the light-harvesting proteins of the PSII whereas the core of PSI and its peripheral antenna remain fixed. The correlation between PSI and PSII cores was calculated (Figure 3-16.c) based on the PsaA/CP43 ratio of the immunotitration of thylakoid proteins (Figure 3-16.d), where PsaA is one of the PSI core subunits, whereas CP43 co-forms the PSII core.

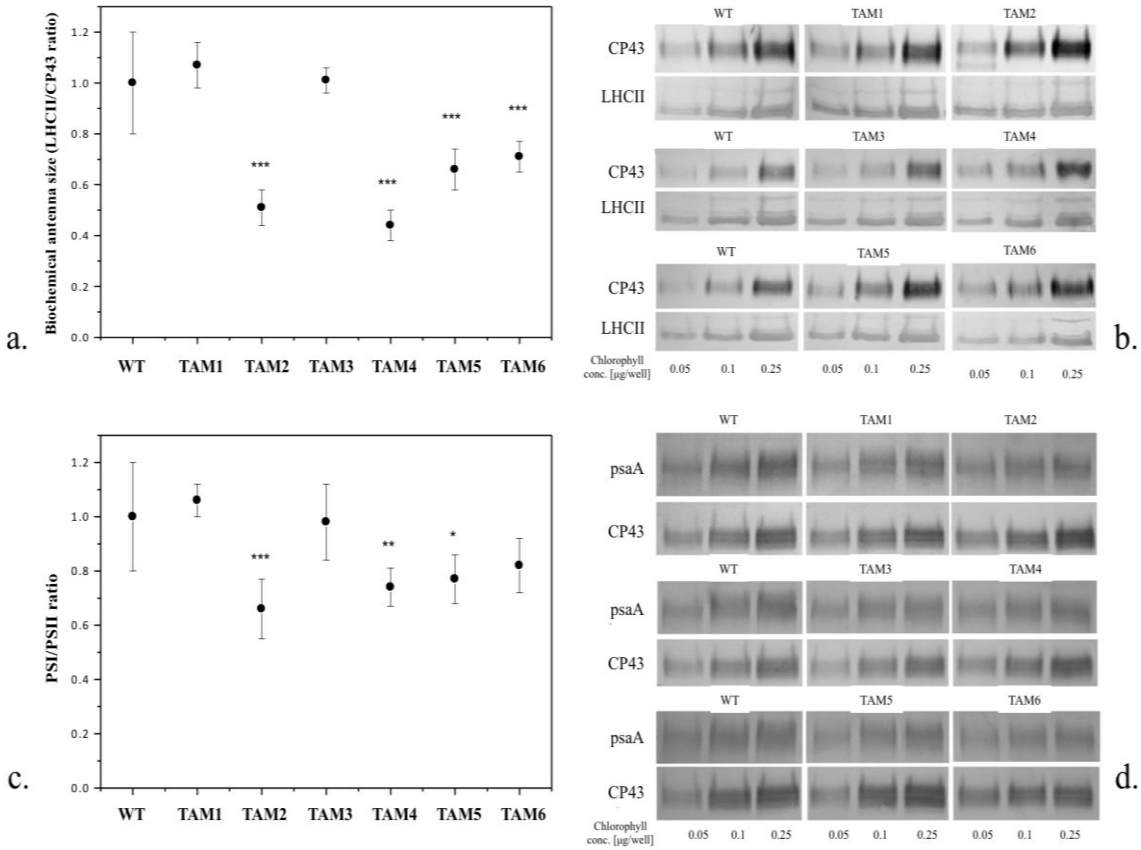


Figure 3-16 Western blot analysis of thylakoid proteins isolated from the truncated antenna mutants of *C. sorokiniana* UTEX1230 (TAM 1-6) and wild type (WT).

a. Biochemical antenna size calculated based on the LHCII/CP43 ratio obtained from semi-quantitative analysis.
b. Immunoblotting used for the quantification of photosystem II (PSII) subunits in the TAM isolates and WT thylakoids. In order to prevent saturation each sample was loaded in three dilutions
c. PSI to PSII ratio calculated from PsaA/CP43 ratio
d. Immunoblotting used for the quantification of PSI to PSII ratio of the thylakoids of TAM isolates and WT.

Cultures were maintained in minimal medium under constant illumination and average light intensity of 100 µmol m^{-2} s^{-1}.

Thylakoids corresponding to 0.05, 0.10 and 0.25 µg of chlorophylls were loaded to each well. Symbols and error bars show means ±SD (n=3). Significantly different values calculated using Student's *t* test of the TAM mutants compared to WT; extremely significant values were based on $p>0.001$ and were marked with three asterisks whereas very significant values were based on $0.05<p>0.001$ and were marked with two asterisks.

The calculated PSI/PSII ratio of the TAM isolates is presented in Figure 3-16.c. TAM2 displays a significant decrease in the PSI/PSII ratio, with reference to WT, whereas this difference is reduced in TAM4 and even more in TAM5. PSI/PSII ratio refers to the relative proportion of absorbed light energy available to both photosystems. This ratio in plants is inclined to decrease with increased light intensity to a point where photoinhibition starts dominating. However, after reaching a certain threshold it starts gradually increasing (Walters and Horton, 1991). One of the possible explanations for the PSI/PSII ratio decrease with increasing irradiance is to avoid the superoxide formation at the PSI donor site (Bonente et al., 2012).

The results of the western blot presented in Figure 3-16.a and Figure 3-16.b and discussed above allowed the protein ratios of the PSII core versus LHCII that establish the biochemical size of the chlorophyll antenna to be calculated, yet this does not reveal the information on its function. The biochemical data do not provide the full picture on LHCII. It has been shown that different families of protein within the internal or external antenna bind various numbers of pigment molecules (Mozzo et al., 2010) or do not bind them at all (Bonente et al., 2008).

The fluorescence emission kinetics of PSII reaction centre oxidation was therefore measured in the presence of an inhibitor of electron transport – DCMU, and the calculated rate of fluorescence is proportional to the functional antenna size of PSII. Figure 3-17.a presents the example of the fluorescence emission kinetics of three dark-adapted DCMU-treated strains: WT versus either TAM1 or TAM4. The calculated functional PSII antenna size is presented in Figure 3-17.b. As it might be expected, TAM2 and TAM4 show significantly reduced functional PSII antenna compared to WT. However, the functional antenna is also significantly reduced in TAM5 and TAM6 as well as in TAM1 and TAM3.

In algae, during the photoacclimation, the functional antenna size does not change, albeit the biochemical size of the LHCII decreases with increased irradiation (Bonente et al., 2012). Therefore, the reduction in LHCII activity presented in Figure 3-17 is a consequence of permanent structural changes. Subsequently, the next experiment was designed to investigate how the LHCII decline influences the mechanisms of photoprotection in algae.

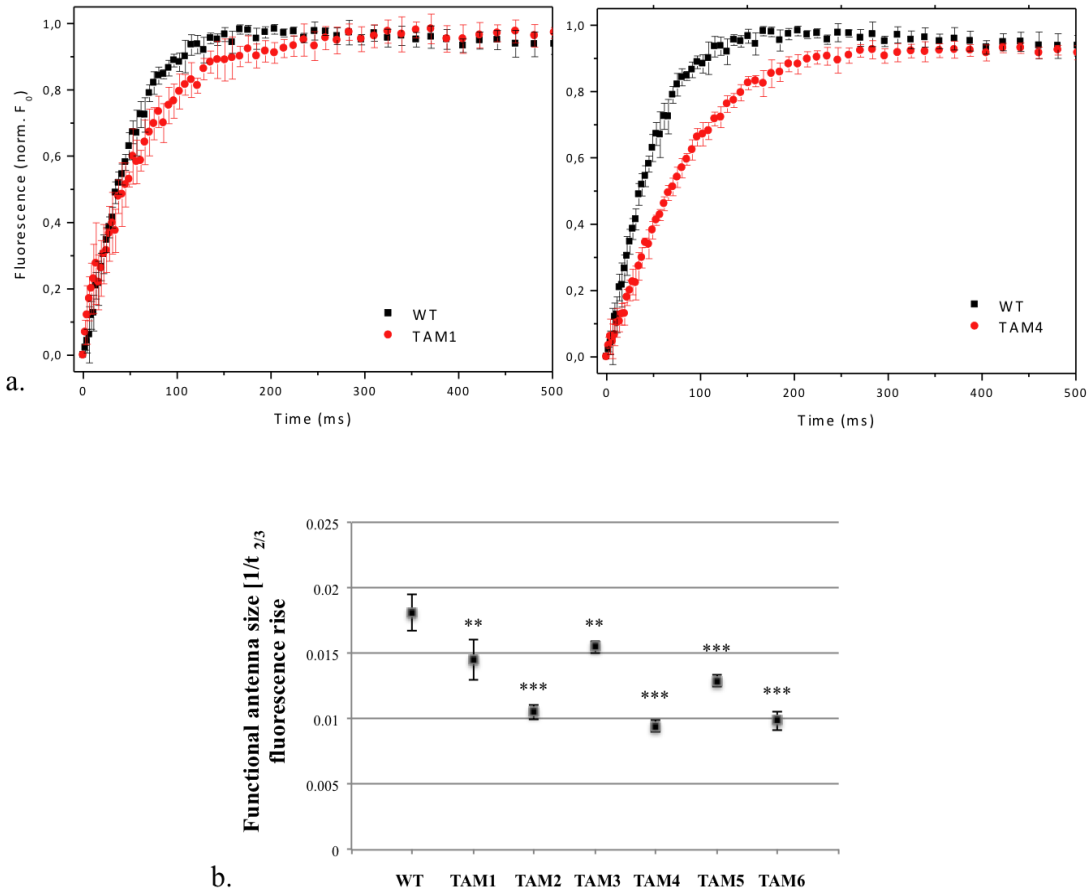


Figure 3-17 Functional antenna size of photosystem II.

a. Fluorescence emission kinetics of PSII from the dark-adapted truncated antenna mutant cells of *C. sorokiniana* UTEX1230 (TAM1 and TAM4) and wild type (WT) maintained in minimal medium under constant illumination and average light intensity 100 µmol m^{-2} s^{-1}. Symbols and error bars show means ±SD (n>3).

b. PSII fluorescence induction analysis in DCMU of the dark adapted cells – measurement of the functional antenna sizes of each TAM strain and WT that was calculated based on the fact that the time required to reach 2/3 of the maximum fluorescence is inversely proportional to PSII antenna Symbols and error bars show means ±SD (n>3), significantly different values calculated using Student's *t* test of the TAM mutants compared to WT; extremely significant values were based on p>0.001 and were marked with three asterisks whereas very significant values were based on 0.05<p>0.001 and were marked with two asterisks.

As mentioned earlier, one of the major photoprotection mechanisms in photosynthetic organisms is a mechanism of non-photochemical quenching (NPQ) that is associated with the PSII core (Finazzi et al., 2004), but also with the antenna system (Havaux et al., 2007). NPQ is associated with an ability to adapt to changing light conditions of photosynthetic organisms that is activated to prevent the ROS formation. When *C. sorokiniana* UTEX1230 was cultured in either rich or minimal medium in low light of average intensity 30 µmol m^{-2} s^{-1}, no NPQ induction was observed (data not shown). It has been discussed that various photosynthetic species grown in diverse light conditions show different NPQ responses, indicating that readily reversible NPQ is a reflection of an efficient photoadaptation process and certain strains take several generations maintained in high light to induce it (Bailey et al., 2004, Park et al., 1996). NPQ activity of UTEX1230 WT and TAM2, TAM4 and TAM5 isolates was measured and the results are presented in Figure 3-18. All tested strains were cultured in minimal medium at two light intensities, high light (500 µmol m^{-2} s^{-1}) and low light (50 µmol m^{-2} s^{-1}), for five days. The wild type strain acclimated to low light showed a rapid increase of NPQ up to 0.7 within the first minute after exposure to a saturating light intensity of 1200 µmol m^{-2} s^{-1}, whereas this increase was substantially lower in all TAM isolates (Figure 3-18.a and b) reaching between 0.1 and 0.15. Throughout the illumination process, NPQ remained at the same level in TAM1, TAM2, TAM3 and TAM4 isolates, or somewhat increased to 0.2 in TAM5 and TAM6. Recovery in the dark after 10-minute illumination was considerably longer for wild type than any of the TAM strains. The NPQ activities were entirely different for the tested strains acclimated to the high light (Figure 3-18.c and d). WT showed an increase of NPQ up to 0.425 within the first minute followed by a steady increase up to 0.6 within the next 10 minutes of illumination. TAM1 and TAM5 showed an NPQ rise up to approximately 0.35. For TAM5, NPQ remained on the same level and steadily decreased in TAM1 within next 10 minutes of illumination. The initial NPQ rise in TAM2, TAM3 and TAM4 was approximately twice as high as for WT. NPQ gradually decreased in WT during 10-minute illumination, whereas for TAM2, TAM3 and TAM4, NPQ declined to the WT level. Conversely, the recovery in the dark after 10-minute illumination was substantially faster and longer of all TAM mutants compared to WT, with TAM2, TAM3 and TAM4 being the most efficient.

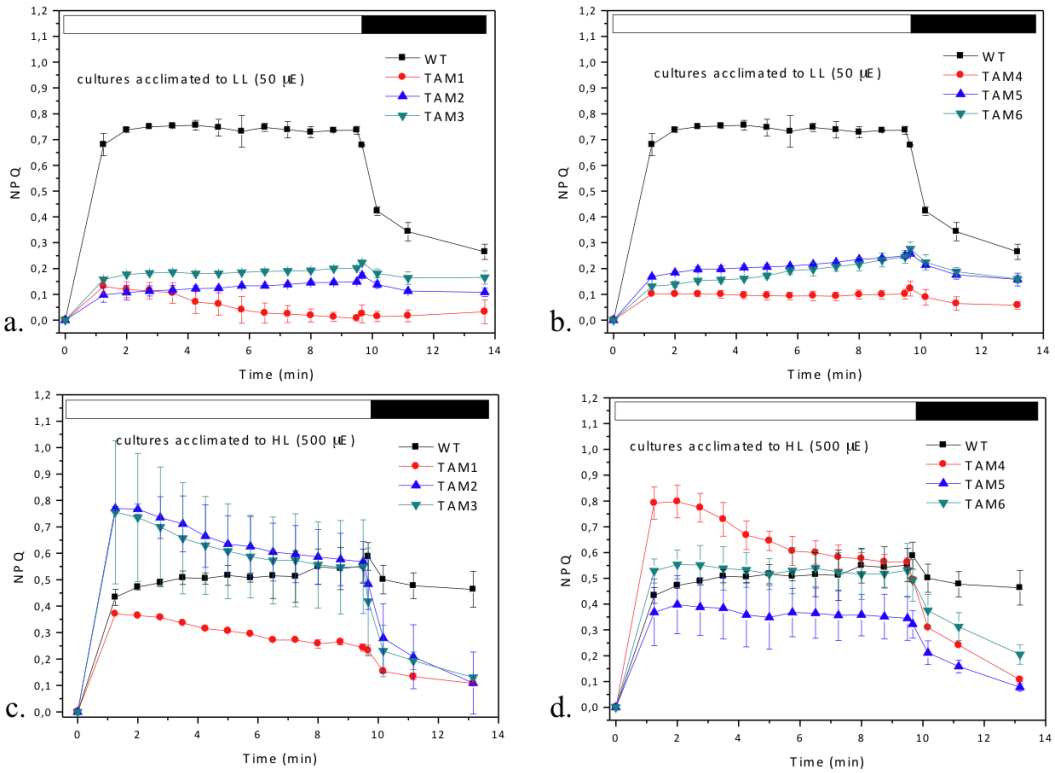


Figure 3-18 NPQ kinetics of truncated antenna mutants of *C. sorokiniana* UTEX1230 and wild type (WT) maintained in minimal medium.

During each measurement the strains were illuminated with saturating light pulse (approximate light intensity 1200 µmol photons m^{-2} s^{-1}), 24^0C (white bar) followed by 4 minutes of dark relaxation (black bar). The strains were acclimated to two light conditions: low light of 50 µmol m^{-2} s^{-1} (a, b) and high light of 500 µmol m^{-2} s^{-1} (c, d) during 2 fresh re-inoculations. Symbols and error bars show means ±SD (n>3)

Light response curves of the WT strain indicate that photoprotection mechanisms of cultures adapted to low light are substantially more efficient when compared to cultures adapted to high light. In cultures adapted to high light, the cells experience photodamage, as NPQ is non-reversible. Conversely, TAM2, TAM3 and TAM4 show a more efficient response to high light than WT, but only if they were previously photoadapted to excess light conditions. Plants adapted to high irradiation are characterised by rapidly reversible NPQ and higher maximum levels of NPQ is linked to high-light grown plants (Park et al., 1996). The actual mechanism of this correlation is unknown, but is most likely associated with the structure and composition of thylakoids (Bailey et al., 2004) whereas the rapid dark relaxation of NPQ has been identified with interconversion of violaxanthin and zeaxanthin (Walters and Horton, 1991). These results suggest that in order to activate photoprotection mechanisms in TAM2, TAM3 and TAM4, the cultures are required to be adapted to high light. On the contrary, the photoprotection mechanisms in WT are available only when cultures are maintained in low light and long exposure to high light causes photodamage.

3.3.5.5 Growth analysis of TAM isolates and wild type

As it was seen in Figure 3-17, all TAM isolates are characterised by reduced functional antenna activity, yet only TAM2, TAM4, TAM5 and TAM6 show the reduction in size within the protein components of LHC (Figure 3-16) and only TAM2 and TAM4 present the most promising pigment profile (Figure 3-13, Figure 3-14, and Figure 3-15). Therefore, the analysis of the growth profile of TAM2, TAM4 and TAM5 at various irradiances was performed and the results were compared to WT.

The cultures in minimal medium were maintained under constant illumination at the range of light intensity between 5 and 100 μmol m^{-2} s^{-1} (Figure 3-19.a), where stacked flasks were illuminated from one side. The specific growth rates (μ) were calculated at the log phase of the algal growth and the results are presented in Figure 3-19.b. The specific growth rates for TAM2 and TAM4 were almost identical and were higher than WT, particularly when the light intensity was low. The μ values for TAM5 were lower compared to WT. Moreover, the specific growth rates for TAM2 and TAM4 were the same in all light conditions, whereas the specific growth rates for WT declined with decreasing light intensity.

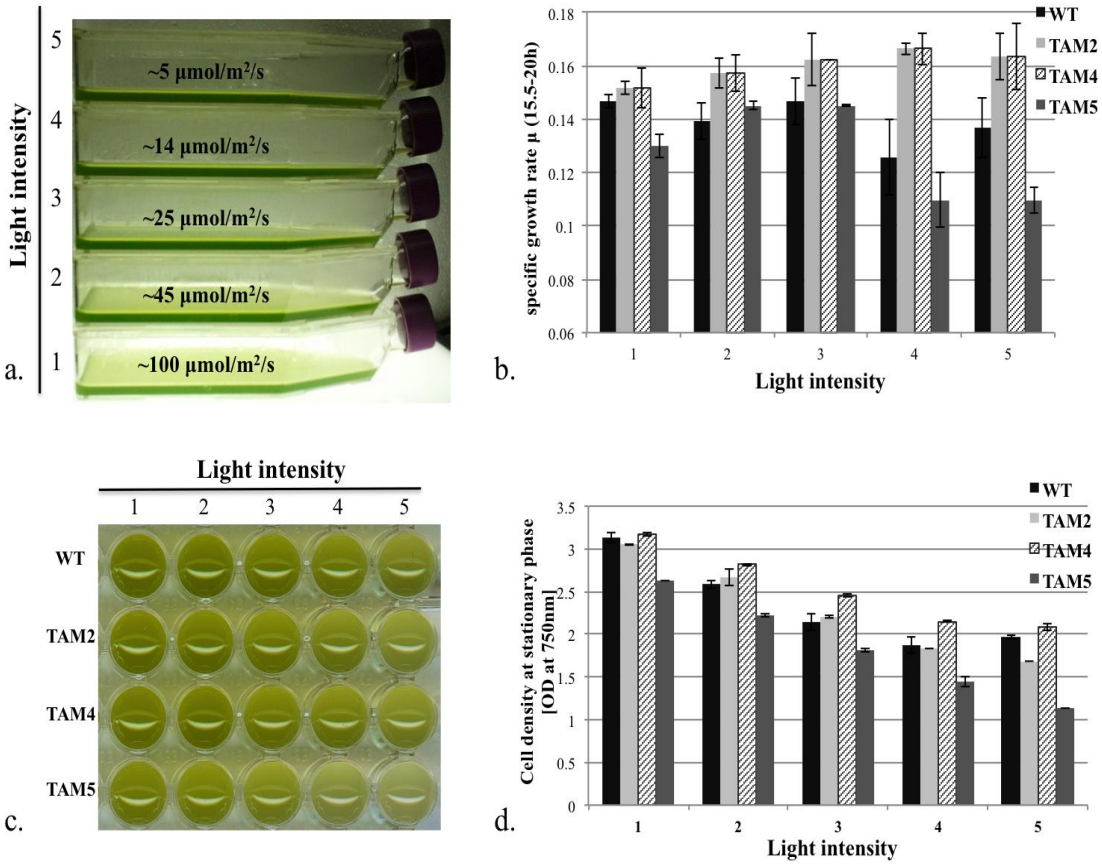


Figure 3-19 Photosynthetic growth analysis of wild type (WT) and three truncated chlorophyll antenna mutants: TAM2, TAM4 and TAM5 of *C. sorokiniana* UTEX1230.

Each strain was maintained in minimal medium (HSM) in five flasks in a form of a stack (a.) illuminated from one side only to record growth rates at the range of light intensity: 5-100 µmol m^{-2} s^{-1}. The specific growth rate at the logarithmic phase was calculated (b) and cell density the cultures at the stationary phase (c) was measured in a form of optical density OD at 750mn (d.) for all light conditions. Error bars show ±SD (n=3).

Algal cultures at the stationary phase (Figure 3-19.c) were further examined for cell density (Figure 3-19.d). TAM4 turned out to be the most productive isolate, as the cell density was the highest out of all tested strains across the rage of irradiance intensity, whereas TAM5 was the least productive. The improved productivity of TAM4 compared to WT can be explained in two ways. Firstly, a reduced chlorophyll antenna is more efficient at capturing the light during photosynthesis, as it has been demonstrated that reduction in the chlorophyll antenna minimises the risk of over-absorption with saturated light. This results in a decrease in the dissipation of excitation energy hence, the efficiency of photosynthesis increases (Polle et al., 2003). Secondly, decreased amount of chlorophyll in cells enables deeper penetration of light deeper into the culture as reduced chlorophyll concentration during algal growth at high irradiance has been proved to increase the productivity (Melis et al., 1998, Kosourov et al., 2011). Additionally, the altered balance chlorophyll/carotenoids in TAM4 mutant may result in increased photoprotection (Fiore et al., 2012).

The *Chlamydomonas reinhardtii Tla1* mutants with truncated light-harvesting chlorophyll antenna showed a 35% reduction of the functional PSII antenna (Polle et al., 2003, Polle et al., 2002), whereas the *C. sorokiniana* UTEX1230 mutants TAM 1-6 were a product of a random UV-induced mutagenesis that resulted in a 45% reduction of the functional PSII antenna. It has been reported that PSII chlorophyll antenna chlorophyll (Chl) b-less mutant of *C. reinhardtii* was reduced to 70%, which resulted in a higher (per Chl) photosynthetic productivity of the cells (Polle et al., 2000). Hence, there is a possibility for further manipulation of the photosynthetic antenna size in the TAM isolates.

3.4 Conclusions and future work

UTEX1230 is the best performing *Chlorella sorokiniana* strain in terms of growth rates in a wide range of temperature conditions and various types of growth medium. The prospective work on UTEX1230 involves testing the lipid productivity in various conditions to find the optimal setting for maximal biofuel yield. Work performed currently in the Purton lab includes nitrogen deprivation as a main induction method for biofuel production in green algae. It has been recently reported that other types of environmental stress also induce lipid productivity, for instance P, S, K, Ca or Mg deficiency resulted in increased lipid production in *C. reinhardtii* and *C. vulgaris* (Deng

et al., 2011a). Additionally, increased salinity has been linked to increased lipid production in marine algae (Peeler et al., 1989) but also in a freshwater alga such as *C. reinhardtii* (Siaut et al., 2011), and future work will investigate whether this can be achieved in UTEX1230.

In recent years the idea of combining wastewater treatment and biofuel production from algae (Christenson and Sims, 2011, Kumar et al., 2010, Woertz et al., 2009, Bordel et al., 2009, Li et al., 2008) or incorporating flue gas in order to increase the biofuel yield (Collet et al., 2011, Li et al., 2008) has been of great interest. Therefore, the next step would be establishing if using wastewater or flue gas in culturing *C. sorokiniana* for biofuel is plausible.

Chlorella sorokiniana UTEX1230 turned out to be the fastest growing *Chlorella* strain and by far one of the most robust algal strain in terms of growth and potential lipid productivity. In order to further increase the productivity of UTEX1230, six mutants with truncated chlorophyll antenna were isolated and evaluated in terms of biochemical activity and their productivity in changing light conditions.

The chlorophyll a/b ratio and total chlorophyll concentration are good initial indicators for estimating the LHC size, which suggested that out of 6 identified mutants, TAM2 and TAM4 are the isolates with truncated antenna. Those results were further confirmed with the biochemical and functional activity of the LHC and supplemented with analysis of growth in various light conditions, which indicated that TAM4 is able to achieve the highest cell density, particularly in cultures under decreased irradiation. TAM1 possesses the closest biochemical resemblance to WT, although functional antenna size appears to be significantly reduced when compared to WT. TAM2 and TAM4 strains presented strong similarities across the range of experiments based on examination of the biochemical and functional properties of their chlorophyll antenna implying that their LHC structure and capabilities are alike. Having said that, the growth experiments at the range of light intensity proved that cell productivity was improved only in TAM4 in respect to WT and the other TAM isolates. All experiments investigating the chlorophyll antenna presented in this work are limited to LHCII. There is a possibility TAM2 and TAM4 differ in their LHCI activity, and it explains the difference in photosynthetic growth parameters between both isolates. Therefore, a follow up investigation would be to analyse the LHCI activity of TAM4 and also TAM2 strains, but more importantly to test the lipid profile and productivity of both isolates in comparison to WT.

Furthermore, both TAM2 and TAM4 isolates carry the yellow-in-the-dark phenotype that is a result of a mutation in one of the genes required for chlorophyll synthesis. It is plausible this mutation may influence the chlorophyll proportion and is associated with the NPQ and chlorophyll fluorescence. Therefore, it is necessary to exclude the fact that the yellow-in-the-dark phenotype has any association with the reduced chlorophyll concentration. Several revertants of TAM2 and TAM4 conferring a green-in-the-dark phenotype were isolated. The future work will involve comparing the initial biochemical results such as relative fluorescence, total chlorophyll content, chlorophyll a/b and chlorophyll/carotenoid ratios we obtained in original TAM2 and TAM4 isolates with these isolated TAM2 and TAM4 revertants.

Finally, in order to maximise the potential of *Chlorella sorokiniana* UTEX1230, it is necessary to develop molecular techniques for genetic transformation. As the literature lacks reports of reliable transformation methods for this species, attempts to develop molecular transformation techniques for the UTEX1230 strain are reported in the next chapter.

CHAPTER 4 DEVELOPMENT OF TRANSFORMATION METHODS FOR *CHLORELLA SOROKINIANA*

4.1 Introduction

4.1.1 General overview

The creation of molecular-genetic techniques is crucial in research of any algal species and both chloroplast (Boynton et al., 1988) and nuclear (Debuchy et al., 1989) transformation methods have been powerful tools since their first successful reports in *Chlamydomonas reinhardtii*. Not only do they enable more advanced fundamental research into aspects of algal physiology and metabolism (Kindle, 1990), but also provide the potential for improvement and modification of algal metabolism to exploit algae as a factories for biofuel production (Beer et al., 2009) or for production of recombinant proteins for the pharmaceutical and biotechnology industries (Mayfield et al., 2007).

4.1.2 Chloroplast transformation overview

As discussed in Chapter 1, algal and plant chloroplasts are semi-autonomous organelles containing their own genomic information where many important biosynthetic pathways occur. They also function as storage organelles and are able to assemble and accumulate an abundance of many types of macromolecules such as soluble or membrane-associated proteins, chlorophylls, carotenoids, lipids and starch (Purton, 2006). Moreover, chloroplasts are reported to be able to accumulate high level of foreign proteins (Miyagawa et al., 2001), which does not interfere with the chloroplast metabolism. Therefore chloroplasts appear to be an ideal sub-cellular compartment for metabolic engineering and recombinant product accumulation.

Chloroplast metabolism could be manipulated in two ways – via either chloroplast or nuclear genes. In principle, chloroplast transformation is more desirable than nuclear manipulation and the advantages of chloroplast transformation over nuclear transformation have previously been discussed in Chapter 1. However, chloroplast transformation method has been successfully developed for only a few microalgal species, such as *Chlamydomonas reinhardtii* (Boynton et al., 1988), *Porphyridium ssp.* (Lapidot

et al., 2002), *Haematococcus pluvialis* (Steinbrenner and Sandmann, 2006) and *Euglena gracilis* (Doetsch et al., 2001).

Designing a process of a transformation method involves several steps. Firstly a suitable and effective selectable marker needs to be chosen. There are numbers of commonly applied selectable markers developed for chloroplast transformation method and they are characterised in

Table 4-1.

Almost all known selectable markers suitable for chloroplast transformation can be classified in four categories. Firstly, the selection method can be based on restoration of photosynthetic activity where a deleted chloroplast gene essential for photosynthesis is reintroduced into the genome together with the gene of interest (GOI) (Figure 4-1.b). Second and third groups of markers confers resistance to antibiotics (Li et al., 2010) and herbicides (Randolph-Anderson et al., 1998). Both classes are relatively diverse and differ in various properties such as dominance but also cell-autonomy (Day and Goldschmidt-Clermont, 2011). The dominant markers *aadA* or *aphA-6* are capable of inactivating the antibiotics spectinomycin and streptomycin (Spc + Str) or kanamycin and amikacin (Kan + Amk), respectively. Alternatively, markers can be endogenous genes such as *rrnS* or *rrnL* containing recessive point mutations that result in alleviation of the sensitivity to Spc, Str or erythromycin respectively. As algal chloroplast is highly polyploid, the selective phenotype of the recessive marker is revealed only if sufficient copies of the chloroplast DNA contain the recessive gene, whereas dominant selection is displayed at the very early stage after chloroplast transformation. Furthermore, a specific phenotype is limited exclusively to the cell that expresses the specific recessive gene. On the other hand the dominant gene that confers the ability to inactivate the antibiotic or herbicide may influence the growth of adjoining cells by partially reducing the concentration of the surrounding harmful chemical (Day and Goldschmidt-Clermont, 2011). This property may have the effect of reducing the ability to obtain homoplasmic transformants (Figure 4-1.d).

Table 4-1 Selectable markers available for plastid transformation.
Adapted from (Day and Goldschmidt-Clermont, 2011).

Type of marker Examples	Selection type	Organism the marker was developed for	References
Photosynthesis			
atpB	Photoautotrophy	*C. reinhardtii*	(Boynton et al., 1988)
tscA	Photoautotrophy	*C. reinhardtii*	(Goldschmidt-Clermont, 1991)
psaA/B	Photoautotrophy	*C. reinhardtii*	(Redding et al., 1998)
petB	Photoautotrophy	*C. reinhardtii*	(Cheng et al., 2005)
petA, ycf3, rpoA	Photoautotrophy	*N. tabacum*	(Klaus et al., 2003)
rbcL	Photoautotrophy	*N. tabacum*	(Kode et al., 2006)
Antibiotic^{R}			
rrnS	Spectinomycin + Streptomycin	*C. reinhardtii*	(Kindle et al., 1991)
		N. tabacum	(Svab et al., 1990)
		S. lycopersicum	(Nugent et al., 2005)
rrnL	Erythromycin	*C. reinhardtii*	(Newman et al., 1990)
aadA	Spectinomycin + Streptomycin	Over 20 species	(Day and Goldschmidt-Clermont, 2011)
nptII	Kanamycin	*N. tabacum*	(Carrer et al., 1993)
		G. herbaceum	(Kumar et al., 2004a)
aphA-6	Kanamycin + amikacin	*C. reinhardtii*	(Bateman and Purton, 2000)
	Kanamycin	*N. tabacum*	(Huang et al., 2002)
		G. herbaceum	(Kumar et al., 2004a)
Herbicide^{R}			
psbA	DCMU	*C. reinhardtii*	(Newman et al., 1992)
bar	Phosphinothricin	*N. tabacum*	(Iamtham and Day, 2000)
AHAS	Sulfometuron methyl	*Porphyridium sp.*	(Lapidot et al., 2002)
EPSP	Glyphosate	*N. tabacum*	(Ye et al., 2003)
HPPD	Diketonitrile	*N. tabacum*	(Dufourmantel et al., 2007)
Metabolism			
BADH	Betaine aldehyde	*N. tabacum*	(Daniell et al., 2001)
codA	5-fluorocytosine	*N. tabacum*	(Serino and Maliga, 1997)
ARG9	Arg autotrophy	*C. reinhardtii*	(Remacle et al., 2009)
ASA2	Trp analogues	*N. tabacum*	(Barone et al., 2009)

The fourth type of the selective marker is a system where an impaired metabolic activity of the recipient line is restored. *ARG9*, encoding for N-acetyl ornithine aminotransferase is a suitable example of this system since introduction of an *ARG9* cassette to the chloroplast of an arginine auxotrophic strain of *C. reinhardtii* (*arg9-2*) resulted in restored arginine phototrophy (Remacle et al., 2009). In addition, reporter genes such as those for GFP (Franklin et al., 2002) or luciferase (Minko et al., 1999) have been developed and could, in principle, be used to identify transformant lines. However, the low efficiency of chloroplast transformation (approximately 10^{-6} in *C. reinhardtii*) makes such a method impractical.

The second step in developing chloroplast transformation for a new species involves choosing effective cis elements – i.e. promoters and Un-Translated Regions (UTRs) that allow efficient transcription and translation of foreign genes (Hallmann, 2007). Typically, these elements are taken from highly expressed endogenous chloroplast genes such as those encoding subunits of the photosynthetic apparatus. The combination of promoter – 5'UTR – coding sequence of gene of interest (GOI) – 3'UTR is often referred to as "expression cassette" that is a self-contained DNA structure ready to be inserted anywhere within the algal chloroplast genome (Purton, 2006).

When all the cloning steps in the creation of the suitable transformation vector are accomplished, the most appropriate method of delivering the DNA to the genome needs to be chosen. The record of all available chloroplast transformation methods has been discussed in detail in chapter 1.

4.1.3 Nuclear transformation overview

Figure 4-1.c presents the mechanism of insertion of the transgene (GOI) into the nuclear genome whereas Figure 4-1.e represents the schematic illustration of the nuclear transformant of *Chlamydomonas*.

Since its successful report in *C. reinhardtii* (Debuchy et al., 1989), nuclear transformation has been an efficient method for genetic manipulation for over 25 microalgal strains (Walker et al., 2005b, Hallmann, 2007, Coll, 2006). This suggests that successful delivery, integration and expression of foreign DNA is more challenging when considering the chloroplast genome. Moreover the generation of nuclear transformation constructs does not necessarily require cloning of genomic DNA of the recipient strain,

and this can be very convenient when designing a transformation method for strains with genomes that have not been sequenced yet. In contrast, chloroplast transformation requires genome elements for homologous recombination and for expression of the foreign gene. Although it is commonly accepted that during the nuclear transformation of most algae a random insertion of the transgene occurs, there are reports on homologous recombination taking place during nuclear transformation in algae such as *Cyanidioschyzon merolae* (Minoda et al., 2004) and in the moss *Physcomitrella patens* (Schaefer, 2002). Therefore, there is a chance it could be possible to control the foreign gene integration within the nuclear genome in other algal species and make the nuclear transformation method more desirable.

Previous reports on nuclear transformation in *Chlamydomonas* investigated using heterologous selectable markers. For instance, the recipient *Chlamydomonas arg7* mutants were rescued with the yeast *ARG4* gene (Rochaix and van Dillewijn, 1982), acquired resistance to the aminoglycoside antibiotic G418 through transformation with the bacterial aminoglycoside 3'-phosphotransferase (APH) gene (Hasnain et al., 1985) or the paromomycin resistance obtained through transformation with the aminoglycoside 3′-phosphotransferase typeVIII (*APHVIII*) encoding gene from *Streptomyces rimosus* (Sizova et al., 1996). Heterologous transgene inserted into the algal nucleus however, tend to be inefficiently expressed. This has been linked to gene silencing, biased G/C content that affects the chromatin structure and hypermethylation of the inserted DNA (Cerutti et al., 1997) but also the absence of the polyadenylation signal found commonly in algal nuclear genes (Silflow and Youngblom, 1986). This problem is diminished when the transgene presents a selective advantage. Nevertheless, genes isolated from foreign species are poor selectable markers for nuclear transformation. Therefore using the homologous selectable markers in theory should increase the efficiency of nuclear transformation.

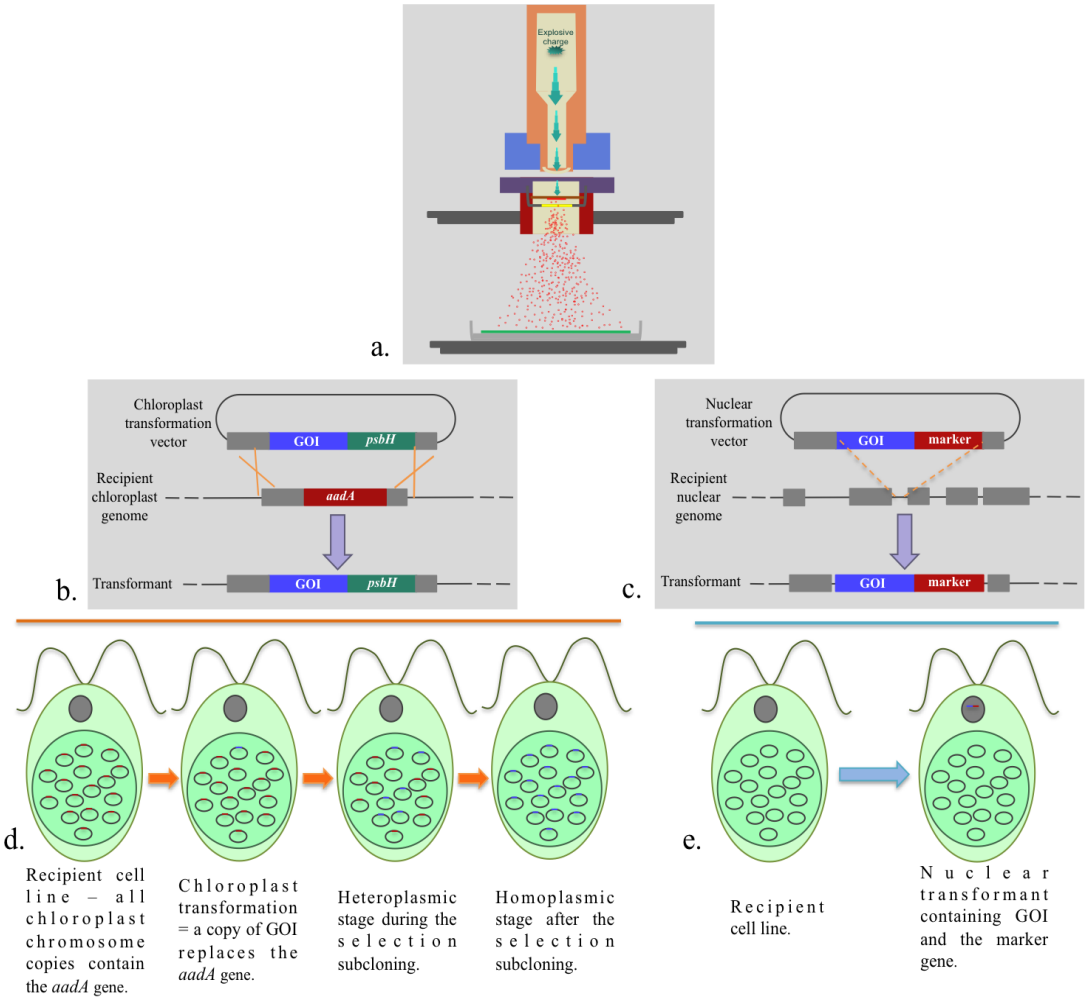


Figure 4-1 Schematic illustration of algal chloroplast (cp) and nuclear (n) transformation in *Chlamydomonas* using the microparticle bombardment method.

a. High-pressure helium accelerates the DNA-coated microcarriers that are shot into a lawn of algal cells at a velocity of approximately 400 m/s.

b. Chloroplast transformation vector contains the gene of interest (GOI) and a selectable marker (here *psbH* gene essential in photosynthesis) flanked by cpDNA that enables successful integration into the plastid genome.

c. Nuclear transformation vector contains the gene of interest (GOI) and a selectable marker that is randomly inserted into the nuclear genome.

d. Initial integration of the GOI+marker construct into one copy of the plastid genome resulting in heteroplasmic cell. Several rounds of selection enables the recovery of a homoplasmic cell line.

e. A single copy of the GOI+marker construct integrated in the nDNA is necessary for successful nuclear transformation.

Since the first successful report of nuclear transformation of an *C. reinhardtii arg*7 mutant with the wild-type *ARG7* gene (Debuchy et al., 1989) and first re-integration of the previously deleted *atpB* gene in to the algal chloroplast (Boynton et al., 1988) there has been an increasing interest in developing transformation systems for other unicellular algae in order to maximize their metabolic capabilities such as production of recombinant proteins, engineered photosynthesis, improved growth capabilities (Rosenberg et al., 2008) but also in a foreseeable future of biodiesel production. One of the promising green algal species from the metabolic point of view is *Chlorella sorokiniana*, as highlighted in chapter 3; chapter 4 describes the various attempts made to develop chloroplast and nuclear transformation methods for this fast growing microalga.

4.2 Background and aims of the project

Chapter 3 presented the potential application of *C. sorokiniana* in economic biofuel production. In order to maximize the alga's potential the recombinant DNA strategies need to be developed. Despite many attempts, there are few reports of stable transformation methods for any *Chlorella* species. The majority of the successful reports on genetic manipulation of *Chlorella sp.* are on *C. ellipsoidea* initially with transient expression of luciferase in protoplasts (Jarvis and Brown, 1991), but eventually a stable transformation was achieved (Chen et al., 2001, Kim et al., 2002). A transformation method was also established for a marine strain *Chlorella sp. MACC/C95* (Han et al., 2005) with a GUS reporter gene under the viral CaMV35S promoter. Previous report on genetic manipulation of *C. sorokiniana* presented the attempts to restore the nitrate reductase activity of the recipient cell line through introduction of the *C. vulgaris* version of this gene (Dawson et al., 1997). No recent evidence on genetic manipulation of *C. sorokiniana* has been reported for the past fifteen years. Hence, the goal of this chapter is to develop successful methods for chloroplast and nuclear transformation of the UTEX1230 strain.

4.3 Results and Discussion

4.3.1 Development of chloroplast transformation methodology

4.3.1.1 Antibiotic resistance characteristics of *C. sorokiniana* UTEX 1230

As shown in

Table 4-1 there are several available markers for chloroplast transformation in green algae. The most common however, are *aadA* and *aphA-6* that confer resistance to spectinomycin + streptomycin and kanamycin respectively.

In order to establish the most suitable selection conditions for chloroplast transformation, the UTEX1230 strain was tested for resistance to the three different antibiotics: spectinomycin (Spc^{100}, 100 µg/ml), streptomycin (Str^{50}, 50 µg/ml) and kanamycin (Kan^{100}, 100 µg/ml), using as controls two strains of *C. reinhardtii* - a Spc+Str-sensitive wild type (WT) and Spc+Str-resistant line. Figure 4-2 presents the outcome of growth tests on medium containing the various antibiotics. This evaluation demonstrates that the tested *Chlorella* strain is resistant to Kan^{100}. Furthermore adding Str^{50} appears to have no effect on UTEX1230 growth whereas addition of Spc^{100} does not inhibit completely its growth. A combination of Spc^{100} + Str^{50} however, appears to completely suppress UTEX1230 growth. Hence it was decided to use the *aadA* selectable marker coding for aminoglycoside 3' adenyl transferase enzyme, which confers streptomycin and spectinomycin resistance (Goldschmidt-Clermont, 1991).

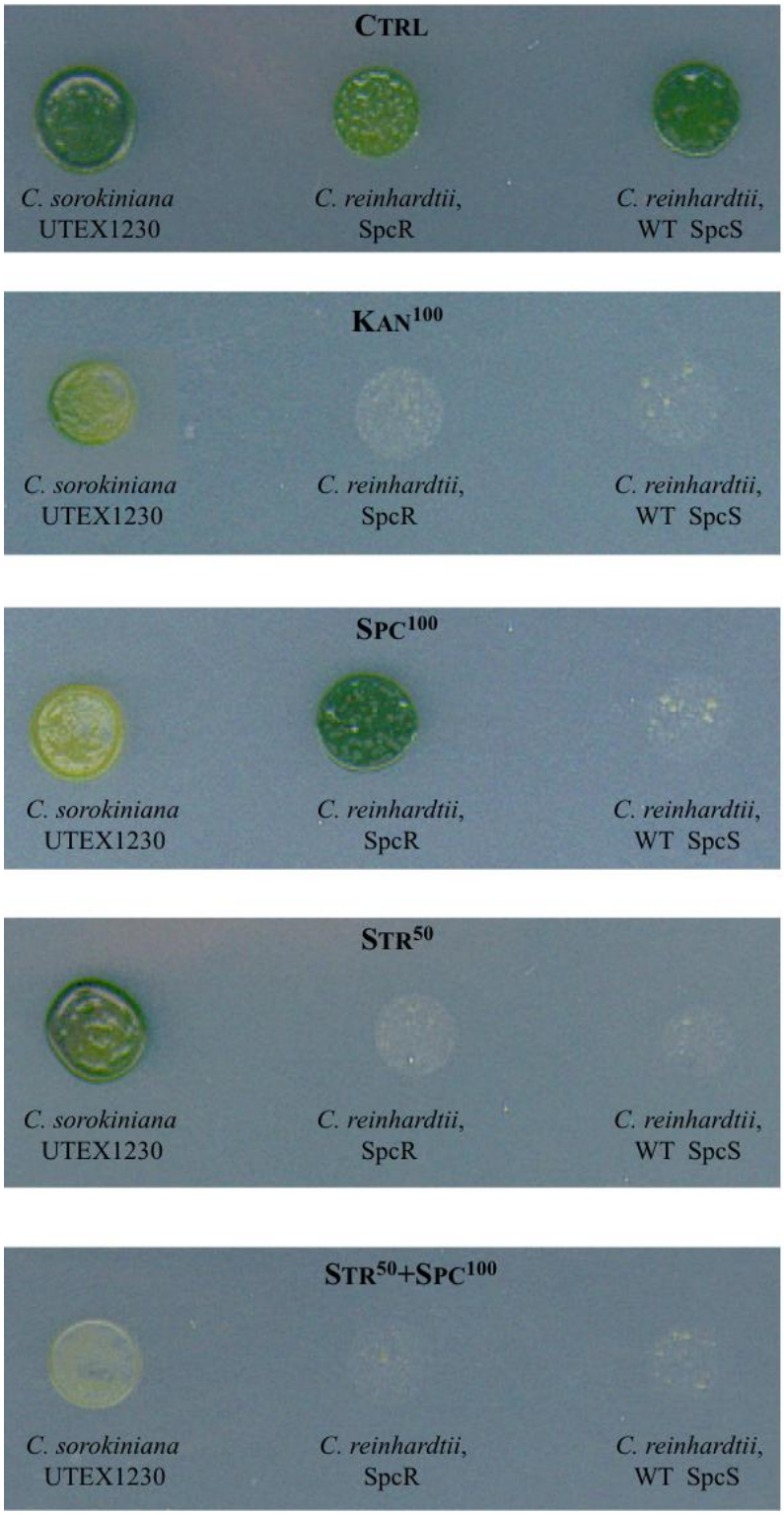


Figure 4-2 Spot test evaluation of antibiotic resistance of *Chlorella sorokiniana* UTEX1230.

Cell suspensions of *C. sorokiniana* UTEX1230 and two strains of *C. reinhardtii*: spectinomycin resistant (SpcR) and spectinomycin sensitive wild type (SpcS WT) were applied on TAP plates containing three types of antibiotics: streptomycin (Str^{50}), spectinomycin (Spc^{100}), kanamycin (Kan^{100}) and also a combination of spectinomycin and streptomycin (Str^{50} + Spc^{100}). Streptomycin concentration was 50 µg/ml whereas spectinomycin and kanamycin were 100 µg/ml.

4.3.1.2 Cloning and sequencing of a chloroplast DNA fragment of *C. sorokiniana* UTEX 1230

As discussed earlier and shown in Figure 4-1.b, the insertion of foreign DNA into the chloroplast genome occurs via homologous recombination. Therefore, it is necessary to isolate a fragment of chloroplast DNA of *C. sorokiniana* UTEX1230 to create the homologous regions flanking the marker gene within a transformation vector. A major limitation in the development of chloroplast transformation method for *C. sorokiniana* is lack of the available chloroplast genome sequence. Consequently, it was decided to isolate a suitable region of the *C. sorokiniana* genome by PCR using primers based on the published genome of the related species *C. vulgaris* [*C. vulgaris* C-27 chloroplast DNA, complete sequence, accession no: AB001684]. Two degenerate oligonucleotides psbB.F and petB.R were designed to amplify the *psbB-petB* region from gDNA isolated from *C. sorokiniana* UTEX1230. The *psbB-petB* fragment was cloned into the *Sma*I site of the pUC9 vector to create plasmid pUC.Cs.C3.1, and was sequenced using the following primers: RevM13, FwII.Cs.BPB, FwIII.Cs.BPB, FwIV.Cs.BPB, FwM13, RevII.Cs.BPB, RevIII.Cs.BPB and RevIV.Cs.BPB. Figure 4-3 shows the sequencing strategy for the *psbB-petB* insert and the orientation of the *psbB*, *clpP* and *petB* genes within the pUC.Cs.C3.1 plasmid. The sequenced 3.14 kb fragment was published in GenBank under the accession number FJ623757 and the full sequence can be found in the appendix B.

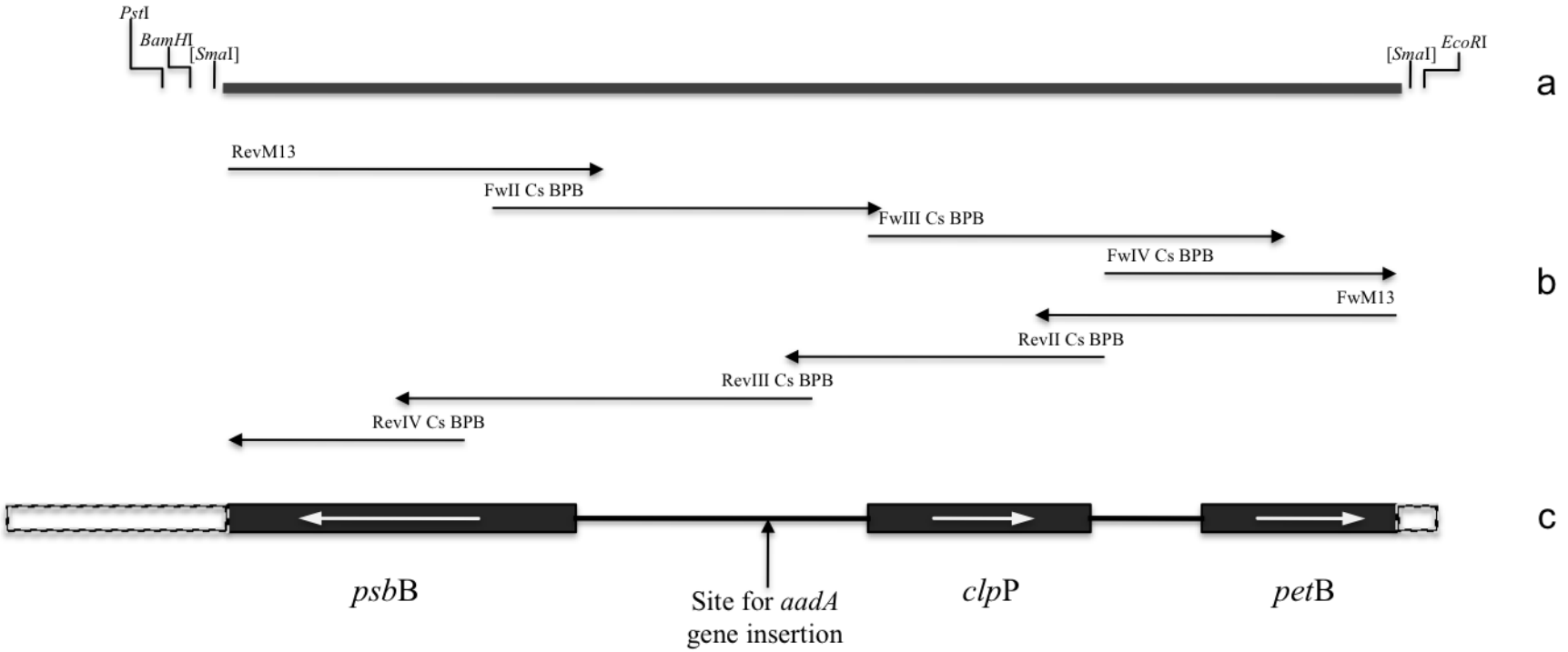


Figure 4-3 Detailed map and the sequencing steps of the *psbB-petB* insert.

a) The orientation of the cloned chloroplast DNA fragment in the pUC9 plasmid
b) Primers used for sequencing of the *psbB-petB* insert
c) The orientation of the *psbB*, *clpP* and *petB* genes in the cloned DNA fragment; *clpP* and *petB* genes are transcribed in the same direction, whereas *psbB* gene is transcribed from the opposite strand; parts of the *psbB* and *petB* genes not included in the *psbB-petB* insert appear white.

Despite this cloning success, a comparison of the *C. sorokiniana* and *C. vulgaris* gene maps reveals remarkable fortune in the design of the degenerate primers, as it become evident that the gene order within the cloned fragment is inverted (Figure 4-3) compared to *C. vulgaris* (Wakasugi et al., 1997) where not only is *psbB* in the same orientation as *clpP* and *petB* but it is also separated by nine ORFs.

This feature has the potential to play a significant role in establishing an additional diagnostic methodology for *Chlorella* species, however further analysis is required. Inverted genes within the chloroplast genomes of related species are features commonly occurring among even closely related algal strains. It has been reported that the *Chlamydomonas pitschmannii* chloroplast genome contains three noticeable differences in gene arrangement compared to the genome of its close relatives *C. eugametos* and *C. moewusii* (Boudreau and Turmel, 1995). Those changes are associated with either inversions of the genes within the single copy region or are linked to expansion/contraction activities of the inverted repeat.

4.3.1.3 Creation of the pU.Cs3-*aadA* plasmid

The detailed steps of the plasmid pU-Cs3-*aadA* creation are presented in Figure 4-4. The amplified *psbB*-*petB* fragment from chloroplast genome was cloned into *Sma*I site of pUC9 vector and pUC.Cs.C3.1 plasmid was created. In order to avoid the accidental disruption of the gene transcription in the transformed cell line of *C. sorokiniana*, the region between the *psbB* and *clpP* genes that are transcribed from opposite DNA strands was chosen for insertion of the selectable marker gene. However, there were no suitable restriction sites available in this *psbB – clpP* region. Consequently, a unique *Eco*RV restriction site was generated using site directed mutagenesis (see subchapter 2.5.9) and pUC.Cs.C3.1-RV plasmid was created.

The *aadA* cassette conferring spectinomycin resistance was obtained by double digestion of the pUC-atpX-AAD plasmid (Goldschmidt-Clermont, 1991) with *Eco*RI and *Sma*I restriction enzymes. The 5' overhang originated after the *Eco*RI digestion was subjected to DNA Polymerase I Large (Klenow) Fragment activity to form a blunt end. The isolated cassette contains the coding region for the aminoglycoside 3' adenyl transferase enzyme flanked by the promoter/5'UTR region of *atpA* and the 3'UTR of *rbcL* from *C. reinhardtii*. This was inserted into the newly created *Eco*RV restriction site between *psbB* and *clpP* genes within the pUC.Cs.C3.1-RV plasmid to generate the pUC.Cs.C3.1-RV-*aadA* plasmid.

Subsequently, the pUC.Cs.C3.1-RV-*aadA* plasmid was subjected to double digestion with *Eco*RI and *Pml*I that resulted in deletion of a 0.18 kb fragment of the *petB* gene. The 5' overhang created after the *Eco*RI digestion was treated using DNA Polymerase I Large (Klenow) Fragment to form a blunt end and the shortened plasmid was re-ligated. This partial *petB* gene deletion was performed in order to develop an efficient PCR-based method to identify true *C. sorokiniana* transformants. In this method, two primers were designed: a primer whose sequence is within the *aadA* gene: Cs.3.1.F [GGATCGTTTAAAACTACTGTATGAACTCTATACC] and a primer within the deleted fragment of the *petB* gene: Cs3.Pml-Eco.R [GAACACCTGTTACAATTTTTACCGCC]. The created method allows for identification of the correct chloroplast transformants from the free plasmid and from the transformants with random integration of the *aadA* gene within the genomic DNA.

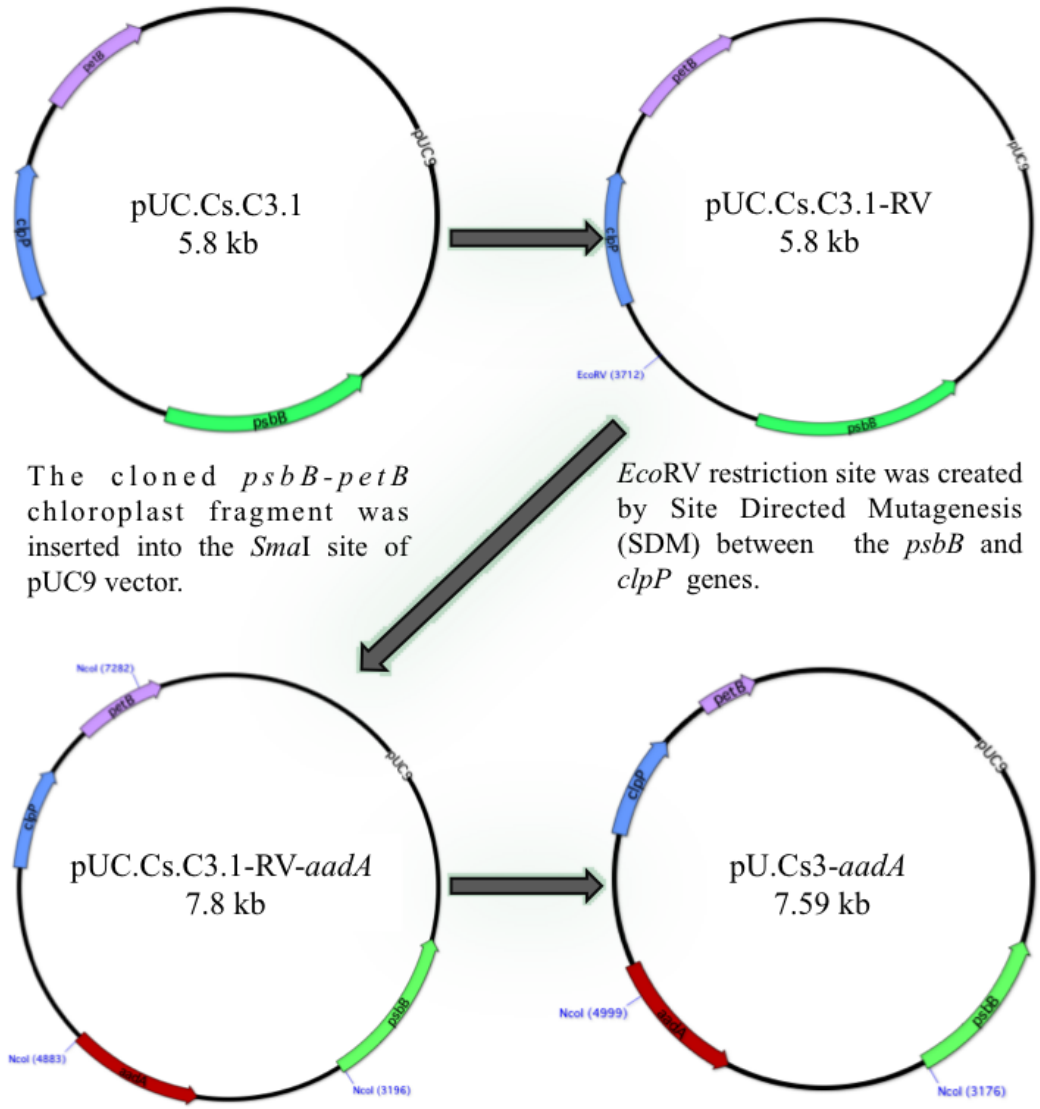


Figure 4-4 Schematic illustration of the pU.Cs3-*aadA* plasmid generation

4.3.1.4 Chloroplast transformation of *C. sorokiniana* UTEX1230 with the pU.Cs3-*aadA* plasmid

The pU.Cs3-*aadA* plasmid was introduced into the *C. sorokiniana* chloroplast using the microparticle bombardment method (Figure 4-1). A week after bombardment, single colonies that appeared on the selective TAP plates containing Spc^{100}+Str^{50} were streaked onto new selective plates with both antibiotics. Three rounds of re-streaking of thirty isolates on fresh TAP plates containing Spc^{100}+Str^{50} were performed in order to increase the probability of selection of homoplasmic transformants. Eventually, eight promising cell lines: 2, 13, 27, 102, 103, 104, 201 and 204 were isolated for further analysis.

4.3.1.5 Analysis of created putative transformants

Total genomic DNA was prepared from the eight selected isolates and PCR analysis was performed using the Cs.3.1.F and Cs3.Pml-Eco.R primers. Unfortunately, the attempts to amplify the targeted DNA fragment in any of the putative transformant lines were unsuccessful (data not shown). A 2-step PCR method was therefore designed to amplify the *aadA* gene within the putative transformant lines. During the first PCR amplification primers atpA.F [CAAGTGATCTTACCACTCAC] and rbcL.R [GCTGCTGCATGTGAAGTTTG] were used and the product from this reaction was later employed as a template for a second PCR with primers atpA.F2 [CAAGTGATCTTACCACTCAC] and rbcL.R2 [GCTGCTGCATGTGAAGTTTG]. Both sets of primers were designed to target the UTR regions within the *aadA* cassette: *atpA* 5'UTR and *rbcL* 3'UTR and the results are presented in Figure 4-5.a (first PCR reaction) and Figure 4-5.b (second nested PCR reaction). In both cases three controls were applied – H_2O where no DNA was added, pUC.Cs.C3.1-RV – plasmid that does not contain the *aadA* cassette used as a template for negative control, pU.Cs3-*aadA* – plasmid containing the *aadA* cassette that was used as a positive control.

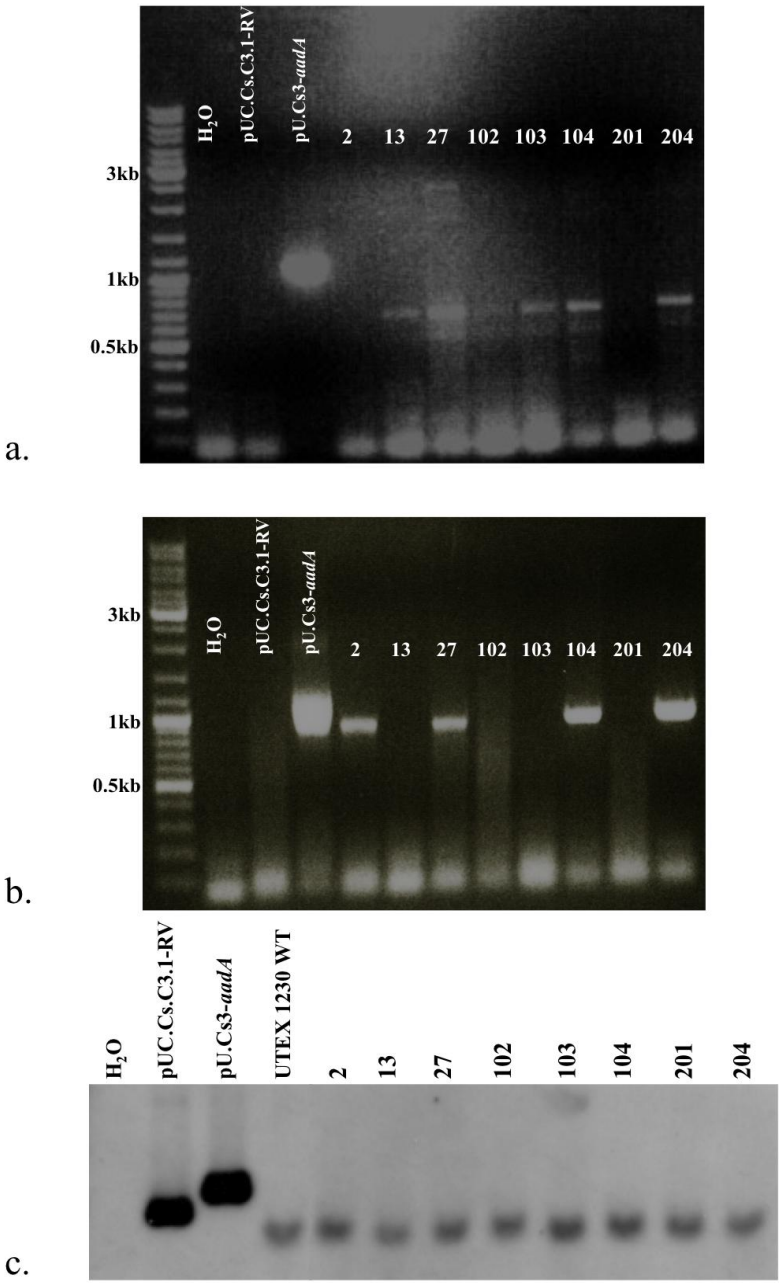


Figure 4-5 Molecular analysis of the putative transformants of *C. sorokiniana* UTEX1230.

H_2O, pUC.Cs.C3.1-RV and **UTEX 1230 WT** – negative controls
pU.Cs3-*aadA* – positive control
2, 13, 27, 102, 103, 104, 201, 204 –selected putative transformants of *C. sorokiniana*

a. First step PCR performed with primers atpA.F and rbcL.R, expected size 1.1 kb
b. Second (nested) PCR amplification with atpA.F2 and rbcL.R2 primers where PCR product from previous reaction was used as a DNA template, expected size ~0.9 kb
c. Southern blot analysis of the putative transformants. Control plasmids and genomic DNA isolated from WT and putative transformants were digested with *Pst*I enzyme. The template for the probe was designed to detect the flanking region around the *aadA* cassette of the pU.Cs3-*aadA* plasmid. Expected band for WT and plasmid pUC.Cs.C3.1-RV is of ~1.4 kb whereas for transformed cell lines and plasmid pUC.Cs.C3.1-RV::*aadA* two bands of ~ 1.6 kb and ~1.8 kb are anticipated to appear.

Several bands appeared in the first PCR reaction of the putative transformants (Figure 4-5.a) yet the second PCR reaction resulted in amplification of the correct size band in four out of eight tested isolates: 2, 27, 104 and 204 (Figure 4-5.b).

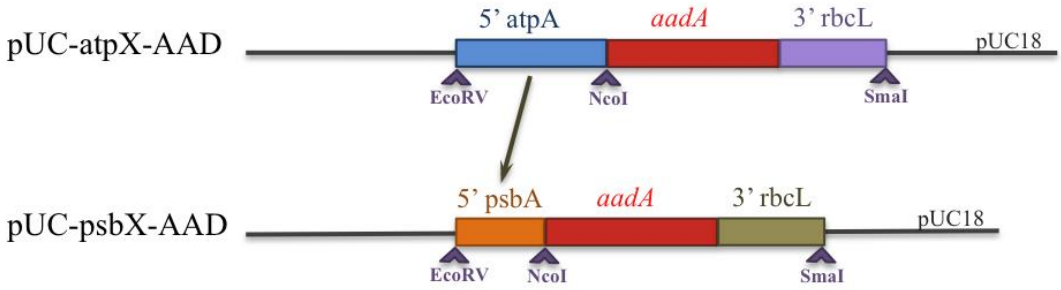
The two-step PCR method confirmed the presence of *aadA* cassette in four putative transformants: 2, 27, 104 and 204. In order to confirm the presence of this marker in the expected location of chloroplast DNA, a Southern blot analysis was designed. The PCR product of the following primers: C3.1-RV:PstI.F [AATGTGGTGTGCTGAAATGC] and C.3.1-RV:PstI.R [AAAGGGACTTTAGGGACACC] served as a template to create the probe for Southern blot analysis. This probe was designed to detect the flanking region *psbB-petB* around the *aadA* insertion within the chloroplast DNA of the putative transformants 2, 27, 104 and 204. The results are presented in Figure 4-5.c. In this evaluation, two bands of approximately 1.6 kb and 1.8 kb are expected for positive control (plasmid pU.Cs3-*aadA*) and true transformants. Unfortunately, all of the putative transformants showed a band of approximately 1.4 kb characteristic for WT and negative control pUC.Cs.C3.1-RV plasmid.

PCR analysis is an extremely sensitive method for detecting the presence of a specific DNA fragment. It is plausible the designed PCR reaction amplified free pU.Cs3-*aadA* plasmid as it has been reported that free plasmids are able to persist in the algal cells cell for generations after they were successfully introduced into their chloroplast (Boynton et al., 1988). It is also possible the *aadA* cassette was incorporated into another position in the genomic DNA of the cell lines that showed the presence of this selectable marker during the PCR analysis (Figure 4-5.b). Another credible reason for lack of success of the chloroplast transformation performed on *Chlorella sorokiniana* could be the unsuitable promoter as the *aadA* cassette described earlier was controlled by the 5'UTR and *atpA* promoter from *C. reinhardtii*. Current knowledge on gene expression and regulation in chloroplasts is still very limited, however it is widely accepted that this process is controlled by various post-transcriptional factors encoded by the nuclear DNA and involves diverse aspects of mRNA modification such as processing, stability and translation. It has been demonstrated that such post-transcriptional control is associated with interactions between the 5'UTR of particular genes with these nuclear-encoded factors (Rochaix, 1996), therefore choosing the appropriate 5'UTR region controlling the expression of the transgene is a crucial step in the assembly of a successful transformation vector.

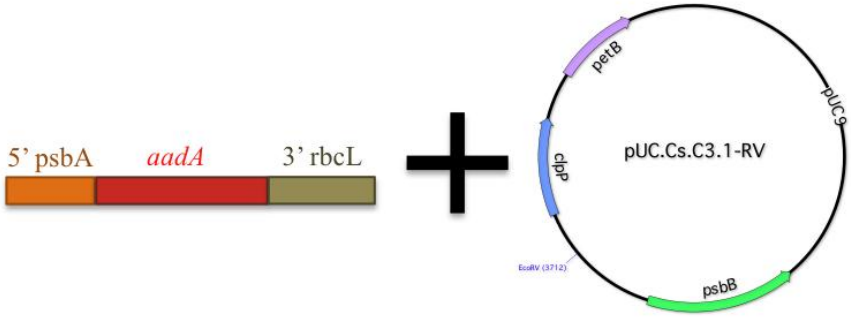
In addition, naturally occurring spontaneous mutations conferring resistance to streptomycin and spectinomycin are not helpful in the identification of true transformants of *C. sorokiniana*. The translational apparatus of chloroplasts is prokaryotic in origin. Consequently, mutations within the genes coding for ribosomal proteins or rRNA can result in resistance to certain antibiotics, for instance changes within the 16S rRNA gene (*rrnS*) results in spectinomycin or streptomycin resistance (Newman et al., 1990, Svab and Maliga, 1993). Cell lines transformed with *aadA* marker are capable of tolerating very high concentration of Spc and Str even up to 500 μg of each antibiotic per 1 ml of growth medium (Svab and Maliga, 1993). Therefore increasing the concentration of the selective element may result in decreased occurrence of the spontaneous mutations leading to resistance to Spc and Str.

4.3.1.6 Creation of pUC.Cs7.3–psbX–*aadA* plasmid

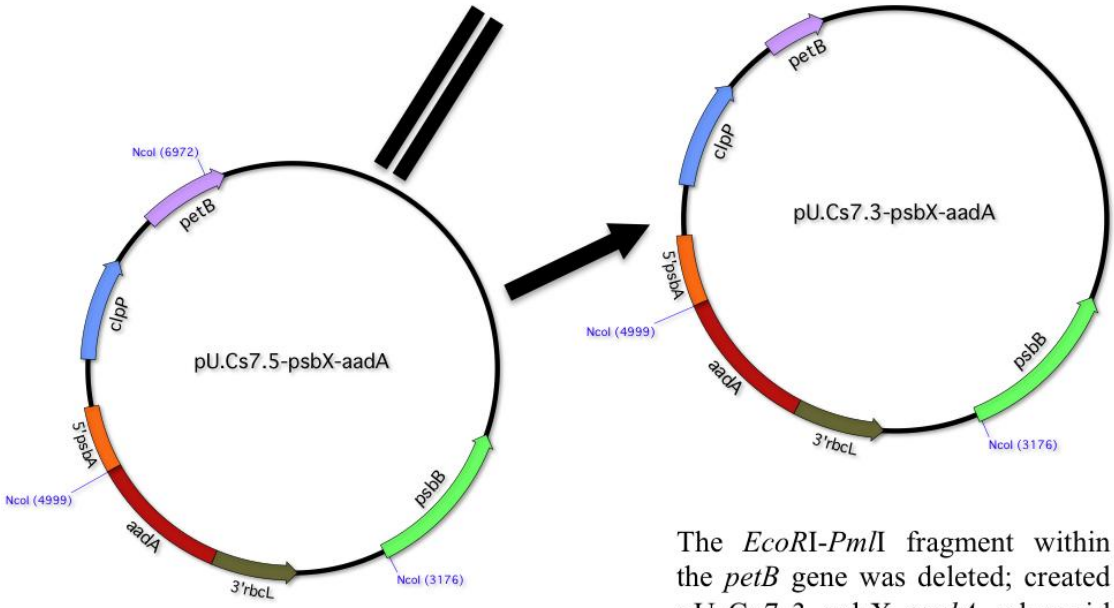
The lack of success in isolating true transformants following the transformation with pU.Cs3-*aadA*, in which *aadA* expression is controlled by the *C. reinhardtii atpA* promoter/5'UTR, resulted in the design of an alternative cloning strategy. The *atpA* element was replaced with a promoter/5'UTR sequence from the *psbA* gene of *C. sorokiniana* to give new plasmid pUC.Cs7.3–psbX–*aadA*. The *psbA* promoter/5'UTR sequence was amplified using the psbA_NcoI.F [CCAACCCATGGATAAACGGTTTTCAGTGC] and psbA_EcoRV.R [GCTTGATATCTAAGTTTAGAGAGGGGTGATCG] primers where chloroplast DNA isolated from *Chlorella sorokiniana* H-1983 strain served as a template. The primers were designed to introduce two restriction sites – *Nco*I and *Eco*RV in order to facilitate cloning. The full sequence of the isolated *psbA* promoter is presented in the appendix D. The cloning steps for the creation of the pU.Cs7.5-psbX-*aadA* plasmid are presented in Figure 4-6.



Replacement of the 5'*atpA* promoter region with a *psbA* promoter region from *C. sorokiniana* within the *aadA* cassette.



aadA gene flanked by 3'*rbcL* and the introduced 5'*psbA* promoter region was inserted into the *Eco*RV site of the pUC.Cs3.1-RV plasmid



The newly created pU.Cs7.5-psbX-*aadA* plasmid contains three *Nco*I restriction sites

The *EcoR*I-*Pml*I fragment within the *petB* gene was deleted; created pU.Cs7.3-psbX-*aadA* plasmid contains two *Nco*I restriction sites.

Figure 4-6 Schematic illustration of the pU.Cs7.3-psbX-*aadA* plasmid construction.

The original *atpA* element was removed from pUC-atpX-AAD (Goldschmidt-Clermont, 1991) after double digestion with *Eco*RV and *Nco*I and the *C. sorokiniana psbA* region was cloned into the vector to give plasmid pUC-psbX-AAD. The new *aadA* cassette was then cloned into the pUC.Cs.C3.1-RV vector. As before, the newly created pU.Cs7.5-psbX-*aadA* plasmid was subjected to double digestion with *Eco*RI and *Pst*I and religated, resulting in the deletion of a 0.18 kb fragment of *petB* within the plasmid.

Successful integration and expression of a transgene depends on the presence of various necessary factors within the transformation vector. Firstly, the gene needs to be controlled by an appropriate 5'untranslated region (UTR) that is necessary in RNA structure stabilisation via created *cis* elements but also very important for ribosome binding and initiation of translation. Additionally, the transgene needs to be controlled by an efficient promoter and both 5'UTR and promoter regions tend to be derived from the same endogenous gene.

It was decided to keep the *C. reinhardtii rbcL* 3'UTR element in this new marker. The 3'UTR region is vital for the effective processing and stabilisation of transcripts (Purton, 2006). It has been demonstrated however, that the choice of 3'UTR has little impact on expression level of transgenes and using 3'UTR regions isolated from various algal housekeeping genes did not affect the accumulation of foreign protein (Barnes et al., 2005, Goldschmidt-Clermont et al., 2008).

4.3.1.7 Chloroplast transformation of *C. sorokiniana* UTEX1230 with the pUC.Cs7.3–psbX–*aadA* plasmid and analysis of putative transformants

The pUC.Cs7.5-psbX-*aadA* plasmid was introduced into *C. sorokiniana* cells by microparticle bombardment, as before. In addition to the biolistic method, an alternative method of the vector delivery – vortexing with glass beads – was tested. The latter method, when performed on cell wall deficient *C. reinhardtii* strains, is an efficient and reliable technique for both nuclear and chloroplast transformation (Kindle, 1990). The cells were vortexed with glass beads and transforming DNA for three time periods of 1 min, 2 min or 5 min. In addition, a further modified method was a one minute heat shock treatment at 45^0C. This additional step has been proven to increase the nuclear transformation efficiency in *Chlorella* sp. (Hawkins and Nakamura, 1999). Cells that were subjected to either of the transformation procedures were plated on TAP plates containing increased concentration of selective antibiotics with up to 300 μg/ml of spectinomycin and 150 μg/ml of streptomycin. Single colonies appeared approximately a week after the treatment (Figure 4-7.a).

Approximately 200-300 colonies appeared on each of the transformation plates. There were no differences in the number of colonies between each of the transformation conditions, and also there were no substantial differences between "+DNA" and "–DNA" plates. However, no statistical analysis was performed because of the high number of colonies. In total, 200 colonies were subjected to three cycles of re-streaking on TAP plates containing increased concentration of the antibiotics: Spc^{300} and Str^{150}. Interestingly, approximately 20% of the re-streaked isolates failed to grow on fresh selective medium. All isolates that retained their Spc+Str resistance phenotype were subsequently isolated for further tests where the isolated genomic DNA was subjected to PCR analysis using the Cs.3.1.F and Cs3.Pml-Eco.R primers. An example of the analysis of putative transformants with positive and negative controls is presented in Figure 4-7.b. Unfortunately, none of the tested isolates showed the presence of the *aadA* gene in their chloroplast genome as the PCR analysis showed their genotype to be the same as WT.

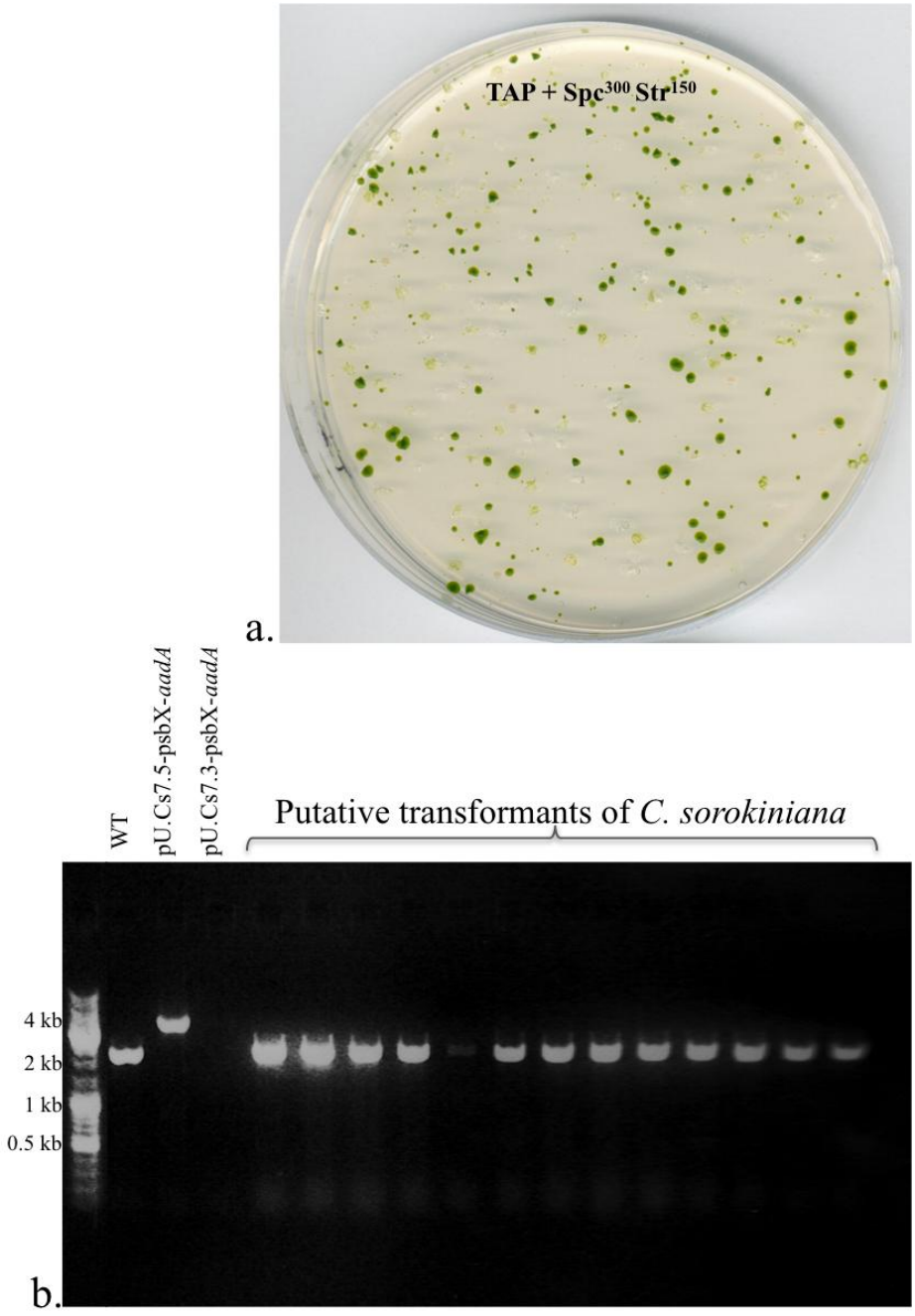


Figure 4-7 Analysis of the putative transformants of *C. sorokiniana* UTEX1230.

a. Putative transformants of *C. sorokiniana* appearing after approximately a week after a chloroplast transformation
b. PCR analysis (primers used: Cs.3.1.F and Cs3.Pml-Eco.R) of the putative transformants of *C. sorokiniana*. WT – wild type of *C. sorokiniana*, pU.Cs7.5-psbX-*aadA*- positive control, pU.Cs7.3-psbX-*aadA*- negative control. Expected bands for WT was ~2.2 kb whereas for positive control and transformants ~3.9 kb.

The possibility that true transformants are present among the colonies that appeared on the selective plates cannot be excluded even though the PCR analysis did not confirm this. Unfortunately, the fact that similar number of colonies appeared on transformed and control plates suggests that the isolates most likely gained their antibiotic resistance randomly and not necessary through the inserted *aadA* gene. This acquired resistance to antibiotics is a result of either undirected spontaneous mutation or direct and specific adaptation acquired as a response to environmental selection (Luria and Delbruck, 1943).

This increased number of false positive results could be explained by an escalated mutation rate to both antibiotics. Spontaneously acquired resistance to the antibiotics spectinomycin or streptomycin in *C. reinhardtii* are associated with different point mutations within the 16S chloroplast ribosomal gene (Harris et al., 1989). It has been demonstrated the antibiotic resistance that occurs as a result of mutations within the domains of ribosomal RNA results in specific changes in primary and secondary structure of rRNA (Sigmund et al., 1984, Moazed and Noller, 1987). In *Chlamydomonas* there are four well-defined chloroplast loci prone to mutations resulting in streptomycin resistance of which three are within the 16S gene whereas spectinomycin resistance is associated with three mutations within a single locus of the 16S gene (Harris et al., 1989).

The pace at which algae adopt resistance to certain antibiotics or herbicides varies among species but also from gene to gene within the same strains (Lopez-Rodas et al., 2001). As might be expected, the mutation rate resulting in resistance to two antibiotics is substantially lower compared to the mutation rate to each individual chemical but again it varies from species to species. It has been demonstrated that many organisms have effective internal regulatory mechanisms that control the rates that mutation occur in response to environmental adaptations (Kepler and Perelson, 1995).

Transformation efficiency is expected to be low when performed on species with no optimised transformation protocol. Various factors can contribute to transformation efficiency such as plasmid concentration, plasmid form (linear or circular), and optimal cell number or growth stage of the recipient cells. Additionally, for the glass bead method there is also vortexing time and the size of the glass beads.

4.3.1.8 Testing the *C. sorokiniana* UTEX1230 *psbA* promoter in *E. coli* and *C. reinhardtii*

The question arises as to whether the cloned region containing the *psbA* promoter/5'UTR isolated from *C. sorokiniana* is actually functional. Therefore, its operational capability in both *E. coli* and *C. reinhardtii* was tested. Firstly, the transformation plasmid pBa3-AX (Hallahan et al., 1995) that has been proven to be functional in *C. reinhardtii*, was modified and inserted two *aadA* constructs under two types of promoter/5'UTR: the *C. reinhardtii atpA* element and newly isolated *C. sorokiniana psbA* element. The cloning strategy and orientation of each gene within the created transformation vectors are presented in Figure 4-8. Both *aadA* cassettes were inserted into the *Eco*RI site within the pBa3-AX plasmid in two separate cloning events. The sequences of the created transformation vectors were confirmed by sequencing and are presented in the appendix F.

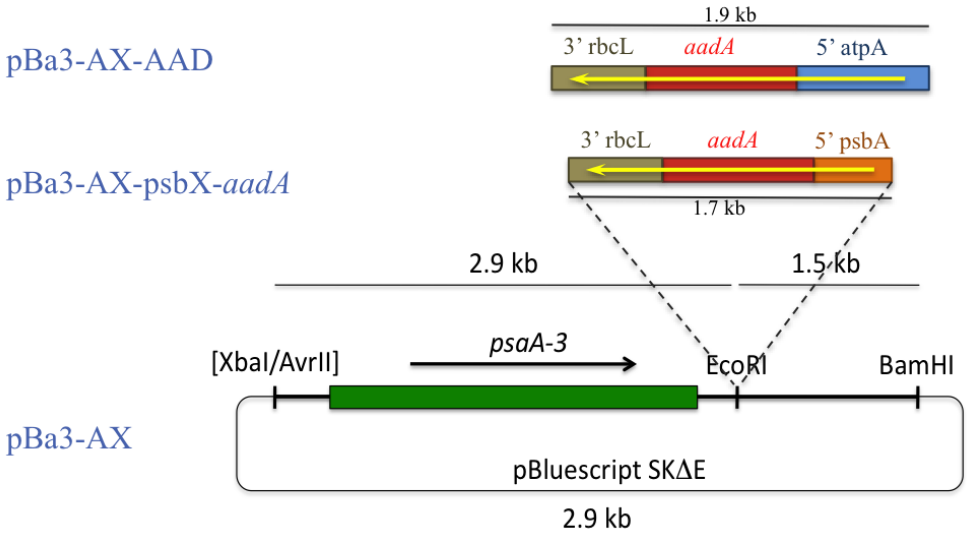


Figure 4-8 The map of pBA3-AX plasmid constructs and its derivatives – pBa3-AX-AAD and pBa3-AX-psbX-*aadA*.

The *atpA* promoter of *C. reinhardtii* from pBa3-AX-AAD plasmid was replaced with *psbA* promoter of *C. sorokiniana* in the pBa3-AX-psbX-*aadA* plasmid.

Firstly, the efficiency of the *psbA* promoter/5'UTR region was tested in *E. coli*. Competent cells were transformed with three transformation vectors: pBa3-AX, pBa3-AX-AAD and pBa3-AX-psbX-*aadA* and the activity of the promoters were tested by plating the transformed cells on selective media containing spectinomycin, ampicillin or a combination of both antibiotics. The results are presented in Figure 4-9.

Both promoters *atpA* in pBa3-AX-AAD and *psbA* in pBa3-AX-psbX-*aadA* proved to be functional in *E. coli* as cells transformed with both plasmids showed resistance to spectinomycin (Figure 4-9.a). The identity of both plasmids used of *E. coli* transformation were confirmed by a test digest with *Eco*RI, *Bam*HI or a double digestion with *Bam*HI and *Nco*I restriction enzymes. The results are presented in Figure 4-9.b.

Testing functionality of created expression cassettes in *E. coli* can be potentially a fast and efficient control method. The growth rate of *E. coli* is substantially faster than the algal strains and does not require using any complicated transformation techniques hence the results are obtained faster than in algal transformation.

The functionality of the *psbA* promoter/5'UTR from *C. sorokiniana* H-1983 was subsequently tested in two strains of *C. reinhardtii*. The first test involved using the cell wall defective strain cw15.J3. Cw15.J3 strain was transformed using the glass beads method with the two transformation vectors pBa3-AX-AAD and pBa3-AX-psbX-*aadA* containing *aadA* cassettes under the different promoters and plated the transformed cells on selective medium containing spectinomycin. After approximately 10 days, colonies started to appear on the plates transformed with the pBa3-AX-AAD plasmid, whereas cells transformed with the pBa3-AX-psbX-*aadA* vector failed to grow on selective medium (Figure 4-10). This result clearly shows the *C. sorokiniana psbA* promoter and 5'UTR is not functional in *C. reinhardtii*.

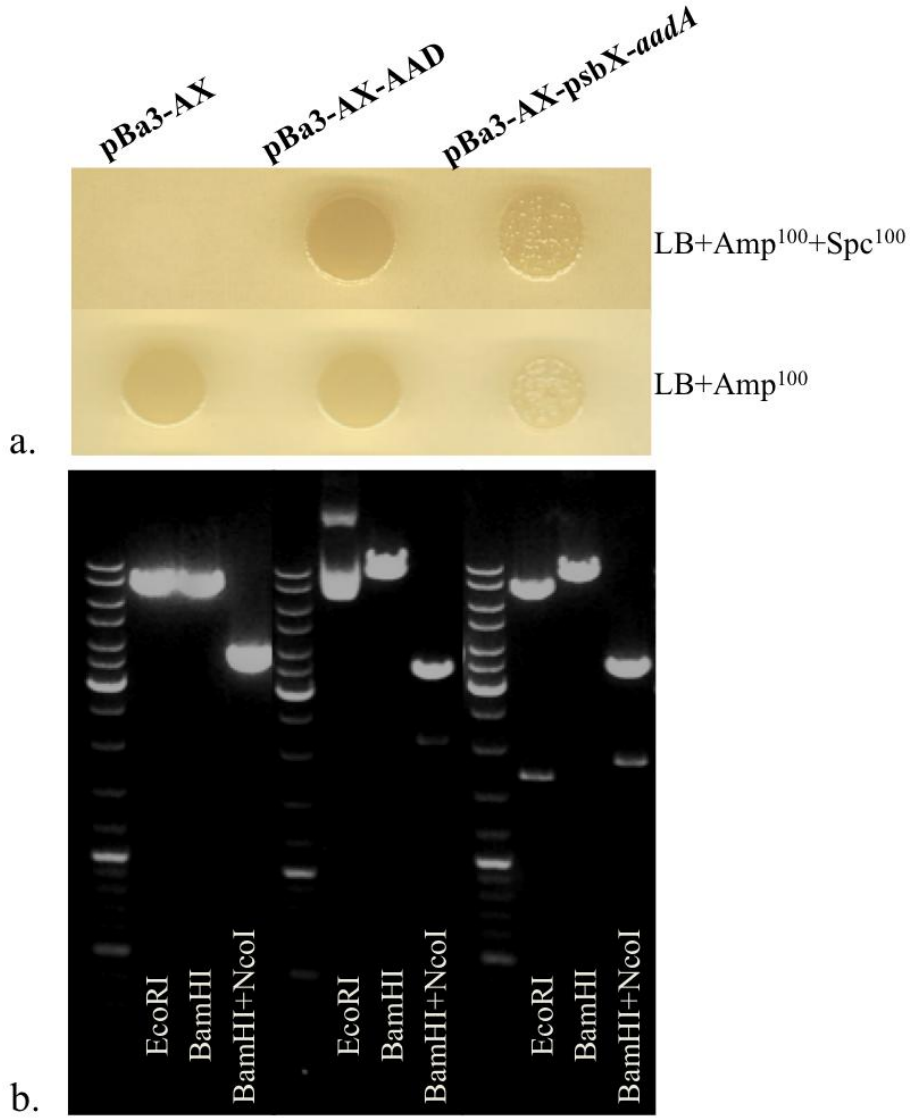
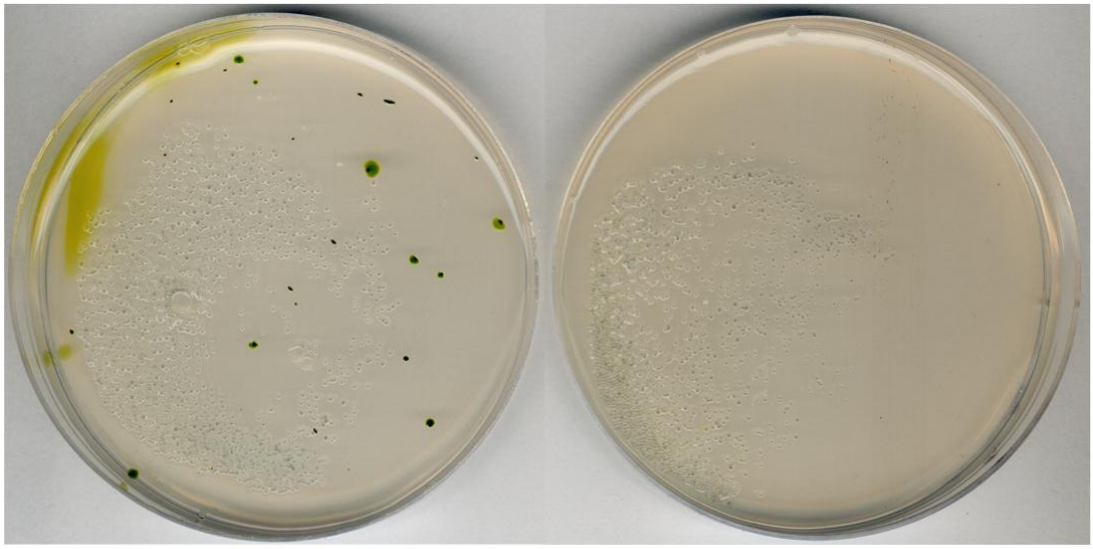


Figure 4-9 Analysis of the pBa3-AX-psbX-*aadA* plasmid activity in *E. coli*.

a. Spot test of *E coli* transformed with pBa3-AX, pBa3-AX-AAD and pBa3-AX-psbX-*aadA* plasmids. Both *Chlorella* and *Chlamydomonas* promoters (*psbA* and *atpA* respectively) appear to be functional in *E. coli* as pBa3-AX-AAD and pBa3-AX-psbX-*aadA* transformed *E. coli* are able to grow on LB plates containing ampicilin (Amp^{100}) and spectinomycin (Spc^{100}); LB+Amp^{100} is used as a control plate and pBa3-AX control plasmid.
b. Test digest of the pBa3-AX, pBa3-AX-AAD and pBa3-AX-psbX-*aadA* plasmids with *Eco*RI, *Bam*HI and *Nco*I restriction enzymes.



pBa3-AX-AAD pBa3-AX-psbX-*aadA*

Figure 4-10 *C. reinhardtii* CW15.J3 cell line transformation with pBa3-AX-AAD and pBa-AX-psbX-*aadA* plasmids.

Putative transformants of *C. reinhardtii* cw15.J3 appeared only after transformation with plasmid pBa3-AX-AAD containing the *C. reinhardtii atpA* promoter. The colonies appeared on selective plates containing spectinomycin approximately 10 days after a chloroplast transformation using glass beads method. Plates transformed with the pBa-AX-psbX-*aadA* plasmid containing the *C. sorokiniana psbA* promoter show no sign of transformed cell lines.

Using heterologous promoters and 5'UTR regions isolated from different species during genetic modification is not a novel idea. Successful incorporation into *C. reinhardtii* nuclear genome of *aphVIII* gene under the *hsp70A-rbcS3* hybrid promoter and 5'UTR region isolated from *V. carteri* has been reported (Hallmann and Wodniok, 2006). Moreover, the same report demonstrated functionality of the 5' and 3' UTRs of *C. reinhardtii* in *V. carteri*. Similarly, a heterologous *RbcS* promoter amplified from *D. tertiolecta* effectively induced *ble* activity in *C. reinhardtii* (Walker et al., 2005a). Cloning heterologous UTR regions that were isolated from genes of model organisms such as *C. reinhardtii* is particularly applicable during development of transformation for new species without a sequenced genome. For instance, a promoter isolated from the *rbcS2* gene of *C. reinhardtii* successfully drove expression of *aphVIII* gene in *Gonium pectorale* (Lerche and Hallmann, 2009). Furthermore, introducing a heterologous promoter may result in increased transgene expression compared to the homologous promoter. For instance, transformation of *C. reinhardtii* with *CzPSY* gene amplified from the *Chlorella zofingiensis* genome containing its original 5' and 3' UTRs resulted in significantly enhanced carotenoid biosynthesis compared to the untransformed *C. reinhardtii* cell line (Cordero et al., 2011). In plants, chloroplast transformation very often involves plasmid construction with heterologous promoters which can result in very efficient transgene expression (Leelavathi and Reddy, 2003). On the other hand, heterologous promoters may result in decreased transgene expression compared to endogenous regulatory elements. Transformation of tobacco and lettuce with a construct containing a heterologous *psbA* promoter driving expression of the anthrax protective antigen (PA) or human proinsulin (Pins) fused with the cholera toxin B-subunit (CTB) was substantially less efficient compared to the endogenous *psbA* 5′ UTR (Ruhlman et al., 2010).

The experiment presented in Figure 4-10 demonstrated that the *C. sorokiniana* promoter/5'UTR is not functional in *C. reinhardtii*. Consequently, an attempt was made to test if the altered *aadA* cassette is incorporated into the chloroplast genome of *C. reinhardtii*. KRC1001-11A-Δ*psaA* strain containing partially deleted *psaA*-3 gene was used as a recipient strain for chloroplast transformation with pBa3-AX-psbX-*aadA* expression vector (Figure 4-11). This strain has non-functional PSI and additionally to being photosynthetically impaired it is also light sensitive, and should be rescued to phototrophy since the pBa3-AX-psbX-aadA plasmid carries a WT copy of *psaA-3*.

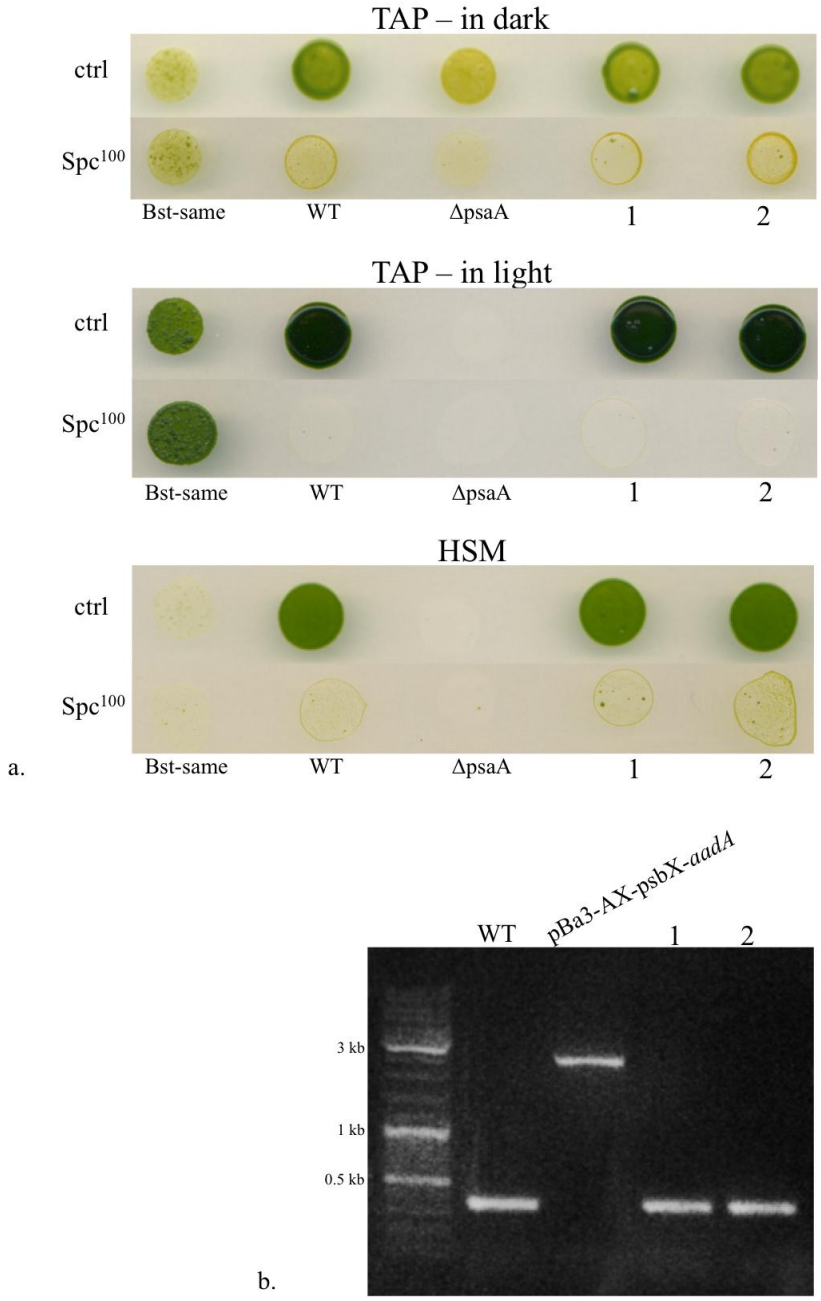


Figure 4-11 Analysis of two transformed cell lines of *C. reinhardtii* strain Δ*psaA* KRC1001-11A.

a. Spot test of two transformed cell lines of *Chlamydomonas reinhardtii* strain Δ*psaA* KRC1001-11A (1 and 2)
b. PCR analysis of two transformed cell lines of *Chlamydomonas reinhardtii* strain Δ*psaA* KRC1001-11A (1 and 2). Both isolates show wild type genotype with restored *psaA* gene and without *aadA* gene.

1 and 2 transformed cell lines of Δ*psaA* **KRC1001-11A** strain, **Bst-same** – a cell line resistant to spectinomycin with impaired photosynthetic activity, **WT** – wild type of *Chlamydomonas reinhardtii* and Δ*psaA* KRC1001-11A strain used as a recipient cell line for chloroplast transformation.
TAP – medium enriched in acetate, HSM – minimal medium; ctrl – no antibiotic, Spc^{100} – added spectinomycin at 100μg/ml.

The KRC1001-11A-Δ*psaA* recipient strain was subjected to glass bead transformation. After vortexing with pBa3-AX-psbX-*aadA*, the cells were plated on minimal medium under constant illumination. Approximately 2 weeks after the transformation two colonies were isolated and further tested for photosynthetic activity and spectinomycin resistance. Figure 4-11.a presents the growth tests performed on minimum (HSM) and acetate enriched (TAP) media with or without Spc. The isolated transformant had restored photosynthetic activity as they can be maintained on minimal medium yet they do not show resistance to spectinomycin, as expected from the previous transformation experiments.

Subsequently, PCR analysis of the transformants was carried out with the following primers: pBA3_flank.F [GTGATCGTTACACTTTAGATAACTGG] and pBA3_flank.R [GATTTTAAGTCCGTAGCGTCTACC] and the results are presented in Figure 4-11.b. Primers pBA3_flank.F + pBA3_flank.R targeted the flanking region around the *aadA* cassette. The PCR analysis verified the wild type genotype of the putative transformants where the *psaA* gene was restored and *aadA* cassette was absent.

The restoration of *psaA* gene and photosynthetic activity of the KRC1001-11A-Δ*psaA* recipient strain without the incorporation of the *aadA* cassette is surprising, but can be explained if there are additional sites suitable for homologous recombination present within the *aadA* cassette that resulted in integration of the *psaA* gene only without integration of the *aadA* construct.

4.3.1.9 Creation and analysis of photosynthesis-deficient strains of *C. sorokiniana* UTEX1230

Choosing the right selectable marker for *Chlorella sorokiniana* chloroplast metabolic engineering applications turned out to be challenging. Since this strain is impervious to most antibiotics, the majority of the available selective markers, for instance kanamycin or chloramphenicol resistance genes are not compatible with *Chlorella sorokiniana*. It appeared a combination of spectinomycin and streptomycin would sufficiently inhibit algal growth and the chosen *aadA* selectable marker would be optimal for the *C. sorokiniana* UTEX1230 chloroplast transformation. However, the number of spontaneous mutants was overwhelming and it was not possible to further test this marker. Therefore, a classical mutagenesis strategy was used to create a photosynthetic mutant that would

serve as a recipient cell line for chloroplast transformation where restoring photosynthetic activity would be an effective selectable marker. This idea was feasible, since it was shown previously that *C. sorokiniana* was capable of heterotrophic growth (Chapter 3) and therefore such photosynthetic mutants would be viable when maintained on an organic carbon source.

Methods for the isolation of photosynthetic mutants of *Chlamydomonas reinhardtii* are based on mutagenesis with 5-fluorodeoxyuridine (FdUrd) and use of metronidazole as a selective agent (see Material and Methods Chapter). FdUrd is known to reduce the copy number of chloroplast DNA but also increase frequency of mutations within chloroplast genes (Wurtz et al., 1979) whereas metronidazole treatment is effective in the enrichment of photosynthesis-deficient mutants (Schmidt et al., 1977). Following mutagenesis, approximately 3000 isolates were tested for photosynthetic activity. As a result, 27 photosynthetically impaired cell lines were isolated and subjected to video imaging analysis of relative fluorescence (Figure 4-12). The relative fluorescence of dark-adapted putative photosynthetic mutants was compared to *C. sorokiniana* UTEX1230 wild type and two additional *C. reinhardtii* control cell lines: PSI-deficient strain KRC1001-11A-Δ*psaA* (Δ*psaA*) and PSII-deficient strain Bst-same (Bst). The following isolates PSM315, PSM941, PSM1345, PSM1645 showed the highest relative fluorescence and therefore were selected for further analysis.

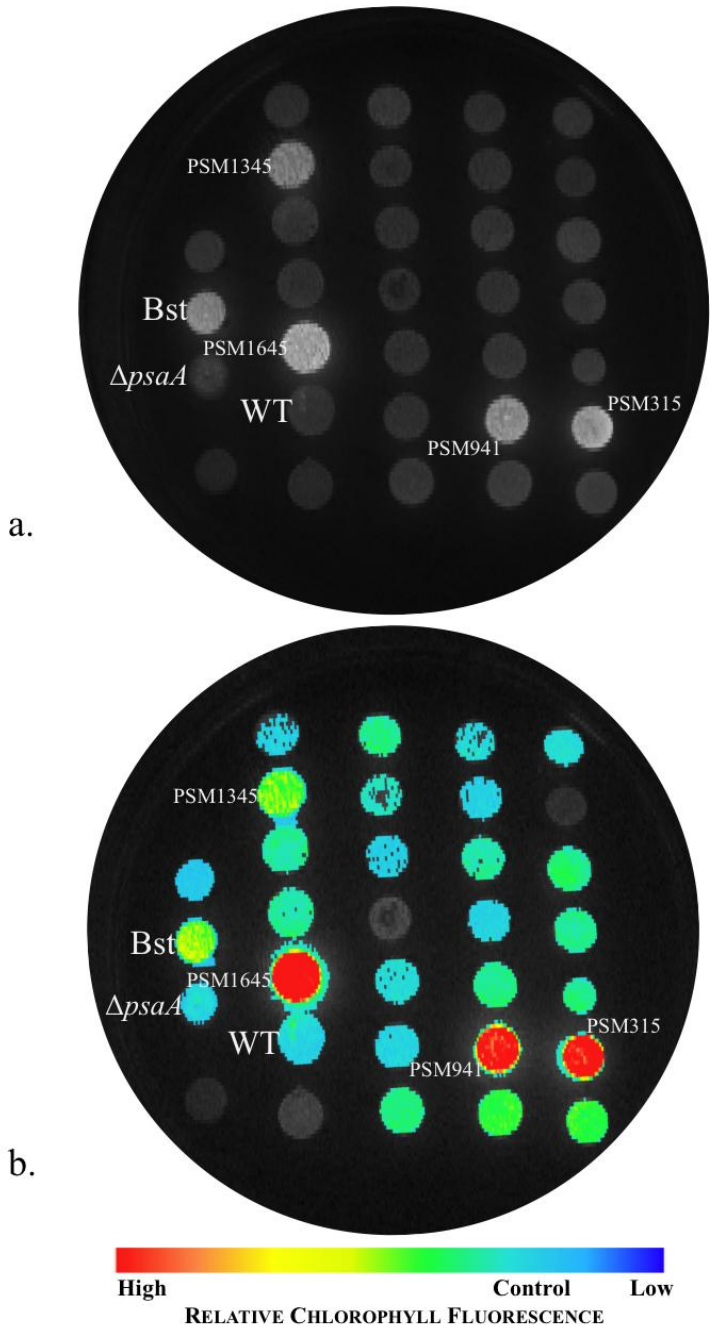


Figure 4-12 Video imaging analysis of relative fluorescence intensity of the putative photosynthetic mutants of *C. sorokiniana* UTEX1230.

Spot test of the 49 mutants of *C. sorokiniana* and controls maintained in dark for seven days on TAP plates.

a. Snapshot of the dark-adapted isolates after a high-intensity light flash
b. Relative chlorophyll fluorescence analysis of the dark-adapted isolates after a high-intensity light flash

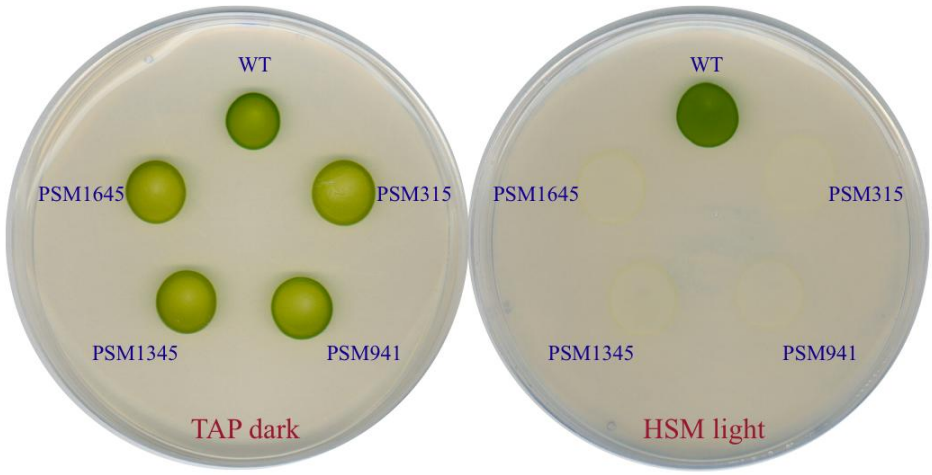
PSM315, PSM941, PSM1345, PSM1645 – putative photosynthetic mutants characterized by higher than wild type chlorophyll fluorescence level.
Bst – photosystem II deficient *Chlamydomonas reinhardtii* mutant, **ΔpsaA** – photosystem I deficient *C. reinhardtii* strain (KRC1001-11A-Δ*psaA*), **WT** – *C. sorokiniana* UTEX1230 wild type

Figure 4-13.a presents growth analysis of the isolated PSM cell lines with *C. sorokiniana* wild type used as control. All four PSM isolates are photosynthetic mutants as they fail to grow on minimum medium. On the other hand their heterotrophic growth in the dark is undisrupted. Wild type *C. sorokiniana* was used here as a photosynthetically active positive control.

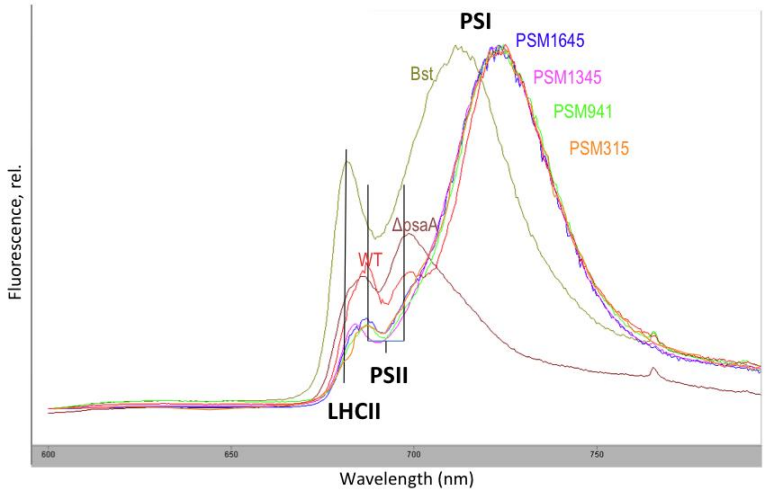
77K fluorescence analysis is specifically designed to detect the difference in fluorescence emission spectra between photosynthetic mutants and wild type of photosynthetic organisms. Results of the 77K fluorescence analysis are presented in Figure 4-13.b. Two *C. reinhardtii* photosynthetic mutants were applied as controls – Bst-same as a photosystem II (PSII) mutant and KRC1001-11A-Δ*psaA* (Δ*psaA*) – photosystem I (PSI) mutant. The fluorescence spectra for PSI and PSII are shifted in *Chlorella sorokiniana* compared to *Chlamydomonas reinhardtii* yet the pattern of the spectra remains the same. The PSII peak is significantly reduced in all PSM cell lines compared to wild type, which indicates that all PSM isolates are PSII mutants.

PSI and PSII are formed by numerous polypeptide subunits into complexes that are not correctly assembled unless all the subunits of each complex are fully synthesized (Larkum et al., 2003). Therefore, western blot analysis was performed to detect the PSI and PSII complexes with the PsaD protein being a component of PSI and the D1 protein a component of PSII. The results of the western blot analysis are presented in Figure 4-13.c. Unfortunately, the western blot data could not confirm the isolated mutants were PSII mutants as the D1 protein was present in all PSM lines.

a.



b.



c.

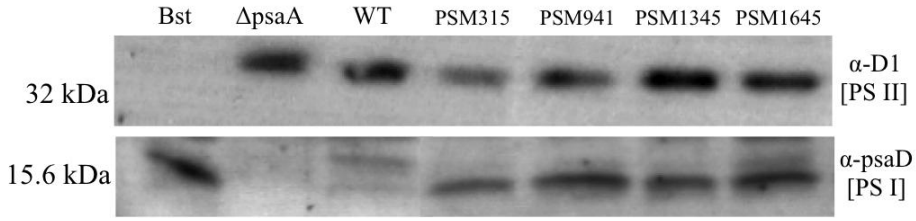


Figure 4-13 Analysis of the putative chloroplast mutants of *C. sorokiniana* UTEX1230.

a. Photosynthetic growth analysis of the PSM isolates and WT on minimum medium HSM; TAP in dark used as a control
b. 77K fluorescence emission spectra of the PSM isolates compared to the controls. Peaks specific for light harvesting complex II (LHCII), PSI and PSII were marked; excitation wavelength was 436 nm
c. Western blot analysis of the PSM isolates aiming to detect D1 and psaD proteins that are part of PSII and PSI complexes respectively

PSM315, PSM941, PSM1345, PSM1645 – putative photosynthetic mutants
Bst – photosystem II (PSII) depleted *Chlamydomonas reinhardtii* mutant
ΔpsaA – photosystem I (PSI) depleted *C. reinhardtii* mutant
WT – *C. sorokiniana* UTEX1230 wild type

In an early report, FdUrd treatment was successfully used to create a herbicide-resistant mutant of *C. reinhardtii* (Erickson et al., 1984). In this report, the induced single T-A to G-C transversion within the *psbA* chloroplast gene resulted in a serine to alanine change within the 32-kDa protein of PSII. This mutation resulted in resistance to 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) and 2-chloro-4-ethylamino-6-isopropylamino-*s*-triazine (atrazine) and therefore allowed the modified *psbA* to be used as a selectable marker. However, this particular change within *psbA*, is not applicable for *C. sorokiniana* UTEX1230 since this strain is resistant to both herbicides DCMU and atrazine (data not shown). Numerous other mutations within *psbA* have been associated with resistance to many herbicides such as triazines, triazinones, ureas, biscarbamates, phenols p-benzo- and naphthoquinones, acridones, NH-thiazol, and phenmedipham (Johanningmeier et al., 2000, Wilski et al., 2006). Therefore, a possibility of using an herbicide resistant marker created by mutation within a chloroplast gene is still plausible for *C. sorokiniana.*

4.3.2 Development of nuclear transformation method for *Chlorella sorokiniana* UTEX1230

In parallel to developing a method for chloroplast transformation, I attempted to create a nuclear transformation method for *C. sorokiniana* UTEX1230. As demonstrated earlier in this chapter, this strain is resistant to many antibiotics and herbicides. Hence, choosing an appropriate and effective selectable marker is a key consideration.

4.3.2.1 Isolation of the cycloheximide-resistant strains.

The most promising selectable marker for *Chlorella sorokiniana* UTEX1230 is an endogenous nuclear gene carrying a dominant mutation resulting in a resistance to a robust antibiotic. Consequently, the chosen gene was that coding for the 60S ribosomal protein L41, since mutations in L41 are known to confer resistance to the antibiotic cycloheximide (CYH) that is an effective inhibitor of cytoplasmic translation in eukaryotes. It has been reported in yeast and in algae that a single change of a specific amino acid where a conserved proline is replaced by glutamine or another residue is responsible for resistance to the antibiotic (Yagisawa et al., 2004, Mutoh et al., 1998, Stevens et al., 2001, Kawai et al., 1992).

Following the plating of ~10^{10} *Chlorella sorokiniana* cells on plates containing CYH at 2 μg/ml, seven resistant mutants emerged because of spontaneous mutation. Figure 4-14 presents growth of the resistant (CYH^R) colonies on TAP medium containing various CYH concentrations ranging from 2 μg/ml to 15 μg/ml. Colony 5 appears to show the highest resistance to this antibiotic since its growth is evident on plates containing up to 15 μg/ml CYH. The minimum inhibitory concentration of CYH for colonies 2, 3, 6 and 7 is approximately 2 μg/ml whereas for colonies 1 and 4 this concentration is approximately 5 μg/ml.

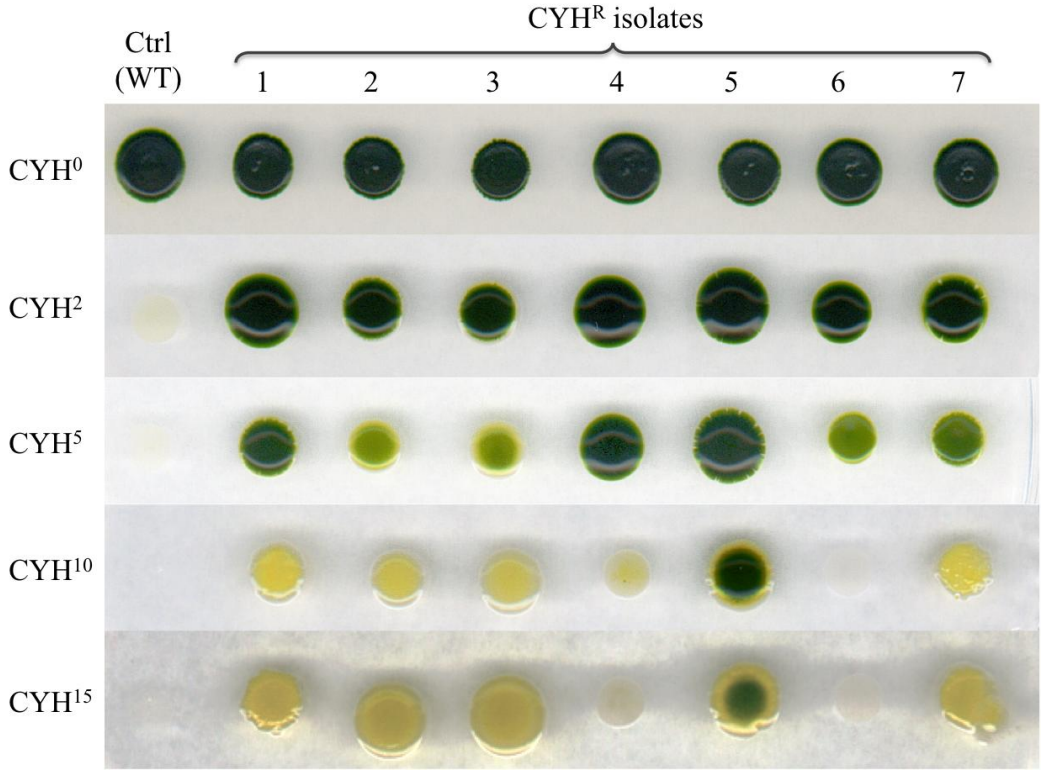


Figure 4-14 Cycloheximide resistance analysis of the cycloheximide-resistant (CYH^R) mutants of *C. sorokiniana* UTEX1230.

CYH^R isolates and wild type (WT) were maintained in TAP (CYH^0) and TAP with four concentrations of cycloheximide – 2 μg/ml (CYH^2), 5 μg/ml (CYH^5), 10 μg/ml (CYH^{10}) or 15 μg/ml (CYH^{15}). The antibiotic resistance was examined after 7 days.

4.3.2.2 Analysis of the isolated putative cycloheximide-resistant strains

The isolated CYH^R colonies were cultured and their DNA extracted (see chapter 2 for reference). To characterise the *RPL41* sequence of the CYH^R mutants compared to wild type, two primers RPL41.F [AAGGTCACACAGTACAAGACCGG] and RPL41.R [CGGAGCACAATCTTCTTGGTGG] were designed based on the *RPL41* sequence from *Chlorella* sp. NC64A [http://genome.jgi.doe.gov/genome-projects/] to amplify an approximately 750bp fragment of the *RPL41* gene containing the expected site of mutation. The PCR was successful and the products were sequenced. The amino acid sequences of *RPL41* of *Chlamydomonas* sp. and *Chlorella* sp. were compared (data not shown) and the protein sequence of the mutants 1-7 and the wild type strain was determined. Figure 4-15 shows this analysis and reveals that the conserved proline is indeed changed through mis-sense mutations to other amino acid residues. The proline to leucine (P>L) mutation displayed the strongest resistance to CYH as seen in colony 5 (Figure 4-14). Colonies 2, 3, 6, 7 resist 2 μg/ml of CYH and this feature is represented by a proline to threonine (P>T) change, whereas colonies 1 and 4 have a proline to serine (P>S) change and are able to grow at CYH concentration up to 5 μg/ml.

The seven isolated cycloheximide-resistant mutants of *Chlorella sorokiniana* possessed the same mutations (P>L and P>S) of RPL41 as that reported in *C. reinhardtii* (Stevens et al., 2001) however, an additional type of mutation P>T has not been previously reported in algae. Nevertheless, this change is less significant compared to the other two, as it is associated with the weakest resistance to cycloheximide.

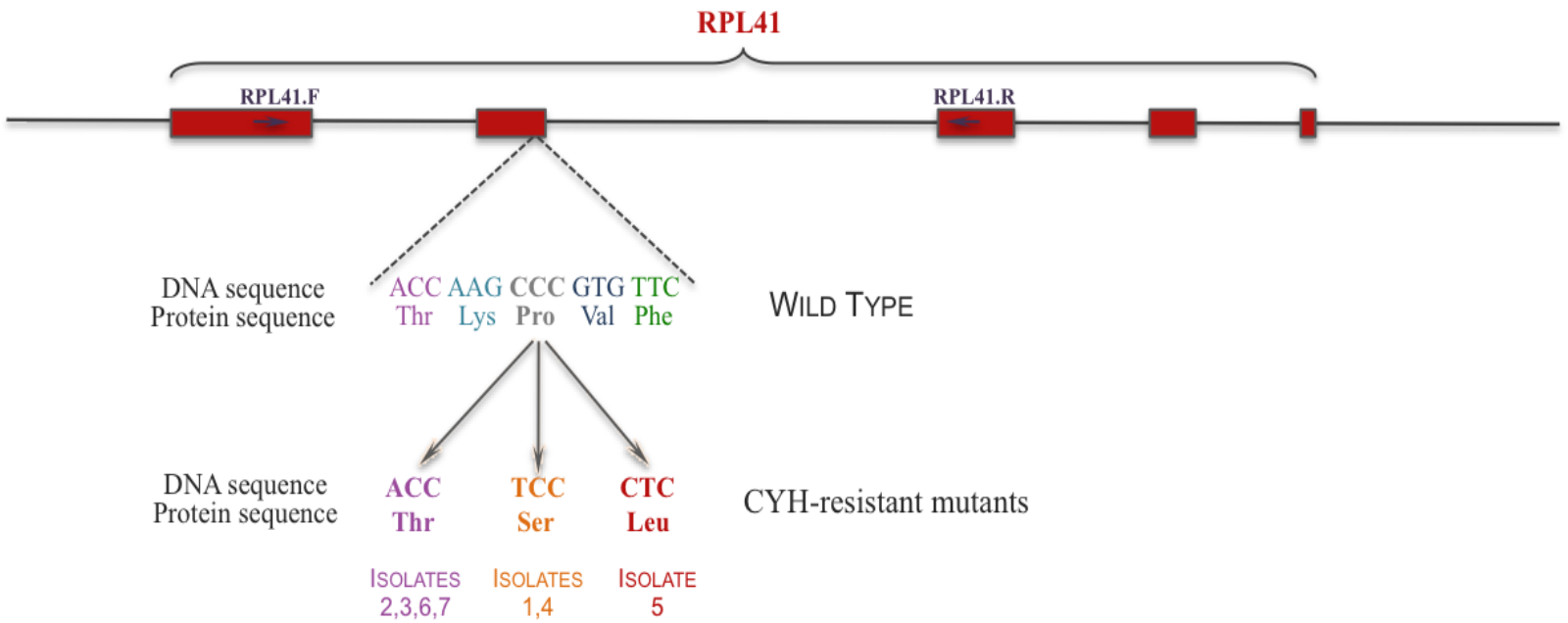


Figure 4-15 Analysis of the genotypes of the cycloheximide-resistant mutants of *C. sorokiniana* UTEX1230.

Proline in the position 56 within the amino acid sequence of *RPL41* in wild type (WT) *Chlorella sorokiniana* is replaced with threonine, serine or leucine and this change is related to a single point mutation where the CCC codon is changed to ACC, TCC or CTC, respectively. All of the mutations correlate with the observed cycloheximide resistant phenotype.

In order to use the P>L variant of the *RPL41* gene as a selectable marker, an attempt was made to amplify the complete gene including the promoter and UTRs. Several PCR-based methods were employed to obtain the full sequence of *RPL41*. An overview of the various amplification strategy attempted is presented in Figure 4-17 whereas the list of oligonucleotides is included in Table 4-3.

As described earlier, a fragment of the *RPL41* gene containing the mutation was successfully amplified using the RPL41.F and RPL41.R oligonucleotides. An inverse PCR method was developed according to the protocol previously described (Ochman et al., 1988) to amplify the 5' and 3' unknown ends of the *RPL41* gene. Firstly, a set of oligonucleotides was designed (41_300_inv.F + 41_600_inv.R) to amplify approximately 300 bp fragment of known region of *RPL41* gene and as a result, the PCR product of this reaction was purified and used as a probe for Southern blot analysis. Southern blot analysis (Figure 4-16) was designed to determine the expected sizes of the products obtained during the inverse PCR amplification. 1 µg of genomic DNA of *C. sorokiniana* was digested with a selection of restriction enzymes (*Pst*I, *Sac*I, *Nco*I, *Hinc*II, *Apa*LI) that cut the known *RPL41* sequence once. Genomic DNA digested with *Apa*LI was chosen for further analysis and this digested DNA was self-ligated as described in section 2.6.3.3 to provide a template for inverse PCR (Ochman et al., 1988). In order to amplify the 5' end of the gene, two sets of oligonucleotides were used: 41_INV_bgn.R and 41_300_INV.F whereas for the 3' end 41_INV_end.F and 41_600_INV.R were applied. Both reactions failed. However, using a second combination of primers RPL41.F + 41_INV_end_2.F on the self-ligated gDNA as a template, resulted in amplification of approximately 500 bp fragment from the 3' end of the gene. A second method attempted to amplify the ends of the gene was thermal asymmetric interlaced PCR method (TAIL-PCR) (Dent et al., 2005, Liu and Chen, 2007, Wu-Scharf et al., 2000).

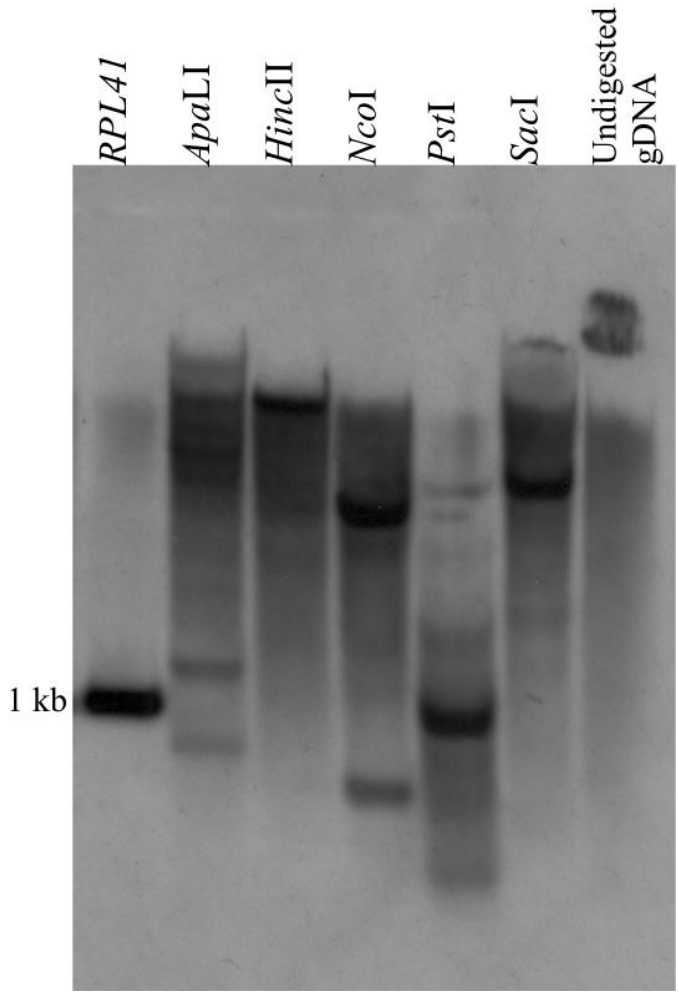


Figure 4-16 Southern blot analysis for the inverse PCR assay.

Genomic DNA (gDNA) was digested with five restriction enzymes: *Pst*I, *Sac*I, *Nco*I, *Hinc*II, and *Apa*LI. *RPL41* functioned as a positive control (PCR product of approximately 1 kb created after amplification with RPL41.F and RPL41.R oligonucleotides) whereas undigested gDNA functioned as a negative control.

The restricted DNAs were electrophoresed, transferred to nylon membrane, hybridized with DIG-labeled probe (PCR product of amplification with 41_300_inv.F + 41_600_inv.R oligonucleotides) and the exposed to the DIG signal Hyperfilm was subsequently developed in automatic film processor.

The protocol was adapted from (Dent et al., 2005) with minor changes in thermal profile of the two PCR reactions that are presented in Table 4-2. The following forward primers were designed: 1_TAIL_PCR.5R for 5' end and 1_TAIL_PCR.3R for the 3' end were used in the primary PCR reaction whereas in the secondary (nested) PCR reaction 2_TAIL_PCR.5R for the 5' end and 2_TAIL_PCR.3R for the 3' end were employed. Two arbitrary degenerate primers RMD227 and RMD228 (Wu-Scharf et al., 2000) were used as reverse primers in both PCR reactions. This method allowed the amplification of the remaining 3' coding sequence and 3'UTR of the gene.

In order to amplify the remaining 5' end of the gene, the RMD227 and RMD228 degenerate primers were replaced with AD7 and AD8 (Mahalingam and Fedoroff, 2001) in the primary and secondary TAIL PCR reactions. Again, the method failed to amplify the remaining *RPL41* sequence. Subsequently, a genome walking method was employed (Reddy et al., 2008). Oligonucleotides 1_TAIL_PCR.5R and 2_TAIL_PCR.5R were used as reverse primers annealing to the known region within the *RPL41* gene. Similarly, the method failed to produce a product. The method of rapid amplification of complementary DNA 5'end (5'RACE) was then tried (Zhang and Frohman, 1997, Zhang and Frohman, 1998, Rapley, 2000). In this method, a mix of four oligonucleotides 5'polyC_A.F, 5'polyC_T.F, 5'polyC_C.F and 5'polyC_G.F where employed together with 5'RACE_known.R. Likewise, the method failed to amplify *RPL41* sequence. In the last attempt, a Splinkerette method was attempted according to the protocol (Potter and Luo, 2010) with minor modifications. In this method, gDNA was digested with *Sac*I and a double-stranded splinkerette oligonucleotide containing a hairpin loop and matching sticky ends (SPLNK_TOP + SPLNK_TCGA_ BOT) was ligated to the digested gDNA. For the primary PCR reaction the SPLNK#1 and 1_SPLNK_41.R oligonucleotides were used whereas the secondary (nested) PCR amplification was performed using the SPLNK#2 and 2_SPLNK_41.R oligonucleotides.

Table 4-2 Thermal profile of the TAIL PCR reactions.

Adapted from (Dent et al., 2005)

REACTION	STEP	THERMAL SETTINGS	NO. OF CYCLES
Primary	1	98°C, 2 min	1
	2	98°C, 1 min; 62°C, 1 min; 72°C, 2.5 min	6
	3	98°C, 1 min; 25°C, 3 min; ramping to 72°C over 3 min; 72°C, 2.5 min	1
	4	98°C, 30 s; 68°C, 1 min; 72°C, 2.5 min; 948°C, 30 s; 68°C, 1 min; 72°C, 2.5 min; 98°C, 30 s; 44°C, 1 min; 72°C, 2.5 min	16
	5	72°C, 5 min	1
Secondary	1	98°C, 30 s; 64°C, 1 min; 72°C, 2.5 min; 98°C, 30 s; 64°C, 1 min; 72°C, 2.5 min; 98°C, 30 s; 44°C, 1 min; 72°C, 2.5 min	13
	2	72°C, 5 min	1

Table 4-3 List of oligonucleotides used in all applications employed in order to amplify the *RPL41* gene

Name	Sequence	Application	Comment	Reference
RPL41.F	AAGGTCACACAGTACAAGACCGG	Initial step		This work
RPL41.R	CGGAGCACAATCTTCTTGGTGG	Initial step		This work
41_300_inv.F	CAGCACCAGGATGTACTTCAAGG	Southern blot probe for inverse PCR		This work
41_600_inv.R	GCTGTAAACATATGGCACATTGCTGC	Southern blot probe for inverse PCR		This work
41_INV_bgn.R	GCAGCAGAGGCAAGAGCAGC	Inverse PCR	Failed	This work
41_INV_end.F	GCTCACTGACATGTTCCATCTCC	Inverse PCR	Failed	This work
41_INV_end_2.F	GGCAAATGCAGCTTAATCACAACC	Inverse PCR		This work
RMD227	NTCGWGWTSCNAGC	3' and 5' TAIL		(Wu-Scharf et al., 2000)
RMD228	WGNTCWGNCANGCG	3' and 5' TAIL		(Wu-Scharf et al., 2000)
1_TAIL_PCR.5R	AACACGGGCTTGGTCTGG	5'TAIL 5'genome walking	Failed Failed	This work
2_TAIL_PCR.5R	ACTGCTTGCGGTCGTAACG	5'TAIL 5'genome walking	Failed Failed	This work
1_TAIL_PCR.3F	CACCAAGAAGATTGTGCTGCG	3'TAIL		This work
2_TAIL_PCR.3F	CACATGCACGCCATCAAGG	3'TAIL		This work
AD7	NTCGASTWTSGWGTT	5'TAIL	Failed	(Mahalingam and Fedoroff, 2001)
AD8	NGTCGASWGANAWGAA	5'TAIL	Failed	(Mahalingam and Fedoroff, 2001)
Walker adaptor 1	GTGAGCGCGCGTAATACGACTCACTA TAGGGNNNNATGC	5'genome walking	Failed	(Reddy et al., 2008)
Walker adaptor 2	GTGAGCGCGCGTAATACGACTCACTA TAGGGNNNNGATC	5'genome walking	Failed	(Reddy et al., 2008)
Walker adaptor 3	GTGAGCGCGCGTAATACGACTCACTA TAGGGNNNNTAGC	5'genome walking	Failed	(Reddy et al., 2008)
Walker adaptor 4	GTGAGCGCGCGTAATACGACTCACTA TAGGGNNNNCTAG	5'genome walking	Failed	(Reddy et al., 2008)
Walker primer 1	GTGAGCGCGCGTAATACGA	5'genome walking	Failed	(Reddy et al., 2008)
Walker primer 2	GTAATACGACTCACTATAGGG	5'genome walking	Failed	(Reddy et al., 2008)
5'RACE_known.R	GAACACGGGCTTGGTCTGG	5'RACE	Failed	This work
5'polyC_A.F	CCCCCCCCCCCCCCDA	5'RACE	Failed	This work
5'polyC_T.F	CCCCCCCCCCCCCCDT	5'RACE	Failed	This work
5'polyC_C.F	CCCCCCCCCCCCCCDC	5'RACE	Failed	This work
5'polyC_G.F	CCCCCCCCCCCCCCDG	5'RACE	Failed	This work
SPLNK_TOP	CCACTAGTGTCGACACCAGTCTCTAA TTTTTTTTTTCAAAAAAA	Splinkerette	Failed	This work, adapted from (Potter and Luo, 2010)
SPLNK_TCGA_ BOT	CGAAGAGTAACCGTTGCTAGGAGAG ACCGTGGCTGAATGAGACTGGTGTCG ACACTAGTGGAGCT	Splinkerette	Failed	This work, adapted from (Potter and Luo, 2010)
SPLNK#1	CGAAGAGTAACCGTTGCTAGGAGAG ACC	Splinkerette	Failed	(Potter and Luo, 2010)
SPLNK#2	GTGGCTGAATGAGACTGGTGTCGA C	Splinkerette	Failed	(Potter and Luo, 2010)
1_SPLNK_41.R	GCTGCGACTGTTGTCGTCCATGG	Splinkerette	Failed	This work
2_SPLNK_41.R	ATGTCACTTCGCCCAACTGCTGTGC	Splinkerette	Failed	This work

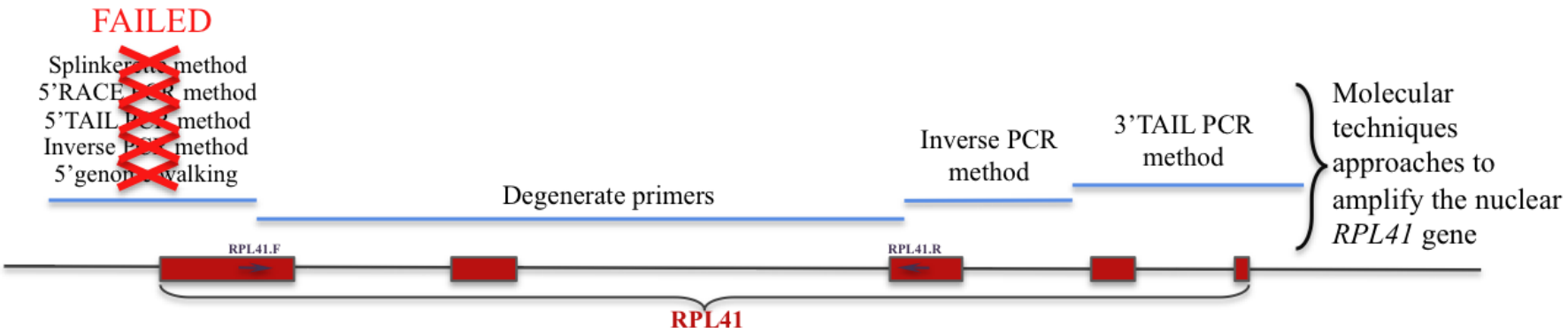


Figure 4-17 Molecular technique approaches to amplify the *RPL41* nuclear gene in *C. sorokiniana* UTEX1230.

RPL41.F + RPL41.R – degenerate primers used in the first method of *RPL41* gene amplification
RACE method – rapid amplification of cDNA ends
TAIL PCR method - thermal asymmetric interlaced polymerase chain reaction
The first part of the nuclear *RPL41* gene was amplified using the RPL41.F and RPL41.R primers. Inverse PCR method allowed for amplification of further ~0.5 kb 3'end whereas 3'TAIL PCR method enabled to sequence the remaining part of 3'end and 3'UTR fragment. I was unable to amplify the remaining 5'end of the *RPL41* gene with any of the available methods.

The full sequence of the amplified *RPL41* gene is presented in the appendix G.

In general, the selection of suitable selectable markers for nuclear transformation in algae such as *Chlorella* is insufficient. The vast majority of the dominant mutations to antibiotics occur within ribosomal genes in algae (Nelson et al., 1994, Stevens et al., 2001). In the presented work, the dominant mutants within *RPL41* that confers resistance to cycloheximide were successfully isolated. The dominant mutations incorporate leucine, threonine or serine in place of proline as seen in Figure 4-15 and the type of mutation is also reflected in the resistance level to cycloheximide with leucine giving the highest and threonine the lowest tolerance to CYH. As a spontaneous mutation naturally occurring within a specific position of *RPL41* would be rare (McCusker and Haber, 1988), the background of the false-positive following transformation would be expected to be low giving this selectable marker an important advantage over the *aadA* marker used for chloroplast transformation.

4.4 Conclusions and future work

The development of a chloroplast transformation method for *C. sorokiniana* was complicated by the need to isolate sections of chloroplast DNA for use as flanking sequence and a promoter/5'UTR in the design of the transformation vector. It proved fortunate that such sections were isolated by PCR since orientation of genes turned out to be rearranged in *C. sorokiniana* compared to the related microalga *Chlorella vulgaris* (Wakasugi et al., 1997). Difficulties during the development of a chloroplast transformation method were not limited to molecular aspects. Firstly, an appropriate DNA delivery method is needed since biolistics might not be approporiate given the very small size of the *C. sorokiniana* cell (2-3 µm in diameter). Electroporation has been demonstrated to be successful for nuclear transformation of *Chlorella* (Wang et al., 2007) and this might prove successful for chloroplast transformation. Optimisation of this transformation method would be required, as *C. sorokiniana* has a very robust cell wall (Muller et al., 2005, Takeda, 1993).

As discussed earlier, choosing the right promoter is crucial for the development of an efficient transformation vector. Several unsuccessful transformation attempts with an expression vector containing the *Chlamydomonas atpA* promoter/5'UTR resulted in a change in the design of the selectable marker. The *Chlamydomonas atpA* promoter/5'UTR was replaced with the *psbA* promoter/5'UTR region from *Chlorella sorokiniana*. Unfortunately, it was not possible to confirm that the cloned *psbA*

promoter/5'UTR was functional in either *C. sorokiniana* or *C. reinhardtii* although this element was shown to drive *aadA* expression in *E. coli*.

An additional issue that became apparent during the research using the *aadA* marker was the very high level of spontaneous resistance seen in *C. sorokiniana*, even when selection on both Spc and Str was employed. This suggested that *aadA* would not be usable as a marker, and other selection strategies should be considered. As a result, an attempt was made to isolate a photosynthesis-deficient mutant containing a mutation within any chloroplast gene essential for photosynthesis. Such a mutant would be potentially ideal recipient cell line for chloroplast transformation with restoration of photosynthetic activity as the selection. Having four photosynthetic mutants already isolated, the next step would be to locate the mutations within their chloroplast genome and to create a suitable transformation vector.

For the development of a nuclear transformation method for *C. sorokiniana*, cycloheximide resistance emerged as a very effective selection method as even low concentration of this antibiotic completely inhibits algal growth. Furthermore, mutants with specific mis-sense mutations within the endogenous *RPL41* gene were recovered and shown to confer resistance to the drug. Unfortunately, none of the five tried methods for isolating the 5' end of *RPL41* were successful. However, introducing modifications in RNA extraction procedure for the 5'RACE method (Hoang et al., 1995) or optimizing the primers and thermal profile for the TAIL-PCR method (Liu and Chen, 2007) could potentially allow the recovery of the 5' region and thus the cloning of the whole gene. When this is achieved, the next step in the development of the nuclear transformation method would be the vector construction and choosing a suitable DNA delivery method.

CHAPTER 5 *CHLAMYDOMONAS REINHARDTII* AS A PLATFORM FOR BIODIESEL PRODUCTION

5.1 Introduction

Microalgae are attractive as a source of biofuel, since they are characterized by very high growth rates and oil content, can be grown on uninhabitable land and potentially utilize waste or salt water (Scott et al., 2010). However, current technology in the commercial production of algae-based biofuels is not sufficiently advanced and several practical aspects have to be controlled for biofuels to be economically viable. The main issues include ensuring the consistent biomass production in outdoor cultivation systems, controlling contamination of algal cultures with unwanted species and also developing efficient and low-cost methods for harvesting the algal biomass and extracting oil.

Biological research on competitive biofuel production from microalgae is currently focused on two aspects. The first group of projects revolves around obtaining the maximum productivity from isolated algal strains and this part was thoroughly discussed in Chapter 3. Additional aspects of this approach include technical strategies of decreasing the cost of biofuel extraction from algal cells. The second approach involves genetic engineering through mutagenesis or introduction of transgenes in order to improve the biofuel production in algae and this topic is discussed in depth in the current chapter.

The successful genetic engineering approach is predominantly associated with access to a genome database, the development of suitable marker and expression vectors and the availability of an efficient gene delivery method. To date, 20-30 algal species have been genetically modified (see chapter 4 for reference), however *Chlamydomonas reinhardtii* remains the model organism for the majority of applications of algal genetic engineering.

Lipid metabolism pathways have been thoroughly analyzed in plants, yet in algae this process is still poorly understood. Numerous genes involved in lipid metabolism in plants have homologs in algal genomes (Guschina and Harwood, 2006), therefore applying genetic engineering strategies developed for plants in algae could result in increased lipid productivity. Figure 5-1 presents a schematic illustration of the metabolic pathways of algal lipid biosynthesis occurring in chloroplasts, adapted from (Radakovits et al., 2010).

The first approach in attempts to increase lipid content in algal cells is associated with the lipid productivity. Since there is a direct correlation between fatty acid production and lipid content in plants (Ohlrogge and Jaworski, 1997), the up-regulation of enzymes involved in free fatty acid synthesis could potentially result in increased lipid productivity (Roessler et al., 1994). One of the key enzymes involved in the early stage of fatty acid synthesis is ACCase (acetyl-CoA carboxylase) catalyzing the conversion of acetyl-CoA to malonyl-CoA. It has been shown that down-regulation of ACCase in *Brassica napus* resulted in a reduced lipid content of seeds (Sellwood et al., 2000). Several attempts to over-express ACCase in photosynthetic eukaryotes have been recorded (Klaus et al., 2004, Madoka et al., 2002, Roesler et al., 1997). Interestingly, lipid production was significantly increased in plants that are naturally rich in starch, such as *Solanum tuberosum*, when transformed with an acetyl-CoA carboxylase gene (Klaus et al., 2004), whereas up-regulation of ACCase in oleaginous plants such as *Brassica napus* did not increase their lipid content (Roesler et al., 1997). KAS (3-ketoacyl-ACP synthase) is an alternative key enzyme in fatty acid synthesis catalyzing the conversion of malonyl-ACP to 3-ketoacyl-ACP using the acyl-carrier protein (ACP) as a carbon source. The reports on transformation of tobacco or rapeseed and Arabidopsis with the transgenic 3-ketoacyl-ACP synthase III (KAS III) isolated from spinach did not however result in increased lipid content of the transformed plants (Dehesh et al., 2001).

Alternatively, in order to increase the lipid production in algal cells maintained in photobioreactors under constant illumination or with an organic carbon source, the reduction of the rate of lipid metabolism could improve its lipid productivity. In *Saccharomyces cerevisiae* the disruption of the fatty acyl-CoA synthetase gene resulted in increased extracellular secretion of free fatty acids (Michinaka et al., 2003). This idea was tested in *Arabidopsis thaliana* seedlings, where the activity of another key enzyme in fatty acid β-oxidation, 3-ketoacyl-CoA thiolase (KAT), was suppressed (Germain et al., 2001). As a result, there was no degradation of lipid bodies during germination of the KAT-deficient mutants, the seedlings however failed to grow immediately after their emergence.

Figure 5-1 Schematic illustration of the fatty acids biosynthesis pathways occurring in the microalgal chloroplast.

Adapted from (Radakovits et al., 2010)

ACP, acyl carrier protein; CoA, coenzyme A.

A third strategy for genetic engineering of algal lipid profile is altering the length of fatty acid chains. On average, the length of fatty acid chain in algal storage lipids varies between C_{14} and C_{20} and in the vast majority of algal strains lipids predominantly consist of C_{16} and C_{18} (Basova, 2005, Caramujo et al., 2008, Sushchik et al., 2003, Yamasaki et al., 2012), whereas the optimal length for biodiesel production is $C_{12} - C_{14}$. Fatty acyl-ACP thioesterases are the enzymes controlling the carbon chain length during the fatty acid synthesis (Jones et al., 1995), therefore overexpression of certain types of thioesterases could potentially alter the fatty acid composition in algae. This approach is discussed in depth in the current chapter.

The final strategy involves a biodiesel synthesis that does not include an alteration of the generated lipids. Lipids produced by algae cannot be utilized directly as biodiesel but have to be transformed to fatty acid methyl esters (see chapter 3 for details) and this process requires an additional energy supply. However, a novel idea has been proposed to create an optimal environment suitable for a direct transesterification of fatty acids within the transformed cell (Kalscheuer et al., 2006). In this report, *E. coli* was transformed with three transgenes: pyruvate decarboxylase and alcohol dehydrogenase isolated from *Zymomonas mobilis* and acyltransferase from *Acinetobacter baylyi*. This modification of *E. coli* metabolism resulted in ethanol synthesis within the bacterial cell, that was subsequently utilized during the esterification of coenzyme A thioesters of fatty acids, where bacteria were cultured in aerobic conditions on medium supplemented with glucose and oleic acid.

An alternative to the direct biodiesel synthesis approach is the synthesis of alkanes in algal cells. *Botryococcus braunii* has been of great interest for decades because of its unusual ability to synthesize high quantities of long-chain unsaturated hydrocarbons ranging from 15-75% dry weight depending on the strain and growth conditions (Brown et al., 1969, Metzger et al., 1985, Sim et al., 2001, Metzger and Largeau, 2005, Rao et al., 2007). Owing to its exceptionally high hydrocarbon content, *B. braunii* for many years was regarded as a key species in renewable fuel production. However, the major obstacle in its further exploitation was the very long generation time that varies between 3-6 days (Ramirez and Corbacho, 2005, Volova et al., 1998) and a very complex and long-chain hydrocarbon content ranging from C_{23} to C_{37}, that require further processing in order to be used as a biofuel (Metzger et al., 1985, Metzger and Largeau, 2005, Metzger et al., 2007). Therefore, one of the strategies is to generate a fast growing algal strain producing

medium-chain alkanes that can be directly utilized as a fuel source and this approach is discussed in detail in the current chapter.

5.2 Background and aims of the project

In principle, alkanes (saturated hydrocarbons) can be produced from fatty acids (FA) by a two-step process where the FA is first transformed to the corresponding aldehyde by an acyl reductase, and then this aldehyde is converted to the alkane by an aldehyde decarbonylase (Park, 2005). The schematic illustration of this proposed pathway is presented in Figure 5-2. The idea that natural alkanes found in plants and animals are derived from FA through a decarbonylation process has been accepted for many years now and was confirmed by the fact that the majority of alkanes are odd-carbon numbers whereas fatty acids in plants and algae are predominantly even-carbon numbers (Cheesbrough and Kolattukudy, 1988, Cheesbrough and Kolattukudy, 1984, Dennis and Kolattukudy, 1992, Dennis and Kolattukudy, 1991).

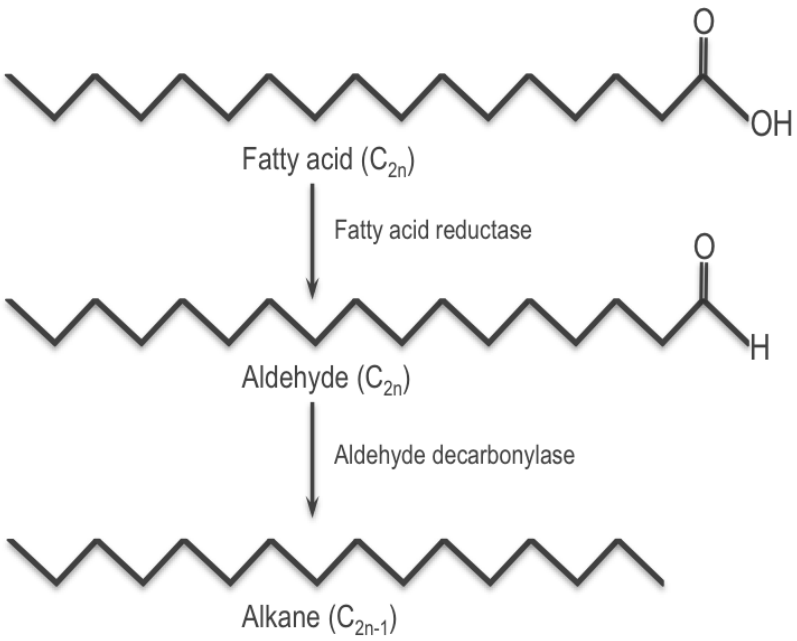


Figure 5-2 Schematic illustration of the reactions involved in the proposed alkane biosynthesis pathway from fatty acids.

The initial information on the possible alkane synthesis pathway in eucaryotic organisms originated from the proposed alkane metabolic pathway in bacteria. Activation of alkanes occurs via oxidation of one of the terminal methyl groups to the corresponding alcohol and this process is catalyzed usually by an alkane hydrolase complex. The alcohol is further oxidized to the corresponding fatty acids by dehydrogenases (Rojo, 2005).

The breakthrough in research on alkane biosynthesis in living organisms came with the report of Schirmer et al. (2010) where the acyl-acyl carrier protein reductase and aldehyde decarbonylase genes isolated from cyanobacteria were successfully integrated into the *E. coli* genome. This resulted in alkane and alkene production and secretion in the transformed bacterial cell line. Therefore, the ultimate aim of the research described in this chapter was to reproduce this finding in algae. The chloroplast organelle of plants and algae is the site of fatty acid biosynthesis, hence the chloroplast genome appears to be an appropriate site for insertion of these genes. The Purton lab has developed a suitable expression system for chloroplast transformation in *Chlamydomonas* (Bateman and Purton, 2000), and in this project the attempts to create transgenic strains expressing both genes will be presented and their alkane profile examined using gas chromatography analysis.

An alternative pathway for alkane synthesis was reported in *Vibrio furnissii* (Park, 2005), where the hexadecanoic acid was a substrate for the pentadecane and hexadecane production. In this report, free FA was transformed to the corresponding aldehyde and was subsequently converted to an alkane. FA synthesis in the chloroplast (Figure 5-1) does not result in free FA, but rather an acyl-ACP – a fatty acid linked by a thioester bond to a derivative of Coenzyme A (CoA) acyl carrier protein. Since a free fatty acid is the substrate in the proposed pathway in *Vibrio furnissii*, overexpression of a specific thioesterase should result in increased substrate concentration for alkane production. Therefore, the aim in this part of the project was to test if overexpression of thioesterase enzyme affects the free fatty acid pool of the transformed algal strain.

5.3 Results and Discussion

5.3.1 Transformation of *C. reinhardtii* chloroplast with synthetic *FatB1* gene

In order to increase the free fatty acid pool in chloroplasts, a specific thioesterase that would target the most common FA-ACP in *C. reinhardtii* was chosen. The most promising thioesterase is that from *Cuphea hookeriana* (Jones et al., 1995). The specificity of this enzyme (termed *Ch*.FatB1) is for 16:0-ACP (*i.e.* palmitoyl-ACP), which appears to be the major FA in *Chlamydomonas* (Wang et al., 2009). Isolation of the gene coding for *Ch*.FatB1 and its expression in *E. coli* resulted in increased accumulation of the free medium-chain fatty acids in the bacterial cultures (Jones et al., 1995).

With the intention of replicating this experiment in the algal chloroplast, the pASapI expression vector developed in the Purton Lab for chloroplast transformation in *C. reinhardtii* was used (Bateman and Purton, 2000). The selectable marker used in this system is based on restoration of the photosynthetic activity of the recipient line and is described in detail further in this chapter.

As a first step, a pASapI–*FatB1* plasmid containing a synthetic version of thioesterase *FatB1* gene from *Cuphea hookeriana* was created. In order to achieve the efficient translation of the transgene, the *FatB1* was codon optimized for the *Chlamydomonas reinhardtii* chloroplast genome and synthesized commercially by GeneArt (www.geneart.com). The Codon Adaptation Index (CAI) for the synthetic *FatB1* gene was calculated using the graphical codon usage software (http://gcua.schoedl.de/) and was approximately 0.8. CAI evaluates the adaptation of a coding sequence towards the selection of codons complementary to the common tRNAs and the achieved 0.8 value has been recommended for the plastid genes that are highly expressed in the *C. reinhardtii* cells (Morton, 1998).

For cloning purposes, the *FatB1* gene was flanked by *Sap*I and *Sph*I restriction sites, which enabled its incorporation into the pASapI plasmid. Additionally, at the end of the *FatB1* gene before the stop codon (TAA), coding sequence for a HA tag was added. The HA sequence consists of a nine-residue influenza haemagglutinin epitope (YPYDVPDYA) and was added to create a protein detection system for FatB1 in the transformed cells using α-HA antibodies. The sequence of the created pASapI-*FatB1*

expression vector was confirmed by sequencing. Figure 5-3 presents a schematic illustration of the cloning strategy of the *FatB1* gene within the pASapI vector.

pASapI vector enables insertion of the open reading frames (ORFs) of a gene of interest at the start codon (ATG) flanked by the regulatory sequences. The regulatory regions were isolated from two chloroplast genes of *C. reinhardtii*: promoter/5'UTR of the *atpA* gene and 3'UTR of the *rbcL* gene. Figure 5-3 presents a cloning strategy of the pASapI-*FatB1* expression vector construction.

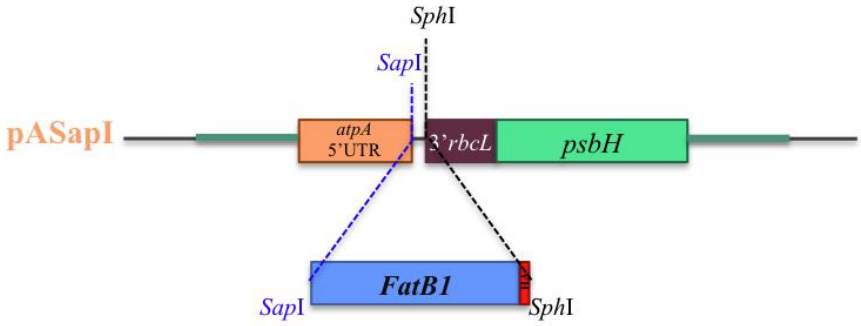


Figure 5-3 Schematic diagram of the creation of pASapI-*FatB1* expression vector.

Synthetic *FatB1* gene codon optimized for *C. reinhardtii* chloroplast containing a HA tag and flanked by *Sap*I and *Sph*I restriction site was introduced into the *Sap*I-*Sph*I site of the pASapI vector. The *psbH* flanking regions are required for homologous recombination with the chloroplast genome and at the same time function as selectable marker.

The created pASapI–*FatB1* expression vector was introduced via biolistic particle method into the chloroplast of the recipient cell line BST-SAME. This recipient cell line is photosynthetically disabled through the disrupting of the photosystem II gene *psbH* by inserting of the *aadA* cassette into a unique *Bst*XI site. The *psbH* gene encodes a crucial protein that is a part of photosystem II. Consequently, the BST-SAME cell line is unable to assemble PSII resulting in impaired photosynthetic activity hence this strain fails to grow on minimum medium. pASapI vector contains the wild type *psbH* gene, therefore the selection for transformed lines occurs via restoration of phototrophic growth. Figure 5-4 presents the mechanism of the *FatB1* gene incorporation into the BST-SAME cell line via homologous recombination.

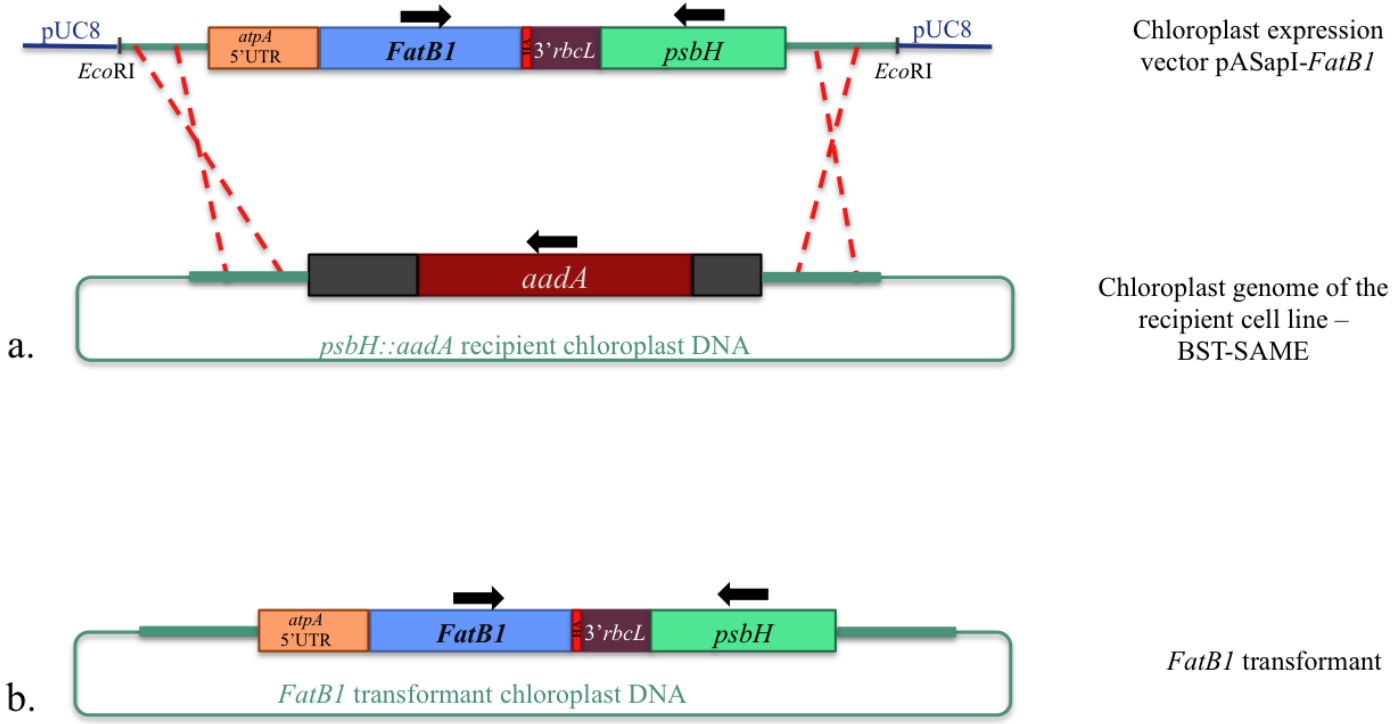


Figure 5-4 Chloroplast transformation of the BST-SAME recipient cell line with pASapI-*FatB1*.

a. Mechanism of the homologous recombination (red dashed line) occurring between the flanking sequences within the pASapI-*FatB1* construct and the corresponding regions in the chloroplast genome of BST-SAME – photosynthetic-deficient *psbH*-deleted recipient cell line containing *aadA* gene conferring spectinomycin and streptomycin resistance.
b. The expected genome of the transformant where the photosynthetic activity was restored by replacement of the *aadA*-disrupted *psbH* gene with a functional copy of *psbH*.

Repeated attempts to transform the BST-SAME cell line with pASapI–*FatB1* expression vector were carried out, yet no transformed colonies appeared on the agar plates containing minimal medium.

The bulk of the synthesized FA in *C. reinhardtii* cells are C_{16} and C_{18} in a 1:1 proportion (Wang et al., 2009). Since the length of fatty acid chain is controlled by fatty acyl-ACP thioesterases, by overexpressing a C_{16} thioesterase, a change in the FA balance was attempted in favour of the shorter chain FA. To date, numerous thioesterases acting on long-chain acyl-ACPs have been isolated from plants, however, all of them can be organized in two groups: FatA principally acting on oleonyl-ACP and FatB preferentially hydrolyzing the saturated FAs (Jones et al., 1995). Since the proportion of saturated to unsaturated FAs synthesized in *C. reinhardtii* is 1:1 (Wang et al., 2009), choosing the FatB type of thioesterase over FatA should not have a negative effect in this experiment.

Transformation with thioesterases isolated from *Cuphea sp* has been well documented in *E. coli* cells (Dehesh et al., 1996a, Dehesh et al., 1996b) and *Brassica* plants (Dehesh et al., 1998). Both experiments however, focused on overexpression of medium chain fatty acids. Overexpression of thioeserases isolated from *Cuphea palustris Cp FatB1* and *Cp FatB2* in *E. coli* cells resulted in increased secretion of 8:0/10:0 and 14:0/16:0 fatty acids respectively (Dehesh et al., 1996a). Similarly, transformation of bacterial cells with *Ch FatB2* isolated from *Cuphea hookeriana* triggered the production of 8:0 and 10:0 fatty acids in *E. coli* and this activity was associated with a significant decrease in the levels of linoleate (18:2) and linolenate (18:3), that are constitutively produced in *E. coli* (Dehesh et al., 1996b).

It has been demonstrated that there is a difference in the activity of the two thioesterases from *Cuphea palustris* in seeds, and FatB2 is kinetically superior to FatB1 (Dehesh et al., 1996a). In *Cuphea hookeriana* there is a characteristic expression pattern of both thioesterases, with Ch FatB1 specific for 16:0-ACP being expressed everywhere in plant tissue, whereas Ch FatB2 shows a strong preference towards 8:0/10:0-ACP and its expression is associated with embryogenesis (Dehesh et al., 1996b). This pattern confirms an earlier publication revealing that medium chain thioesterases are found predominantly in the developing organs of plants (Pollard et al., 1991).

There has also been an attempt to introduce a thioesterase from a diatom *Phaeodactylum tricornutum* (PtTE) into *E. coli* (Gong et al., 2011a). PtTE has no similarities to any of the thioesterases present in bacterial or plant cells. Overexpression of PtTE in *E. coli* resulted in the enhanced content of total FA, and C_{18} was particularly increased. On the other hand, transformation of *Phaeodactylum tricornutum* with a heterologous thioesterase biased towards the production of C_{12} and C_{14}, increased the total lipid production of the transformed algal cell line and also altered the fatty acid composition in the algal triacylglycerols in favour of the shorter chain FAs (Radakovits et al., 2011).

The efficiency of the FA production, additionally to the type of thioesterase, is also associated with the species the enzyme is isolated from. In an experiment aiming to compare the thioesterase activity, acyl-ACP thioesterase *FatB* genes were isolated from *Diploknema butyracea*, *Gossypium hirsutum*, *Ricinus communis* and *Jatropha curcas*. All four genes were introduced into *E. coli* and the production of FA in the transformed strains was compared. Thioesterases isolated from *Ricinus communis* and *Jatropha curcas* species showed the highest production level of FA in bacterial cells (Zhang et al., 2011b). Therefore, in the future attempts to optimize the transgenic thioesterase activity in algal chloroplasts, all factors discussed above should be taken into consideration.

5.3.2 Introduction of aldehyde decarbonylase and acyl reductase into the *C. reinhardtii* chloroplast

5.3.2.1 Construction of chloroplast transformation vectors containing genes for the enzymes acyl reductase and aldehyde decarbonylase

A publication that greatly inspired this part of the project was published in Science by Schirmer et al (2010). It was reported that expression in *E. coli* of two genes isolated from *Synechococcus elongatus* PCC7942: orf1594, that belongs to the short-chain dehydrogenase or reductase family and orf1593, that belongs to ribonucleotide reductase-like family, resulted in the production of odd-chain alkanes and alkenes as well as even-chain fatty aldehydes and fatty alcohols. Pentadecane and heptadecane production was particularly demonstrated. These two genes are found as two-gene operon in many cyanobacteria including *Synechocystis* PCC6803. Therefore, the equivalents of the two genes: aldehyde decarbonylase (sll0208) and acyl reductase (sll0209) from *Synechocystis* PCC6803 were cloned into the *Chlamydomonas* chloroplast transformation system developed in the Purton lab.

The aldehyde decarbonylase and acyl reductase genes, named in this work as "DR17", were amplified using the following primers: DR17.F [CCAGCTCTTCTATGCCCGAGCTTGCTG] and DR17.R that additionally introduced a Human influenza hemagglutinin [HA] tag [GGCATGCTTAAGCATAGTCAGGAACATCGTATGGATAAAGAGCTACTAAAGGGCAAAAG] onto the reductase. The HA tag was added to create a protein detection system for acyl reductase in the transformed cells using α-HA antibodies. Furthermore, the primers were designed to introduce *Sap*I and *Sph*I restriction sites at the ends of the amplified DR17 fragment what enabled cloning of the amplified DR17 fragment into pASapI (containing *atpA* promoter/5'UTR), pPSapI (containing *psbA* promoter/5'UTR), pCSap2 (containing *chlL* promoter/5'UTR) and pSSapI (containing *psaA* promoter/5'UTR) plasmids. All expression vectors contain the *rbcL* 3'UTR and were created in the Purton Lab (Ninlayarn, 2012). Figure 5-5 presents the cloning strategy of expression plasmids pASapI-DR17, pPSapI-DR17, pCSap2-DR17 and pSSapI-DR17.

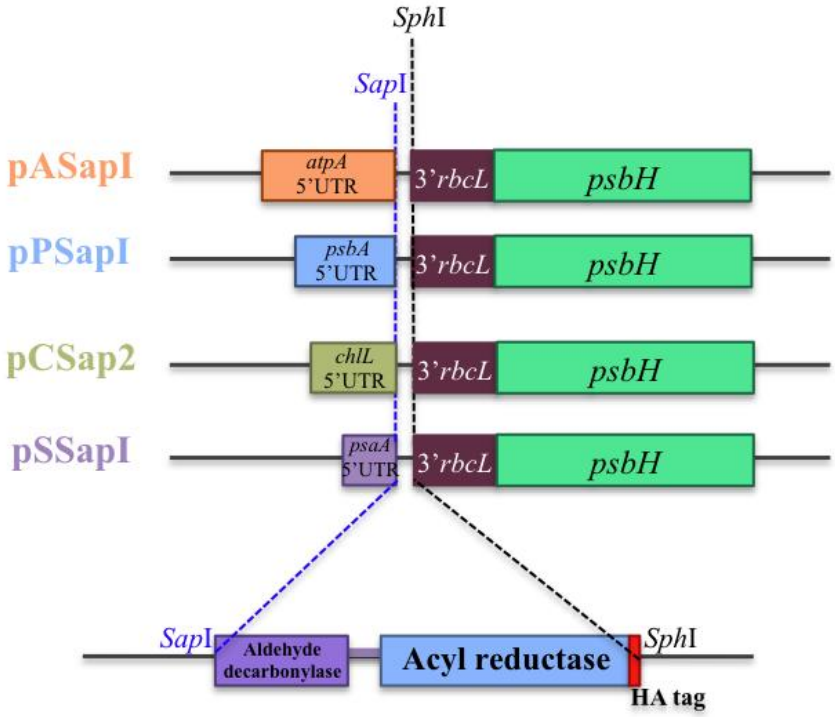


Figure 5-5 Schematic diagram of the DR17 plasmids creation under different promoter/5'UTRs.

Aldehyde decarbonylase and acyl reductase (DR17) genes amplified from *Synechocystis* PCC 6803 using DR17.F and DR17.R primers that introduced an HA tag to acyl reductase and also two restriction sites *Sap*I and *Sph*I. The amplified genes were inserted into four expression vectors pASapI, pPSapI, pCSap2 and pSSapI containing *atpA*, *psbA*, *chlL* and *psaA* promoters/5'UTR respectively through the *Sap*I and *Sph*I restriction sites.

Characteristic features of the four regulatory sequences isolated from the genes for abundant chloroplast proteins used in this work can be found in

. Sequences of the created expression vectors were confirmed by sequencing. Full sequence of the amplified DR17 operon and the sequences of the created chloroplast expression vectors pASapI, pCSap2, pPSapI and pSSapI can be found in appendix H.

Table 5-1 Evaluation of the promoters/5'UTR used in cloning of the aldehyde decarbonylase and acyl reductase genes into chloroplast expression vectors.

Promoter/ 5'UTR	Position relative to endogenous start codon	Size [bp]	Protein function
atpA	-595 to -1	620	ATP synthase α-subunit
psbA	-240 to -1	240	Protein forming the PSII complex
psaA exon 1	-238 to -1	312	Protein forming the PSI complex
chlL	-451 to +48	531	Chloroplast-encoded component involved in light-independent reaction of chlorophyll formation

As discussed in Chapter 4, efficient chloroplast gene expression is controlled by numerous post-transcriptional steps that involve an interaction of various factors encoded by nuclear genes including those that interact with the 5' UTR of mRNA (Rochaix, 1996, Stampacchia et al., 1997, Rochaix, 2001). The choice of the 5' regulatory region has a substantial impact on the foreign protein synthesis and accumulation, whereas the type of 3'UTR has a minor influence on transgene expression (Barnes et al., 2005). The isolated acyl reductase and aldehyde decarbonylase genes (DR17) were not codon optimized for chloroplast gene expression in *C. reinhardtii*. Moreover, both transgenes are adjacent and appear to form a conserved operon (Schirmer et al., 2010) hence, they are most likely to be controlled by the same 5' regulatory region in *Synechocystis* PCC6803. This feature

was replicated in the expression vectors where the expression of both transgenes is controlled by the same promoter/5'UTR region. Therefore, the optimization of the promoter/5'UTR for the DR17 expression is a vital part of this project.

In addition to the promoters that constitutively mediate the expression of chloroplast genes (*atpA*, *psbA* and *psaA*), the effect of an inducible promoter *chlL* on the expression of DR17 genes was tested. Its expression is possibly regulated by light since the accumulation of ChlL is observed in *C. reinhardtii* cells only when maintained in the dark. It was hypothesized that the *cis* elements that are required for the dark/light control are situated either within the *chlL* promoter/5'UTR or within the region surrounding the start codon (Cahoon and Timko, 2000). The expression efficiency of two transgenes *acrV* and *vapA* encoding antigens from the fish pathogen *Aeromonas salmonicida* under *atpA*, *psbA* and *psaA* promoters including an additional *psbD* element have been compared recently (Michelet et al., 2011). The *psaA- exon1* promoter/5'UTR was shown to be the most efficient regulatory region out of four tested to drive the expression of both *acrV* and *vapA* transgenes, that were codon-optimised for chloroplast expression in *Chlamydomonas*. In this work, the optimisation of alkane production in algae through the choice of promoter/5'UTR region is demonstrated.

The four DR17 constructs were designed for chloroplast transformation using the BST-SAME recipient line. All of the expression vectors include the *psbH* gene in addition to the DR17 expression cassette. The *psbH* gene was introduced to the *psbH*-deficient BST-SAME strain through a homologous recombination process, that results in the restored photosynthetic efficiency of the transformed line. Simultaneously, the *aadA* cassette from the recipient cell line is replaced with the DR17 cassette. The schematic illustration of the mechanism of transformation of the BST-SAME cell line with the DR17 constructs is presented in Figure 5-6.a, whereas the expected orientation of the expression cassettes, including re-introduced *psbH* gene within the DR17 transformed lines, are presented in Figure 5-6.b.

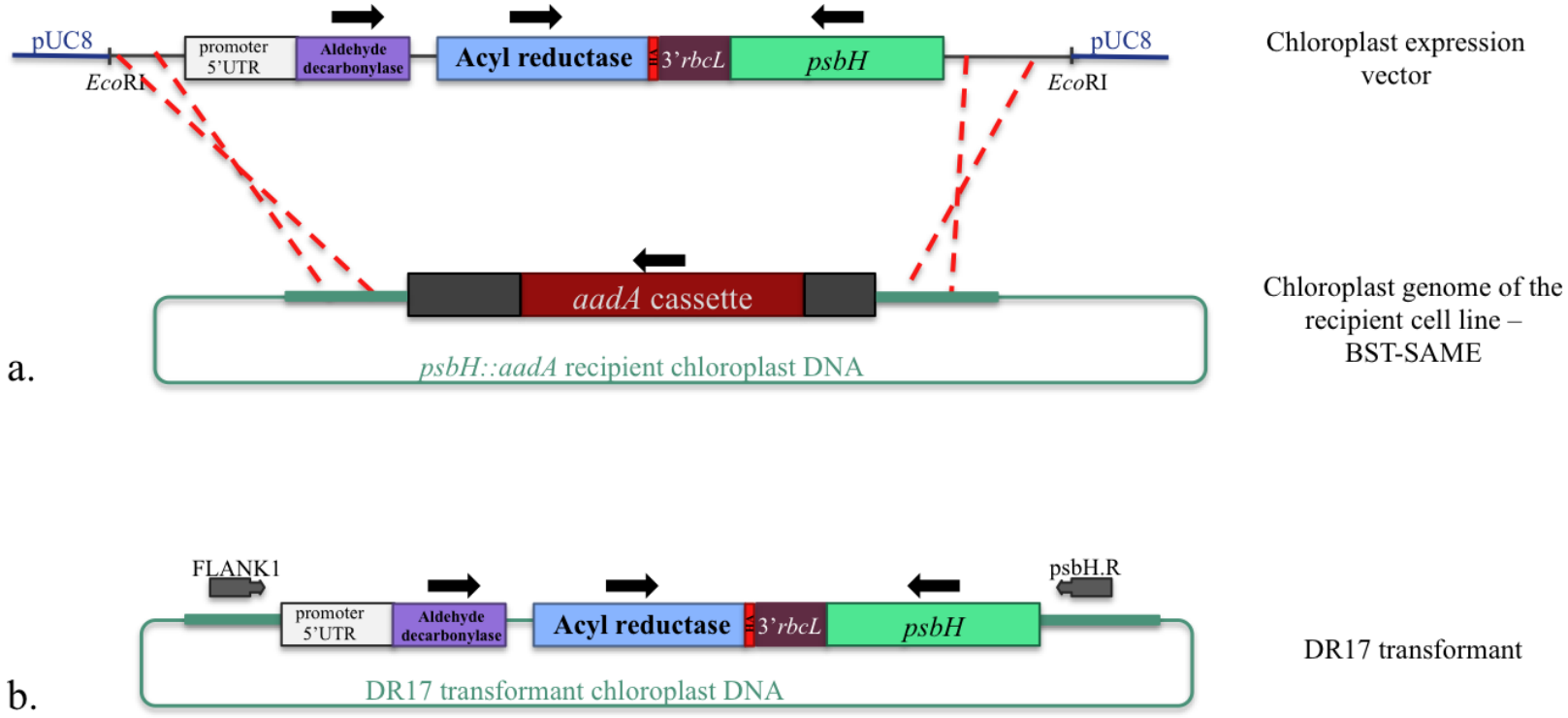


Figure 5-6 Chloroplast transformation of the BST-SAME *C. reinhardtii* recipient cell line with the DR17.

a. Mechanism of the homologous recombination (red dashed line) occurring between the flanking sequences within the p_SapI-DR17 expression vectors containing different promoters/5'UTR and the corresponding regions in the chloroplast genome of the photosynthetic-deficient *psbH*-deleted recipient cell line BST-SAME containing *aadA* gene conferring spectinomycin and streptomycin resistance.
b. The expected genome of the transformants where the photosynthetic activity was restored by replacement of the *aadA*-disrupted *psbH* gene with a functional copy of *psbH*.

FLANK1 and psbH.R – primers used for molecular analysis of the transformed cell lines.

5.3.2.2 Testing the function of the DR17 constructs in *E.coli.*

Many chloroplast promoters are functional in *E.coli*. Therefore, the expression of all four DR17 constructs was tested in *E. coli*. Figure 5-7 presents western blot analysis of the crude cell extracts probed with anti-HA antibodies aiming to detect acyl reductase in the transformed cell lines (a) and semi-quantitative analysis of the protein expression efficiency (b). All four promoter/5'UTR regions work well in *E. coli*, as acyl reductase was present in all the DR17 transformants and the enzyme with HA tag was correctly detected. The semi-quantitative analysis using the Odyssey® system indicated that the *psaA* regulatory region of pSSapI-DR17 transformation vector was the most efficient in driving expression of acyl reductase in bacterial cells, whereas the *chlL* promoter present in the pCSap2-DR17 expression vector was the least effective.

The *psaA* promoter/5'UTR region derived from the chloroplast of *C. reinhardtii* has previously been proven to be the most efficient in *E. coli* out of a number of such elements (Ninlayarn, 2012). It drove the highest resistance level to the antibiotic Zeocin™ through an efficient synthesis of BLE protein that has a capacity to inactivate this antibiotic. Creation of a dual expression vector suitable for both *E. coli* and *C. reinhardtii* is a very attractive idea. It can serve as a primary test to investigate if the created construct is functional and as a positive control for western blot analysis of the algal transformants.

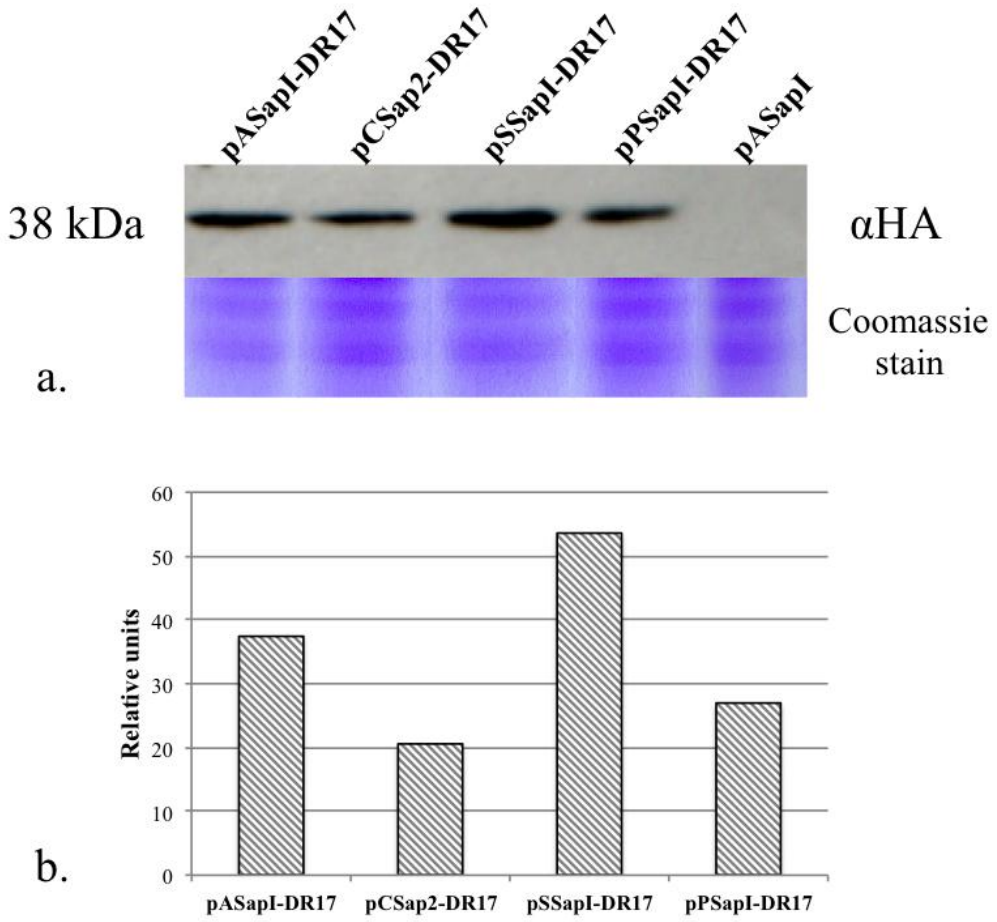


Figure 5-7 Western blot analysis of the *E. coli* lines transformed with the p_SapI-DR17 constructs controlled by different promoter/5'UTR regions.

a. Western blot method was designed to detect the HA tag of acyl reductase. Crude cell extracts were probed with anti-HA antibodies. The expected band size was of approximately 38 kDa. The pASapI vector was used here as negative control whereas Coomassie staining functioned as a loading control.
b. Semi-quantitative analysis of the acyl reductase expression in the transformed cell lines.

It has been reported that the coding regions of cyanobacterial genes can properly function in *E. coli* (Scaife et al., 2009). In this report, four β-carotene ketolase and twelve β-carotene hydroxylase genes derived from five various cyanobacterial species were successfully expressed in both *E. coli* and cyanobacterial hosts. Furthermore, it has been well documented that algal chloroplast promoters operate in *E. coli* (Fargo et al., 1998, Bateman and Purton, 2000, Ninlayarn, 2012). The majority of bacterial regulatory regions require the Shine-Dalgarno (SD) sequence for the initiation of translation However, *E. coli* is also capable of initiating translation activity in the SD-independent process that is common for chloroplast-encoded regulation mechanisms (Fargo et al., 1998).

5.3.2.3 Chloroplast transformation of *C. reinhardtii* with the DR17 vectors.

Chloroplast transformation was performed on BST-SAME recipient cells using the biolistic particle gun method. All four DR17 expression vectors were introduced into the *C. reinhardtii* cell line. Approximately three weeks after bombardment, DR17 transformants appeared. All of them however, were derived solely from the pASapI-DR17 construct and the attempts to transform the chloroplast genome with pCSap2-DR17, pPSapI-DR17 and pSSapI-DR17 were not successful. Five transformants isolated following the transformation with the pASapI-DR17 construct were subjected to three re-streaking cycles to ensure that the chloroplast genomes are homoplasmic and were then used for further analysis.

5.3.2.3.1 Molecular analysis of the transformed lines of C. reinhardtii

Firstly, the correct insertion of the DR17 transgenes within the *C. reinhardtii* chloroplast genome was verified. Figure 5-8.a presents the results of PCR analysis with FLANK1 [GTCATTGCGAAAATACTGGTGC] and psbH.R [GCAACAGGAACTTCTAAAGC] primers.

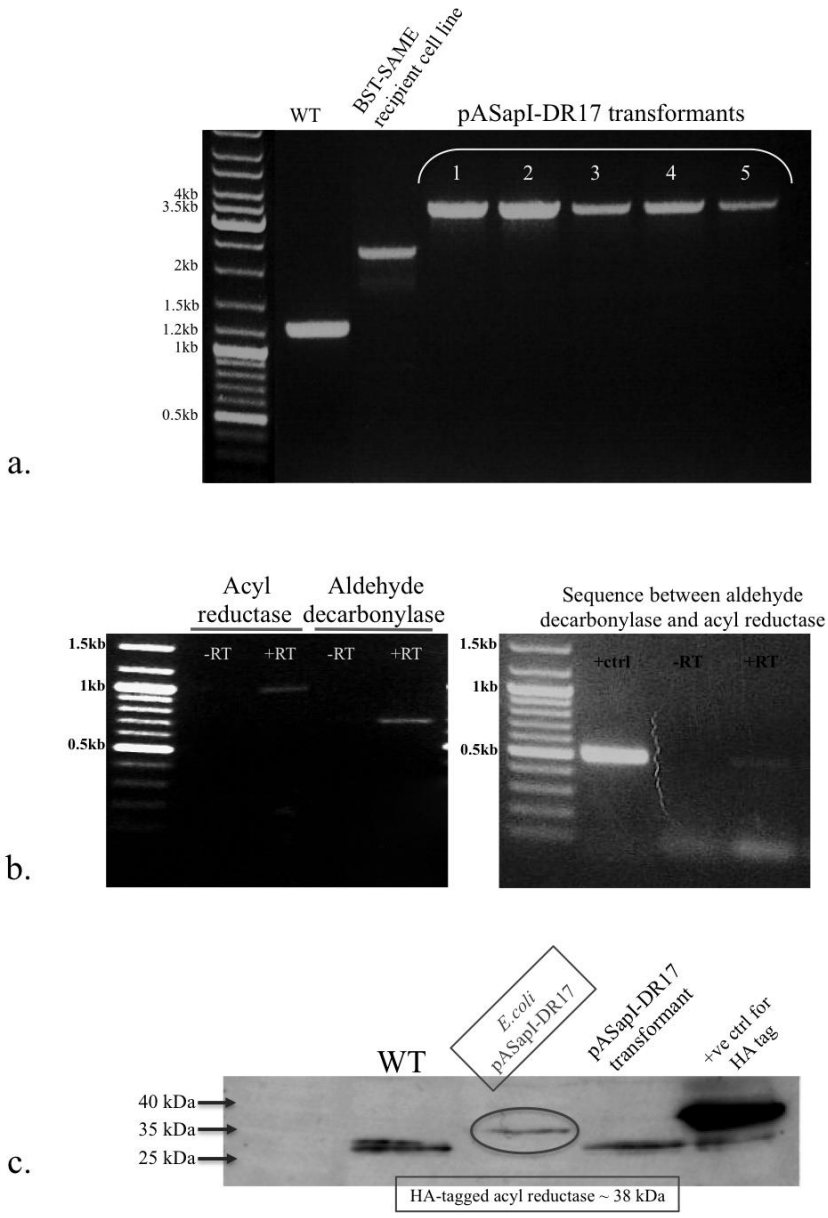


Figure 5-8 Analysis of the pASapI-DR17 transformant of *C. reinhardtii*.

a. PCR analysis of the transformants of *C. reinhardtii* using FLANK1 and psbH.R primers. Expected bands are as follows: for wild type (WT) ~1.2 kb, BST-SAME ~ 2.4 kb, pASapI-DR17 transformants ~3.5 kb. Isolate 1 was chosen for further analysis and this PCR result was confirmed by sequencing.

b. RT-PCR analysis of the pASapI-DR17 transformant 1 to confirm the transcription of the incorporated acyl reductase and aldehyde decarbonylase genes within the chloroplast DNA but also to test if both genes are transcribed simultaneously. Expected PCR products are as follows: ~1 kb for acyl reductase, ~ 0.75 kb for aldehyde decarbonylase and ~0.5 kb for the sequence between both genes.

+RT –cDNA served as a template for this PCR; **–RT** – control for contamination of RNA with gDNA where the isolated RNA served as a template for the PCR; **+ctrl** – pSapI-DR17 plasmid

c. Western blot analysis for the presence of the HA-tagged acyl reductase enzyme. *E. coli* transformed with pASapI-DR17 functioned as positive control for the acyl reductase detection whereas a cell line containing a HA-tagged protein of *C. reinhardtii* was a positive control for the HA analysis.

PCR amplification with FLANK1 and psbH.R aims to amplify the expression cassette targeted to the locus downstream of *psbH* by homologous recombination following transformation of the recipient line (Figure 5-6.b). The PCR analysis confirmed that all five pASapI-DR17 transformants contain aldehyde decarbonylase and alcohol reductase genes in their chloroplast genome since all of them presented the expected ~3.5 kb band. *C. reinhardtii* wild type and BST-SAME strains served as negative controls. pASapI-DR17 transformant 1 was chosen for further analysis to detect an mRNA transcript for both genes (Figure 5-8.b) and protein accumulation in the chloroplasts (Figure 5-8.c).

The aldehyde decarbonylase and acyl reductase genes within the pASapI-DR17 construct were cloned directly from the *Synechocystis* PCC6803 genome and are under the control of the same *atpA* promoter/5'UTR region. Therefore, there was a concern that the transcription of the downstream ORF encoding the reductase could be compromised. In order to confirm the transcription of the aldehyde decarbonylase and acyl reductase genes, reverse transcriptase PCR (RT-PCR) analysis of the two genes in the transformed cell line was performed. RT_D17.F [ATGCCCGAGCTTGCTGTCC] and RT_D17.R [CCAAACCGTAGGAAGACATACG] primers were used to detect the aldehyde decarbonylase sequence within the mRNA transcript, RT_R17.F [ATGTTTGGTCTTATTGGTCATCTCACG] and RT_R17.R [AAGAGCTACTAAAGGGCAAAAGCC] primers were used for acyl reductase detection, whereas RT_D.F [AGTGGAAGATTTTATGATCAGCTACG] and RT_R.R [CTCCACATATTTGCCTTCAATCACC] served to investigate if both genes are co-transcribed. Figure 5-8.b confirms the correct transcription of both transgenes in the chloroplasts of the *C. reinhardtii* cell line, as all three expected bands are present. However, protein analysis of the cell extracts revealed that the acyl reductase is not detectable in the pASapI-DR17 transformant (Figure 5-8.c). In western analysis, wild type *C. reinhardtii* functioned as a negative control, whereas the cell extracts of *E. coli* transformed with pASapI-DR17 was used as a positive control.

In conclusion, the transformants containing the aldehyde decarbonylase and acyl reductase genes were isolated only when placed under *atpA* promoter/5'UTR and transcription of the two genes could be demonstrated. However, the acyl reductase protein in the pASapI-DR17 transformant was not at a detectable level. It is plausible that the other three tested promoter/5'UTRs were more efficient in driving the expression of the aldehyde decarbonylase and acyl reductase than the *atpA* regulatory region, and that the produced alkane concentration was toxic for the host cells.

The DR17 transgenes were correctly transcribed in the pASapI-DR17 *C. reinhardtii* transformant, however the reductase protein was not detectable. Previous research examining the influence of various 5' and 3' untranslated regions within the expression cassette established no correlation between the mRNA level and protein accumulation (Barnes et al., 2005). Similarly, protein synthesis in chloroplasts is frequently unaffected by changes in the gene copy number and the transcript abundance does not affect the translation rate (Eberhard et al., 2002). Moreover, the translation rate of recombinant proteins is decreased when compared to endogenous genes controlled by the same regulatory region, despite the fact that the transcript level of both genes is on similar level (Coragliotti et al., 2011). It is therefore likely that efficient translation of the two genes is affected, possibly due to inappropriate codon usage.

Creation of an inducible expression mechanism would be useful in numerous applications. This machinery would be particularly advantageous in applications where the transgene expression influences life span of the transformed strain. *chlL* regulatory region in pCSapI expression vector was designed as dark-inducible. It has been reported that the same promoter/5'UTR of *chlL* triggered the expression of human growth hormone, even under continuous illumination (Ninlayarn, 2012). This result suggests that the *cis* elements controlling the photoregulation were missing from the created *chlL* promoter/5'UTR region.

5.3.2.3.2 Gas chromatography (GC) analysis of the transformed lines of Chlamydomonas reinhardtii

Negative results obtained during the protein analysis indicated that the concentration of the transgenic acyl reductase enzyme in the transformed cell line was not sufficient for western analysis. However, this finding does not exclude a possibility that the very low levels of the enzymes might be present in the transformants. Hence, an additional method was employed in order to analyze possible changes in the level of alkanes and their derivatives in the transformed cell line. Gas chromatography (GC) analysis of the pASapI-DR17 transformant was performed and the results are presented in Figure 5-9. In addition to the cell extract (DR_pellet) analysis of the transformant, growth medium (DR_supernatant) was also analyzed in order to examine the possibility that the created compounds were secreted by the tested cell line.

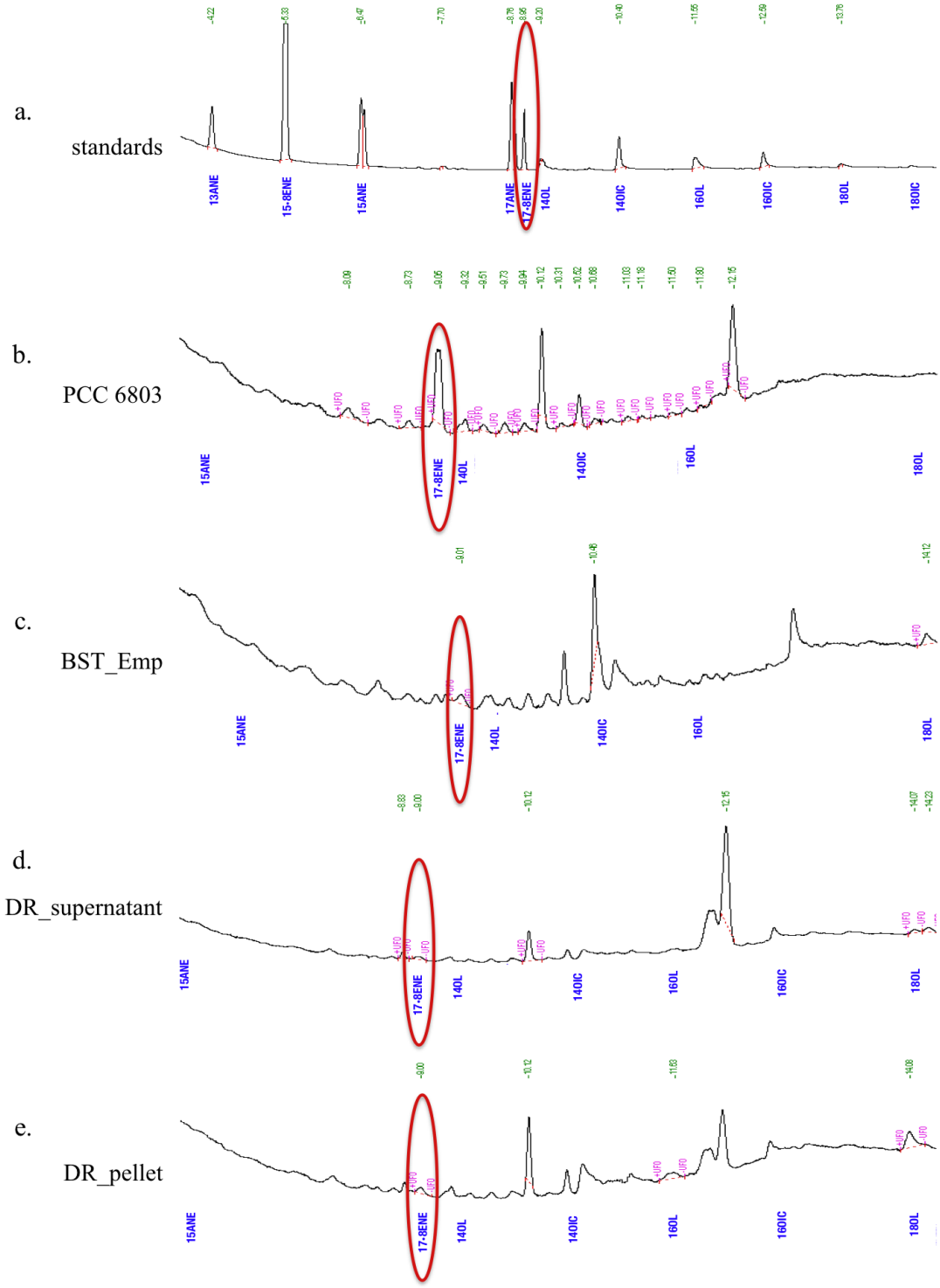


Figure 5-9 GC analysis of the pASapI-DR17 transformants of *C. reinhardtii*.

All the samples were run in parallel with a mixture of standards (a). Cell extract of 7-day old culture of *Synechocystis* PCC 6803 (b) functioned as a positive control for C_{17} alkene (17•8ENE), BST_Emp (c) – cell extract of 4-day old culture of recipient cell line of *C. reinhardtii* transformed with pASapI vector, DR - 4-day old culture of BST-SAME recipient cell line of *C. reinhardtii* transformed with pASapI-DR17. 4-day old culture of DR transformants were analyzed for the presence of alkanes and alkenes in both forms: secreted - DR_supernatant (d) and accumulated in the chloroplasts - DR_pellet (e). Samples of the controls PCC6803 and TNEmp were obtained from cell extracts. Standard mix contained C_{13}, C_{15} and C_{17} alkanes (ANE), C_{15} and C_{17} alkenes (8ENE), C_{14}, C_{16} and C_{18} fatty acids (OIC) and C_{14}, C_{16} and C_{18} alcohols (OL).

A cell extract of *Synechocystis* PCC6803 was used as a positive control for C_{17} alkene detection, whereas BST-Emp (BST-SAME recipient cell line transformed with an empty pASapI vector) functioned as a negative control. The standard mix contained C_{13}, C_{15} and C_{17} alkanes, C_{15} and C_{17} alkenes, C_{14}, C_{16} and C_{18} fatty acids and C_{14}, C_{16} and C_{18} alcohols. The acyl reductase isolated from *Synechocystis* PCC6803 uses C_{18} fatty acids as substrates for aldehyde production, whereas the aldehyde decarbonylase converts the C_{18} aldehydes to C_{17} alkanes. Therefore, in the pASapI-DR17 transformant the characteristic peak for C_{17} alkanes was expected, as seen in the control (*Synechocystis* PCC6803 - Figure 5-9.b). Unfortunately, this peak was absent in both samples from the DR17 transformant. Additionally, there were no notable differences between chromatograms of cell extract and culture supernatant from the analyzed cell line.

Gas chromatography (GC) analysis is an approximate procedure that enables identification of the volatile compounds, rather than precise quantification. Yet, it provides preliminary information on the chemical composition of the samples, based on retention time of peaks of known molecules in a standard mix. In principle, the same molecules should have the same retention times in the same conditions, nevertheless this value is usually approximate. The GC/MS (gas chromatography–mass spectrometry) method allows for very precise identification of even trace amounts of volatile compounds in analyzed samples. This method is more precise than GC analysis, therefore using GC/MS would have provided more detailed information on the metabolism of the pASapI-DR17 transformant.

The biosynthesis of alkanes in plants is poorly understood. The proposed pathway of alkane synthesis within the leaf epicuticular wax layer is presented in Figure 5-10. This pathway resembles the alkane synthesis process in cyanobacteria discussed earlier, where acyl-CoA is transformed to a corresponding aldehyde by acyl reductase and the aldehyde is then a substrate for decarbonylase to form the alkane.

The plant gene *YRE* is speculated to encode an aldehyde-generating acyl-CoA, since the wax composition of *yre-1 Arabidopsis* mutants contains increased levels of fatty alcohols when compared to wild type (Kurata et al., 2003). Similarly, *CER1* is hypothesized to either encode or regulate decarbonylase enzyme, since the wax composition in *cer-1* mutants contains decreased total amounts of alkane, secondary alcohols and ketone (Aarts et al., 1995).

Figure 5-10 Schematic illustration of the wax synthesis pathway in *Arabidopsis*.
Adapted from (Kurata et al., 2003)

An alternative pathway for the wax molecule synthesis has been proposed in the bacterial species *Acinetobacter cacoaceticus* (Reiser and Somerville, 1997). In this pathway (Figure 5-11) acyl-CoA is converted to the corresponding aldehyde by acyl-CoA reductase, and the aldehyde becomes a substrate for the fatty aldehyde reductase catalyzing fatty alcohol synthesis. A fatty alcohol acyl-CoA transferase condenses fatty alcohol and acyl-CoA to synthesise wax esters. *Acinetobacter* reductases from the proposed wax biosynthesis pathway can function as an alternative to the cyanobacterial reductase in alkane production. It has been reported that *acr1* gene isolated from *Acinetobacter baylyi*, coding for the acyl reductase, has been employed in applications for microbial metabolic engineering of FAs (Yan and Liao, 2009). Additionally, fatty alcohols have been identified as a potential precursor in alkane biosynthesis. For instance, a pathway of alkane biosynthesis including fatty alcohols as byproducts has been proposed for *Vibrio furnissii*, where the first two steps occur with fatty alcohols reduced to alkanes (Park, 2005).

Figure 5-11 Proposed pathway of wax ester biosynthesis in *Acinetobacter cacoaceticus.*
Adapted from (Reiser and Somerville, 1997)

5.3.2.4 Construction of two chloroplast transformation vectors containing a single gene coding for acyl reductase (pASapI-R) or aldehyde decarbonylase (pASapI-D)

The expression in *E. coli* of the *Synechococcus elongatus* PCC7942 orf1594 coding for acyl reductase resulted in an increased production of fatty aldehydes and fatty alcohols (Schirmer et al., 2010). The pASapI-DR17 expression vector was created as an operon, where both aldehyde decarbonylase and acyl reductase genes were under control of the same *atpA* regulatory region. The detection of acyl reductase during western analysis of the pASapI-DR17 transformant was unsuccessful (Figure 5-8.c) suggesting that translation of both genes from the same mRNA transcript was inefficient. In order to test this possibility, two separate chloroplast expression vectors were created and in both the *atpA* regulatory region controlled the expression of each gene. The pASapI-R construct contains the acyl reductase (sll0209) gene amplified from *Synechocystis* PCC6803 using the following primers: R_SapI.F [ACGCTCTTCAATGTTTGGTCTTATTGG] and DR17.R [GGCATGCTTAAGCATAGTCAGGAACATCGTATGGATAAAGAGCTACTAAAGGGCAAAAG].
The pASapI-D expression vector contains the aldehyde decarbonylase (sll0208) gene amplified using primers DR17.F [CCAGCTCTTCTATGCCCGAGCTTGCTG] and D_SphI.R [GGCATGCTTAAGCATAGTCAGGAACATCGTATGGATAGACTCCGGCCAAACCGTAGGAAG].

Both reverse primers were designed to introduce a human influenza hemagglutinin [HA] tag at the C-terminus to facilitate detection of the introduced protein in the transformed cell lines. Additionally, both sets of primers were designed to introduce *Sap*I and *Sph*I restriction sites, which enabled incorporation of both genes into the pASapI expression vector. The pASapI-D and pASapI-R constructs were confirmed by sequencing. Figure 5-12 presents the cloning strategy of both expression vectors.

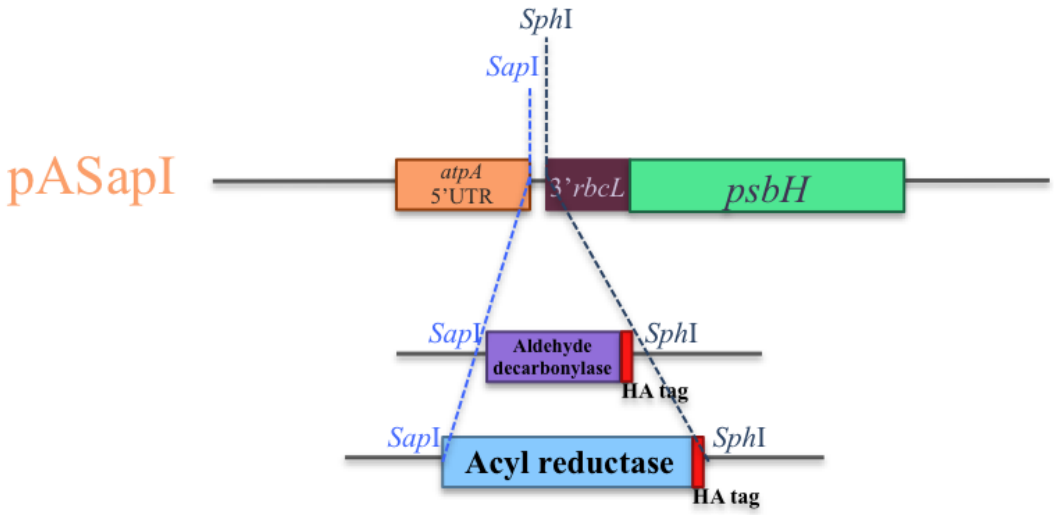


Figure 5-12 Schematic diagram of the creation of the pASapI-D (containing aldehyde decarbonylase) and pASapI-R (containing acyl reductase) plasmids.

The aldehyde decarbonylase and acyl reductase genes were amplified independently from *Synechocystis* PCC 6803 using two sets of primers: DR17.F + D_SphI.R for aldehyde decarbonylase amplification and R_SapI.F + DR17.R for acyl reductase amplification. Both PCR reactions introduced the HA tag and two restriction sites *Sap*I and *Sph*I to each amplified gene. This enabled insertion of the modified genes into the *Sap*I – *Sph*I restriction site of pASapI expression vector.

5.3.2.5 Chloroplast transformation of *C. reinhardtii* with the pASapI-D and pASapI-R expression vectors

The two constructs were introduced into the cell wall-deficient TN72 recipient line using a transformation method based on agitation of a cell and DNA suspension in the presence of glass beads. TN72 was created in the Purton lab (Ninlayarn, 2012) to enable simpler and cheaper chloroplast transformation compared to the biolistic method used earlier in this chapter. Analogously to the previously employed BST-SAME recipient cell line, TN72 contains a disrupted *psbH* gene, which results in its photosynthetic deficiency. The specific design of the pASapI vector and the chloroplast genome of TN72 enables restoration of photosynthetic ability as a selectable marker. Figure 5-13.a presents the mechanism of chloroplast transformation of the TN72 with the two expression vectors: pASapI-D and pASapI-R. During the chloroplast transformation, the *aadA* cassette present in chloroplast genome of TN72, is replaced with the expression cassette containing acyl reductase or aldehyde decarbonylase genes, flanked by *atpA* promoter/5'UTR region and *rbcL* 3'UTR via homologous recombination. Figure 5-13.b presents the expected orientation of the expression cassette in the chloroplast of the transformants. As a result of the chloroplast transformation, a number of transformants were isolated and two transformants derived from each expression vector were subjected to further analysis.

5.3.2.5.1 Molecular analysis of the transformed lines of C. reinhardtii

In order to confirm the presence of the transgenes in the chloroplast genome, PCR analysis was carried out. Figure 5-14.a presents the results of the PCR analysis with FLANK1 [GTCATTGCGAAAATACTGGTGC] and psbH.R [GCAACAGGAACTTCTAAAGC] primers. It aims to amplify the expression cassette between the flanking regions used in the homologous recombination during the chloroplast transformation of the recipient cell line. Wild type and TNEmp (TN72 transformed with the empty pASapI) were employed as negative controls. The PCR amplification confirmed that both transformants have the expression cassettes correctly incorporated, since the pASapI-D transformant presents a band of approximately 2.4 kb whereas the pASapI-R isolate shows a band of approximately 2.7 kb.

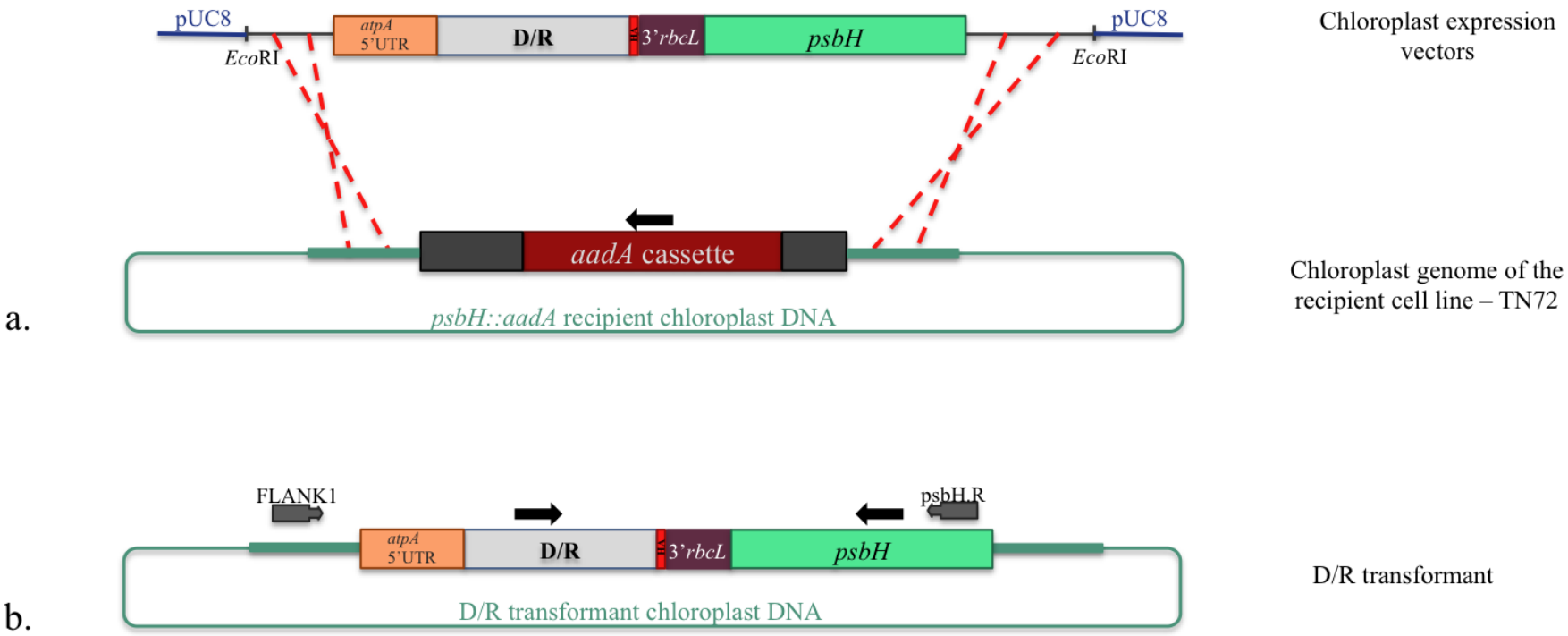


Figure 5-13 Chloroplast transformation of the TN72 recipient cell line of *C. reinhardtii* with the pASapI-D and pASapI-R expression vectors.

a. Mechanism of homologous recombination (red dashed line) occurring between the flanking sequences within the pASapI-D or pASapI-R expression vectors and the corresponding regions in the chloroplast genome of the photosynthetic-deficient *psbH*-deleted recipient cell line TN72 containing *aadA* gene conferring spectinomycin and streptomycin resistance.
b. The expected genome of the transformants where the photosynthetic activity was restored by replacement of the *aadA*-disrupted *psbH* gene with a functional copy of *psbH*.

FLANK1 and psbH.R – primers used for molecular analysis of the transformed cell lines.

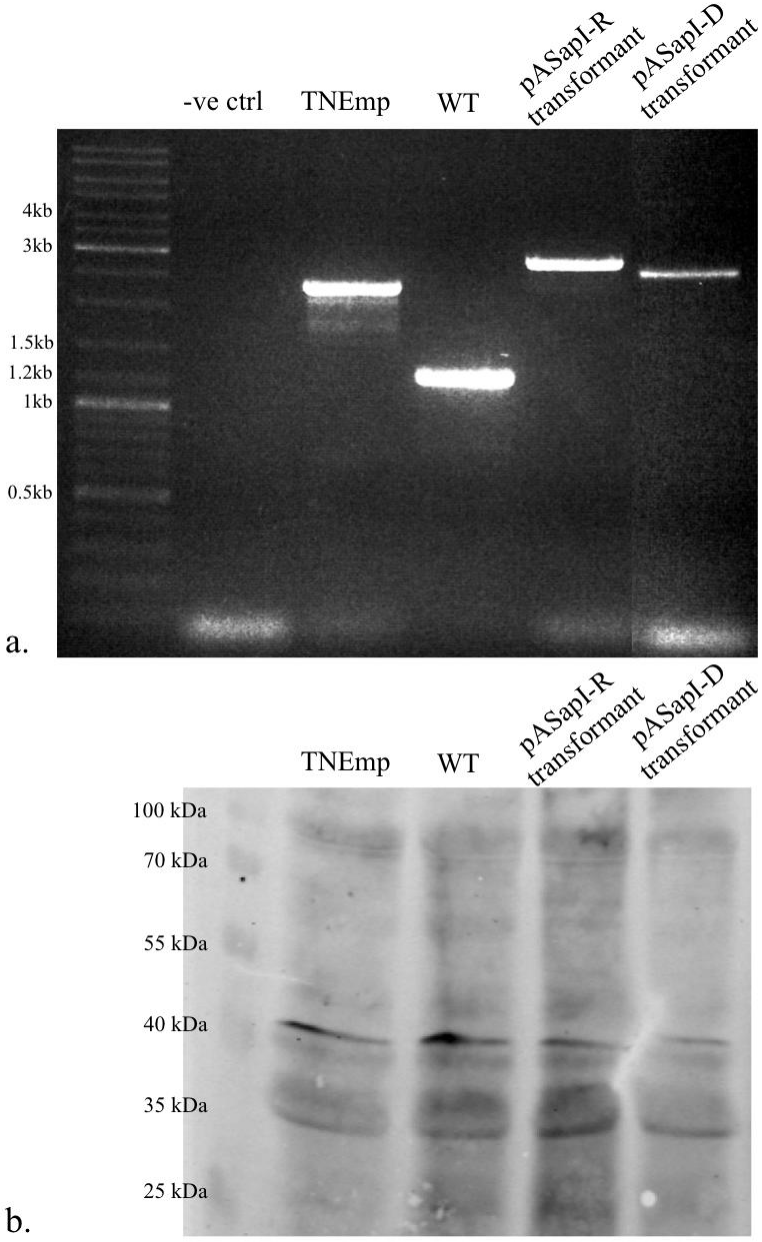


Figure 5-14 Analysis of the pASapI-D and pASapI-R transformants of *C. reinhardtii*.

a. PCR analysis of the transformants using FLANK1 and psbH.R primers. Expected bands are as follows: for wild type (**WT**) ~1.2 kb, TN72 recipient cell line transformed with pASapI vector (**TNEmp**) ~ 2.3 kb, **pASapI-R** transformants ~2.7 kb, **pASapI-D** transformant ~2.4 kb. **–ve ctrl** = PCR reaction without DNA. This result was further confirmed by sequencing.
b. Western blot analysis for the presence of the HA-tagged aldehyde decarbonylase (in **pASapI-D** transformant) and acyl reductase enzyme (in **pASapI-R** transformant). Expected bands are as follows: 27.4 kDa for aldehyde decarbonylase and 37.6 kDa for acyl reductase.

Unfortunately, western blot analysis (Figure 5-14.b) aiming to detect the HA tag affixed to both enzymes failed to identify aldehyde decarbonylase and acyl reductase, since the expected bands of 27.4 kDa (characteristic for aldehyde decarbonylase) and 37.6 kDa (characteristic for acyl reductase) were not detected on the membranes probed with anti-HA antibodies.

5.3.2.5.2 Gas chromatography analysis of the transformed lines of C. reinhardtii

In order to test for any residual activity of the introduced aldehyde decarbonylase and acyl reductase, gas chromatography analysis of the pASapI-D and pASapI-R transformants was performed. The results are presented in Figure 5-15. However, as expected from the western results, there was no difference in fatty alcohol production level between pASapI-R transformant and negative controls – TNEmp and wild type. Moreover, no characteristic alcanols, alkanes or alkenes were present in the analyzed cell extracts of pASapI-D (Figure 5-15.d) and pASapI-R (Figure 5-15.e) transformants. The GC chromatograms of the cell extracts of both transformants show a number of additional peaks in respect to TNEmp (Figure 5-15.c). Yet, it was not possible to identify them with the standard mix that contains a selection of alkanes, alkenes, fatty acids and alcohols (Figure 5-15.a) and more detailed information would be obtained only from GC/MS analysis.

The aldehyde decarbonylase and acyl reductase transgenes, used to induce expression of alkanes in algal chloroplast, were amplified directly from the *Synechocystis* PCC6803 genome. No detectable level of either enzymes present in cell extracts of the transformed algal strains (Figure 5-8.c, Figure 5-14.b) suggests that both genes are poorly expressed in the chloroplast or the enzymes are not expressed at all because of inappropriate codon bias in the genes. The sequences of aldehyde decarbonylase and acyl reductase were not optimized for expression in chloroplast of *C. reinhardtii* and that could be the main reason for the failure of the experiments presented in this part of the chapter.

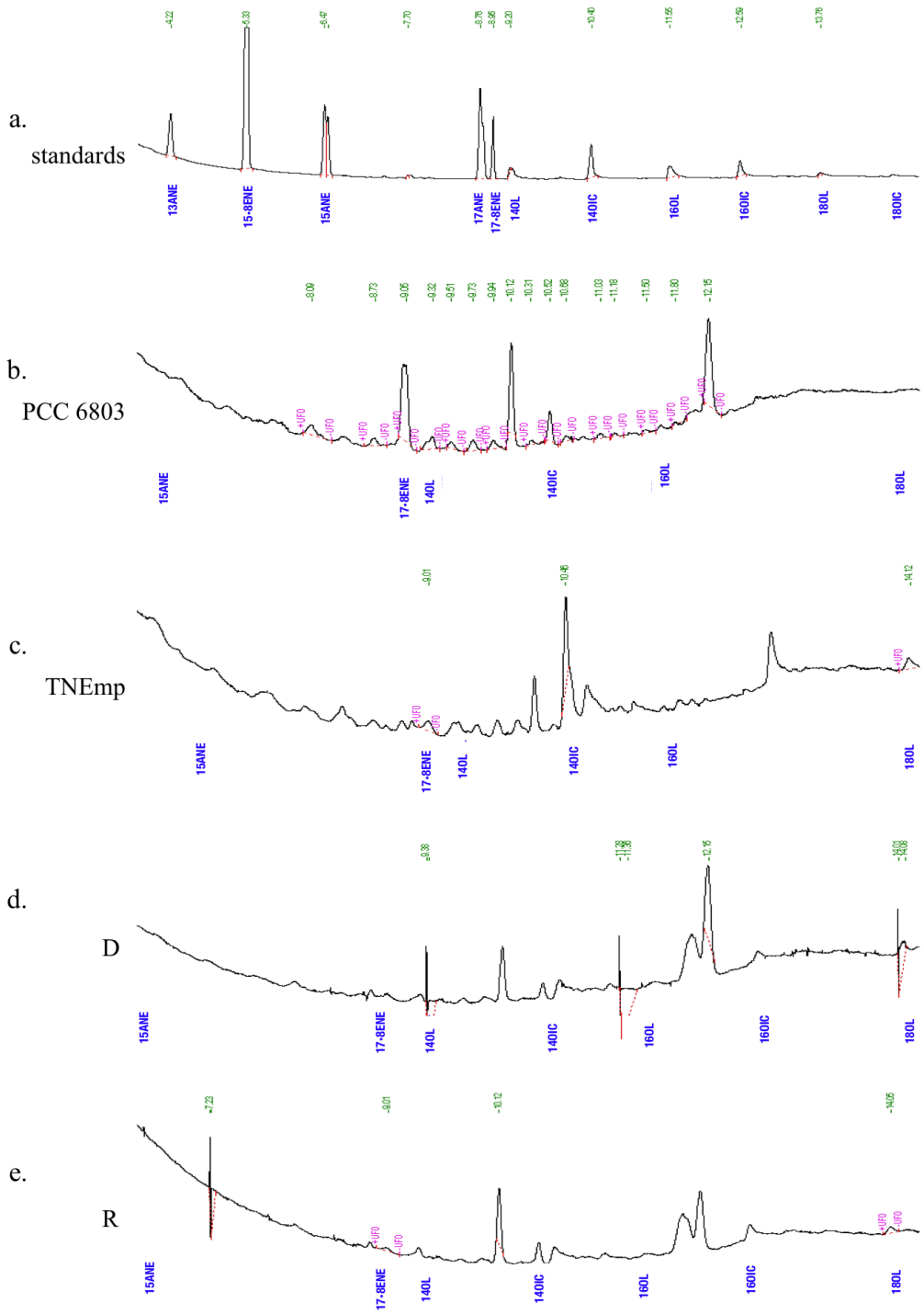


Figure 5-15 GC analysis of the pASapI-D and pASapI-R transformants of *C. reinhardtii.*

All the samples were run in parallel with a mixture of standards (a). Cell extract of 7-day old culture of *Synechocystis* PCC 6803 (b) functioned as a positive control for C_{17} alkene (17•8ENE), TNEmp (c) – cell extract of 4-day old culture of TN72 recipient cell line of *C. reinhardtii* transformed with pASapI vector, D (d) – cell extract of 4-day old culture of TN72 cell line transformed with pASapI-D vector, R (e) – cell extract of 4-day old culture of TN72 cell line transformed with pASapI-R vector. All samples were obtained from the cell extracts.

Standard mix contained C_{13}, C_{15} and C_{17} alkanes (ANE), C_{15} and C_{17} alkenes (8ENE), C_{14}, C_{16} and C_{18} fatty acids (OIC) and C_{14}, C_{16} and C_{18} alcanols (OL).

Codon optimization is commonly applied for the efficient expression of the introduced transgene (Mayfield et al., 2003, Mayfield et al., 2007, Coragliotti et al., 2011). For instance, the expression of codon-optimised ferrochelatase gene *hemH* from *Bradyrhizobium japonicum* and the leghemoglobin gene *lba* from *Glycine max* resulted in 6.8-fold increase of the transgene expression in *C. reinhardtii* when compared to the directly isolated genes from the host organisms (Wu et al., 2011b)

However, various species have specific requirements when it comes to codon optimization of transgenes. For instance, a diatom *Phaeodactylum tricornutum* transformed with various reporter genes coding for GFP showed clear preference towards a *gfp* gene with optimal codon usage for expression in human cells (Zaslavskaia et al., 2000). In some cases, codon optimization is not required for efficient protein expression. This was proven in the experiment aiming to test the influence of codon biased coding sequences on the expression of two chimeric proteins containing zeocin binding protein Ble and codon-optimised and non-optimized variants of green fluorescent protein (GFP) introduced into the chloroplast of *C. reinhardtii* (Heitzer et al., 2007).

5.4 Conclusions and future work

The overexpression of thioesterase *FatB1* specific for 16:0-ACP in the chloroplast of *C. reinhardtii* was not successful. An alteration to this experiment would be the introduction of a thioesterase responsible for medium chain fatty acid synthesis, for instance 12:0-acyl-carrier protein thioesterase (BTE) (Voelker and Davies, 1994, Voelker et al., 1992) since C_{12} and C_{14} are the ideal FA chain lengths for biodiesel production (Kleiman, 1999). BTE thiesterase was successfully introduced into *E. coli* (Voelker and Davies, 1994) and *Arabidopsis thaliana* (Voelker et al., 1992) genomes and that resulted in metabolic alteration in FA synthesis towards the production of C_{12} and C_{14} FAs at the expense of long chain ($\geq$16) FAs, which are synthesized predominantly in both species.

Two genes responsible for the alkane synthesis in cyanobacteria: acyl reductase and aldehyde decarbonylase, were directly amplified from *Synechocystis* PCC6803 into four expression vectors designed for *C. reinhardtii* chloroplast engineering. Although both genes regulated by the same promoter/5'UTR were successfully incorporated into the chloroplast genome, and the transcription of both genes occurred under the *atpA* regulatory region, neither enzyme in the transformed line was at a detectable level. Since

codon optimization of transgenes has been associated with its increased expression (Mayfield et al., 2003, Mayfield et al., 2007, Coragliotti et al., 2011), adaptation of codon usage in acyl reductase and aldehyde decarbonylase genes from *Synechocystis* PCC6803 could be beneficial.

An alternative alkane biosynthesis pathway proposed in plants could be more suitable for *Chlamydomonas* than the pathway identified in cyanobacteria. Therefore, chloroplast transformation in *C. reinhardtii* with two vectors containing codon-optimized sequences of the *YRE* and *CER1* genes from *Arabidopsis* (Kurata et al., 2003) could be more efficient than the genes isolated from *Synechocystis* PCC6803.

Finally, the potential toxicity of alkanes on algal cultures can be overcome by applying an inducible regulatory region controlling the expression of the transgenes. Numerous types of inducible promoters were discussed earlier in this chapter therefore, the next step in optimization of alkane biosynthesis in algal chloroplasts would be creating a new construct containing a promoter activated by an external inducer.

CHAPTER 6. *CHLAMYDOMONAS REINHARDTII* AS A PLATFORM FOR HIGH VALUE PRODUCTS SYNTHESIS

6.1. Introduction

6.1.1. Human Papillomavirus (HPV)

The epitheliotropic pathogen human papillomavirus (HPV) is a double-stranded DNA tumour virus that can cause diverse types of skin lesions varying from common warts or verrucas to cervical neoplasia and cancer (Collier et al., 2002). There are over 100 types of HPV identified based on their DNA sequence where a type is determined if its sequence differs from other types by over 10% (Calleja-Macias et al., 2005, Jenson et al., 2001, Paz De la Rosa et al., 2009). All types of HPV are classified into two groups according to their high- and low-oncogenic risk status (Paz De la Rosa et al., 2009). Low-risk genital HPV types include common types 6, 11, 42, 43, and 44 and are associated with benign cutaneous warts in immunocompetent individuals although HPV-11 is also linked to cancer of the respiratory tract after irradiation (Abramson et al., 1987, Jenson et al., 2001). Conversely, malignant tumours of anogenital tracts are linked to high-risk HPV types such as types 16, 18, 31, 33, 34, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, and 70 (Kohl et al., 2007, Burd, 2003). Cervical cancer is the second most common type of cancer in women (Parkin and Bray, 2006) and HPV type 16 and 18 can be found in over 70% of all cases with type 16 being predominant as it has been found in 50% cases of all cervical cancers (Bosch et al., 2002, Paz De la Rosa et al., 2009).

The HPV capsid is composed of two types of proteins – the major capsid protein L1 and the minor capsid protein L2 that form an icosahedral capsid surrounding the viral genomic DNA (Collier et al., 2002) and they play a significant role in the efficient infectivity of the virus. Figure 6-1 presents the gene orientation within the HPV genome that contains three distinct regions: the long control region (LCR), the early region (including E1-E6 open reading frames) and the late region coding for the L1 and L2 capsid proteins.

Figure 6-1 Organisation of the HPV genome.

Adapted from (Doorbar, 2006)

The HPV genome is a double stranded DNA and is displayed as a black circle. All genes are encoded on one strand. Two gray triangles are representing promoters. Early genes (E6 and E7 in red, E1 and E2 in yellow, E4 and E5 in green) are controlled by either p97 or p670 promoters depending on stages during epithelial cell differentiation whereas late genes (L1 and L2 in blue) are expressed from p670.

Long control region (LCR) contains several binding sites for a regulatory protein E2 and E1.

Infection by papillomaviruses occurs via clathrin-dependent receptor-mediated endocytosis following viral adsorption to the cell surface (Culp and Christensen, 2004, Day et al., 2003). Minor capsid protein L2 assists in the transfer of viral DNA to the nucleus. Ultimately, the virus is transported to the nuclei of the basal layer of the epidermis where the viral DNA does not integrate with the host genome, but forms a stable episome (Day et al., 2003). This phase is associated with expression of viral replication proteins E1 and E2 that enable the initiation of viral DNA replication. Additionally, E2 functions as a DNA-binding protein within the long control region (LCR) of the viral genome and regulates its gene expression (Dell et al., 2003). Viral oncogenes E6 and E7 disrupt the host cell growth and increase the proliferation rates of suprabasal epithelial cells but also E6 acts as an anti-apoptotic protein (Burd, 2003, Doorbar, 2006). Protein E5 is a transmembrane protein predominantly occupying endoplasmatic reticulum and functions as a membrane signaling protein (Figure 6-1) whereas E4 has been identified as a protein participating in the capsid-virion assembly and release of the formed viruses from the host cells as this protein is able to disrupt the keratin network (Wang et al., 2004, Doorbar, 2006).

Major capsid protein L1 is highly conserved and when expressed in host cells it self-assembles into virus-like particles (VLPs), whereas the minor capsid protein L2 is incorporated into a capsid only if it is co-expressed with L1 (Florin et al., 2002). Self-assembly capabilities are also found when L1 is expressed in various heterologous systems (Fernandez-San Millan et al., 2008, Kirnbauer et al., 1992). Each VLP consists of 72 pentamers and each pentamer is composed of five 55kDa L1 proteins (Sapp et al., 1998). VLPs are very efficient in inducing high-titer neutralizing antibodies (Kirnbauer et al., 1992) but also they are capable of binding to and activating dendritic cells *in vivo* (Lenz et al., 2001, Rudolf et al., 2001). The ability to induce the prophylactic immune response has been harnessed during the development of an effective vaccination system against human papillomavirus.

Cytology screening introduced into UK in 1988 substantially decreased the occurrence of cervical cancer. However, approximately 2400 women per year still develop cervical cancer (Szarewski, 2010). Mortality is much higher in developing countries where the cytology screening is too expensive or not efficient and HPV-related cervical cancer is responsible for over a quarter of a million deaths per year (Bosch et al., 2002). The screening test can detect cellular abnormalities relatively early, however it is only a

secondary prevention. A primary prevention against cervical cancer in the form of HPV vaccination has been recently introduced among teenage girls in many developed countries.

There are two licensed HPV vaccines on the market and both of them contain the recombinant L1 protein of HPV. CERVARIX™ vaccine includes L1 proteins from both types HPV16 and HPV18 that were synthesised in cells of the moth *Trichoplusia ni* whereas GARDASIL™ vaccine consists of L1 proteins of HPV16 and HPV18 that are produced in the yeast *Saccharomyces cervisiae* (Szarewski, 2010). Both vaccines are highly immunogenic and safe however, they have been shown to be effective only for about 5 years (Bazan et al., 2009). Both CERVARIX™ and GARDASIL™ are expensive to produce and administer, as the cost reaches approximately $350 per person, which is unaffordable for the majority of developing countries. In order to reduce the costs of vaccines, recent reports introduced a novel idea to use plants for L1 protein production (Biemelt et al., 2003, Fernandez-San Millan et al., 2008, Liu et al., 2005, Maclean et al., 2007, Paz De la Rosa et al., 2009, Varsani et al., 2003, Waheed et al., 2011b). Recombinant vaccine proteins produced in transgenic plants can be 10-100 times cheaper than the proteins obtained in the cell culture systems currently exploited (Kusnadi et al., 1997, Mison and Curling, 2000).

In all reports where tobacco (Biemelt et al., 2003, Fernandez-San Millan et al., 2008, Varsani et al., 2003), tomato (Paz De la Rosa et al., 2009) or potato plants (Biemelt et al., 2003) were successfully transformed with HPV L1 protein gene, the chloroplast transformation vector contained an antibiotic selectable marker – either the *aadA* gene or kanamycin NPTII gene (Liu et al., 2005). Therefore, the selection of transformants was based on introducing an antibiotic resistance gene, whereas the chloroplast delivery method developed in the Purton lab is based on restoring photosynthetic activity. The first system is much more contentious than the latter as it can lead to the escape of antibiotic resistant plant strains to the environment. Additionally, transgenic proteins introduced into the crop plants can be easily spread as debris in the form of dust in the air and can directly affect humans. It has been reported that people subjected to constant exposure to the crop vaccines are very likely to develop a tolerance hence they become susceptible to the disease (Cummins, 2004). Since algal cultures are maintained in closed containers, they are considerably easier to control than crop plants. Also the growth rates of many algal species in optimal conditions are significantly higher than plants hence the

algal productivity can be potentially considerably higher than crop plants. Therefore, establishing a platform for HPV L1 protein synthesis in the algal chloroplast may have a significant impact on development of a cost-effective HPV vaccine.

6.1.2. Cyanovirin-N

Human immunodeficiency virus (HIV) has been a significant global health problem since its discovery and identification in the early 1980s. There are estimated 33.3 million HIV-positive individuals worldwide with 2.6 million new cases reported in 2009 (UNAIDS, 2008). The ongoing HIV epidemic emphasizes increasing demands for advanced prevention methods. Development of an effective vaccination method is very challenging as HIV is characterized by substantial genetic variability and the efficiency of the currently tested vaccines remains ambiguous (Cohen, 2009). Consequently, new prevention approaches have been under investigation and the development of efficient topical microbicides including compounds suppressing cellular receptors or the viral envelope have been particularly analysed (Balzarini and Van Damme, 2007). An empirical screen by the National Cancer Institute of potential microbicides resulted in the discovery of a promising anti-HIV agent cyanovirin-N, naturally produced in the cyanobacterium *Nostoc ellipsosporum* (Boyd et al., 1997).

During a sexual encounter, the infection route for HIV is through the integral layer of the epithelium cells (Pope and Haase, 2003). Consequently, HIV can infect the $CD4^+$ T-cells and dendritic cells of the genital and rectal sub-mucosa area (Veazey et al., 2003, Patterson et al., 1998). Figure 6-2 presents a schematic illustration of the interaction of the HIV envelope proteins prior to infection. The infection is initiated via the viral envelope gp120 protein (Figure 6-2.a) that binds either to the DC-SIGN receptor expressed by dendritic cells (Geijtenbeek et al., 2000) or to the CD4 receptors of T-lymphocytes and a chemokine (CXCR4 or CCR5) co-receptor (Veazey et al., 2003).

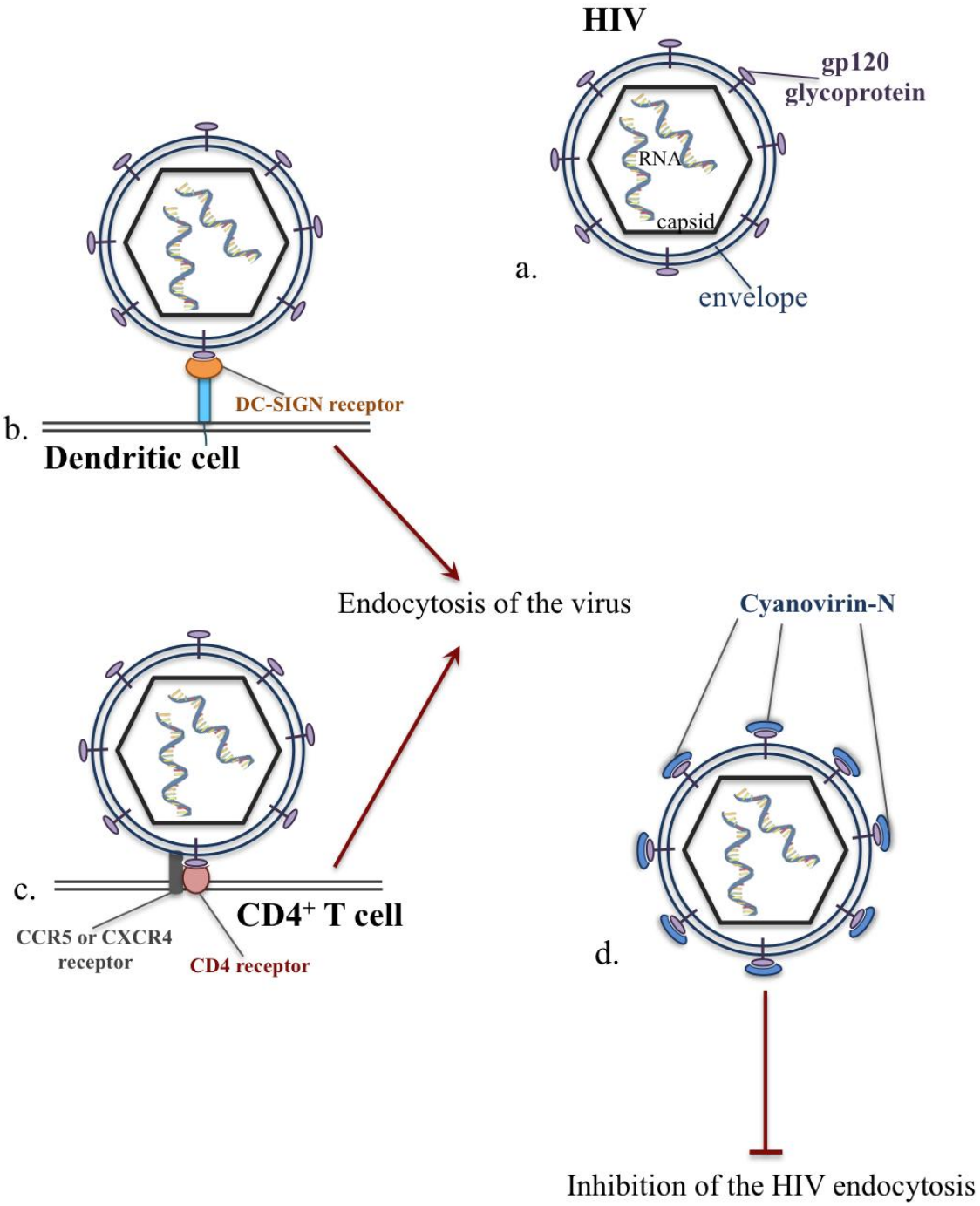


Figure 6-2 Mechanism of the gp120 receptor interaction with human immune cells.

a. Structure of the human immunodeficiency virus (HIV): two single strands of RNA are enclosed by a capsid that is surrounded by an envelope containing glycoprotein receptors gp120
b. Interaction of the gp120 receptor with the DC-SIGN receptor of a dendritic cell resulting in endocytosis of HIV
c. Interaction of the gp120 receptor with the CD4 and CCR5/CXCR4 receptors of a $CD4^+$ T cell resulting in endocytosis of HIV
d. Interaction of the gp120 receptor with cyanovirin-N molecules results in inhibition of the HIV endocytosis by dendritic and $CD4^+$ T cells

Cyanovirin-N (CVN) is an 11 kDa protein that has the capability to inactivate various HIV strains and prevent cell-to cell HIV transmission (Boyd et al., 1997, Buffa et al., 2009). CVN binds with high-affinity to high-mannose oligosaccharides present on the gp120 protein of HIV envelope (Figure 6-2.d), which prevents an interaction between the cell and virus receptors and consequently the endocytosis of HIV (Bolmstedt et al., 2001, Shenoy et al., 2001).

Cyanovirin-N in not only a potent anti-HIV protein but has been found to be also active against herpes simplex virus type 1 (HSV-1) (Tiwari et al., 2009), hepatitis C (Helle et al., 2006), Ebola virus (Barrientos et al., 2003) and almost all strains of influenza A and B virus (O'Keefe et al., 2003). In addition to reducing the viral cytopathic effect, CVN is also responsible for decreasing the viral copies in the infected host cells (Yu et al., 2010). Moreover, CVN is harmless for the host cells, resistant to physico-chemical denaturation and is characterized by a high genetic barrier to resistance (Boyd et al., 1997, Balzarini et al., 2006). All the features of cyanovirin-N listed above suggest that this protein is a very promising novel topical microbicide.

6.2. Background and aims of the project

As discussed earlier, HPV vaccines currently available on the maket are produced in eukaryotic cells therefore their substantial manufacturing costs are a consequence of the expression and maintenance charges. There are currently two strategies under development for more economical HPV vaccine production. The first one has been established in *E. coli* (Schadlich et al., 2009b). The L1 protein is synthesised in a form of pentamers (capsomers), the subunits of VLPs. Capsomers however, induce substantially lower antibody response compared to viruslike particles (Thones et al., 2008). Additionally, the protein samples of *E. coli* are contaminated with endotoxin lipopolysaccharide (LPS) that may cause pyrogenic and shock reactions and its removal can be challenging (Petsch and Anspach, 2000, Schadlich et al., 2009b). A second system has been developed in plants such as tobacco (Biemelt et al., 2003, Fernandez-San Millan et al., 2008, Varsani et al., 2003), tomato (Paz De la Rosa et al., 2009) or potato (Biemelt et al., 2003). This system is very efficient as the productivity of VLPs in leaves and can reach up to 24% of total soluble protein (Fernandez-San Millan et al., 2008) yet it is associated with regulatory and safety concerns and production is usually time consuming.

Microalgal systems are more efficient and safe compared to *E. coli* and plants hence in principle they are more suitable for the HPV vaccine production.

Similarly, cyanovirin-N has been expressed to date in bacteria (Mori et al., 1998, Giomarelli et al., 2002), yeast (Mori et al., 2002) and tobacco plants (Sexton et al., 2006). Since algal genetic engineering systems are superior compared to bacterial, yeast and plant procedures, cyanovirin-N production in algae could be the new approach for the manufacturing of this microbicide.

6.3. Results and Discussion

6.3.1.1. Expression of the transgenic HPV16-L1 gene in the *C. reinhardtii* chloroplast

6.3.1.2. Construction of chloroplast transformation vectors containing a synthetic HPV16-L1 gene

As presented in previous chapter, an expression vector developed in the Purton Lab specifically for chloroplast transformation in *Chlamydomonas* was used in these studies (Bateman and Purton, 2000). The selectable marker used in this system was restoration of photosynthetic activity of the recipient cell line BST-SAME.

Firstly, three expression vectors were created, each containing a synthetic version of *HPV16 L1* gene. The following constructs controlled by three different promoters/5'UTRs were created: pASapI (containing *atpA* promoter/5'UTR), pCSap2 (containing *chlL* promoter/5'UTR) and pSSapI (containing *psaA* promoter/5'UTR) whereas the 3'UTR sequence of the *rbcL* gene was the same for each expression vector (Ninlayarn, 2012).

In order to ensure the efficient translation of the *L1* transgene, the gene was codon optimized for the *Chlamydomonas reinhardtii* chloroplast genome and synthesized commercially by GeneArt (www.geneart.com). A Codon Adaptation Index (CAI) was calculated using the graphical codon usage software (http://gcua.schoedl.de/) and was 0.8 – a value recommended for the plastid genes that are highly expressed in *C. reinhardtii* cells (Morton, 1998). The full sequence of the HPV16-*L1* gene is provided in appendix K.

For cloning purposes, *L1* gene was flanked by *Sap*I and *Sph*I restriction sites, which enabled its incorporation into the pASapI, pCSap2 and pSSapI vectors. All created *L1* expression vectors were confirmed by sequencing. Figure 6-3 presents a schematic illustration of the cloning strategy of the *L1* gene to create the pASapI-*L1*, pCSap2-*L1* and pSSapI-*L1* constructs.

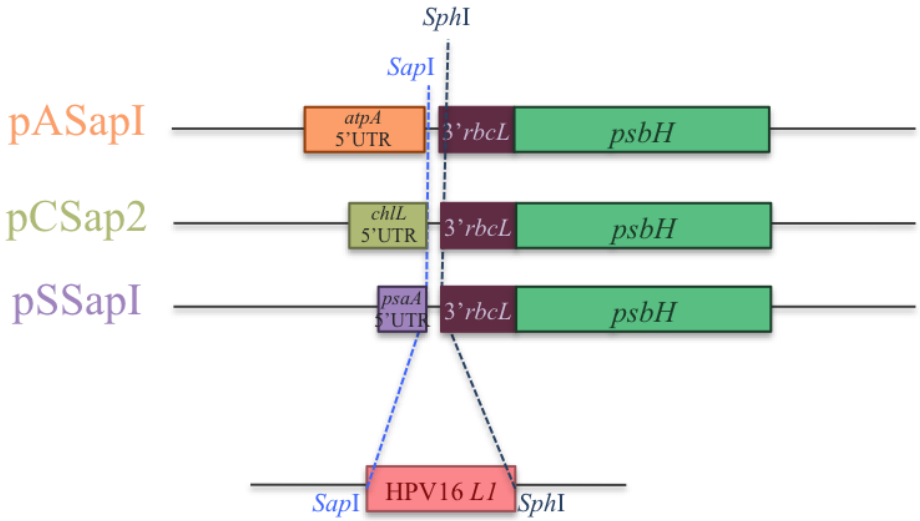


Figure 6-3 Schematic diagram of the p_Sap-*L1* expression vectors creation under the control of three different promoters/5'UTR.

Synthetic *L1* gene codon optimized for *C. reinhardtii* chloroplast flanked by *Sap*I and *Sph*I restriction sites was introduced into the *Sap*I-*Sph*I site of the three vectors: pASapI, pCSap2 and pSSapI.

6.3.1.3. Chloroplast transformation of *C. reinhardtii* with the created L1 expression vectors and analysis of the L1 transformants

The created pASapI-*L1*, pCSap2-*L1* and pSSapI-*L1* plasmids were introduced into the chloroplast of the recipient cell line BST-SAME using the biolistic particle method as described in chapter 5.

Transformants of each construct were isolated and subjected to three cycles of re-streaking to ensure homoplasmy. One isolate of each construct was chosen for the further analysis. Firstly, PCR analysis of the transformants was performed to confirm the presence of the transgene using the FLANK1 and psbH.R primers and Figure 6-4.a presents the results. The obtained results matched the expected outcomes as bands of the following sizes were amplified: for wild type of approximately 1.2 kb, BST-SAME of approximately 2.4 kb, pASapI-L1 transformant of approximately 3.5 kb, pCSap2-L1 transformant of approximately 3.3 kb and pSSapI-L1 transformant of approximately 3.2 kb; these PCR results were confirmed by sequencing.

Subsequently, the phenotype of the transformed *C. reinhardtii* cell lines was tested using western blot analysis and the results are presented in Figure 6-4.b. Whole-cell extract of the transformant samples (pASapI-*L1*, pCSap2-*L1* and pSSapI-*L1*) and negative controls (wild type and BST-SAME recipient cell line) were probed with a commercial anti-L1 antibody. Antibodies to the photosystem II (PSII) reaction center D1 protein functioned as a loading control. The 32 kDa band characteristic for D1 protein was present on the membrane after probing with the anti-D1 antibody however, the 56 kDa L1 protein was below the detectable level in all the transformed cell lines.

In order to test for the transcription of the *L1* transgene, reverse transcription PCR (RT-PCR) analysis was performed and the results are presented in Figure 6-4.c. This assay allows detection of the mRNA as it functions as a template for cDNA synthesis using the reverse transcriptase enzyme. Subsequently, the cDNA serves as a template for PCR. Total RNA was extracted from the transformants samples and an additional DNase treatment was performed in order to ensure the RNA samples were free from genomic DNA (gDNA).

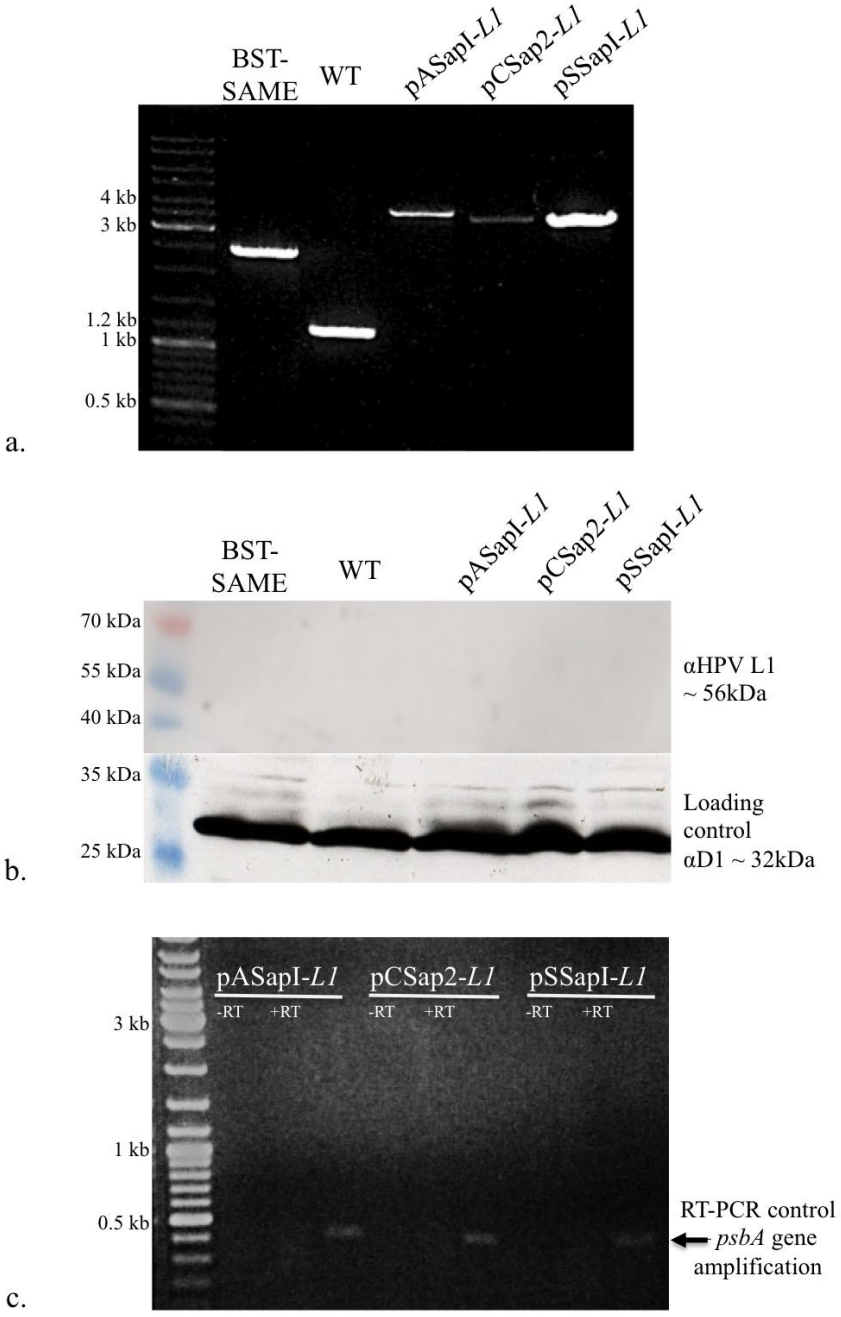


Figure 6-4 Analysis of the *L1* transformants of *C. reinhardtii* under control of three different promoters/5'UTRs.

a. PCR analysis of the *L1* transformants using FLANK1 and psbH.R primers. Expected bands are as follows: for wild type (**WT**) ~1.2 kb, **BST-SAME** ~ 2.4 kb, pASapI-*L1* transformant ~3.5 kb, pCSap2-*L1* transformant ~3.3 kb, pSSapI-*L1* transformant ~3.2 kb; PCR results were confirmed by sequencing.

b. Western blot analysis of the whole cell extracts for the presence of the HPV16-L1 protein. The expected band for L1 protein was of approximately 56 kDa whereas for the loading control D1 of approximately 32 kDa.

c. Reverse transcriptase PCR (RT-PCR) analysis of the p_SapI-*L1* transformants to confirm the transcription of the incorporated HPV16-L1 protein. Expected band for the transformed cell lines is ~1.5 kb. **+RT** – PCR amplification with cDNA operating as a template; **–RT** – negative control for contamination of RNA with genomic DNA where the isolated RNA functioned as a template for the PCR; RT-PCR of ***psbA*** acted as an internal control to test the quality of cDNA.

RT-PCR analysis of the three *L1* transformants was performed using the following primers: RT_L1.F (GTTACGAAAACGACGTTAACGTTTACC) and RT_L1.R (TGATGTAGATGATGTTGTTGGTGTTGC). Expected bands were of approximately 1.5 kb however, no RT-PCR product for the transgenic mRNA was detected (Figure 6-4.c).

The *psbA* chloroplast gene was selected as an internal control for the RT-PCR reaction as it contains introns and this feature is very rare for chloroplast genes (Maul et al., 2002). Primers psbA.F (ATGACAGCAATTTTAGAACGTCGTG) and psbA.R (ACCCATGTAGCAGTATACACCTAG) were designed and previously tested in the Purton lab (Ninlayarn, 2012). A 0.4 kb band was characteristic for the cDNA samples (marked as +RNA in Figure 6-4) whereas the presence of a 1.6 kb band would indicate the tested sample was contaminated with genomic DNA.

The absence of any detectable *L1* mRNA is most likely a consequence of post-transcriptional events occurring in algal chloroplasts. The mRNA is most likely being synthesized since pASapI, pCSapI and pSSapI have been successful in expressing other transgenes (Ninlayarn, 2012), but the *L1* mRNA is probably being rapidly degraded. This process was discussed in details in Chapter 4 and Chapter 5.

The expression of the HPV16 L1 protein in plant systems appears to be very challenging and it seems certain modifications within the *L1* gene are required for its correct transcription and translation. For instance the expression of the functional VLPs of HPV16 in tobacco plants was achieved only when *L1* was adapted towards the human codon usage whereas the otimisation towards tobacco codon usage did not result in accumulation of the L1 protein (Maclean et al., 2007). Transgenic L1 production in yeast was also substantially enhanced from the human codon-optimised gene construct (Kotze et al., 2011). L1 transcripts in transgenic plants appear to be degraded for human- and plant- codon optimized versions of the *L1* gene however, introduction of the translational enhancer Omega derived from the tobacco mosaic virus substantially improved the transgene expression (Biemelt et al., 2003). In some instances a modification of the plant codon-optimised *L1* sequence was needed. For example removal of the 21 carboxyl-terminal amino acids (Warzecha et al., 2003) or 34 C-terminal amino acids (Muller et al., 1997, Paz De la Rosa et al., 2009) resulted in enhanced L1 productivity. Similarly, modifications within the *L1* sequence are also required for its correct and efficient expression in *E. coli* (Schadlich et al., 2009a).

6.3.1.4. Attempts to improve the expression of the L1 protein in the *C. reinhardtii* chloroplast

Numerous recombinant proteins have been produced in the *Chlamydomonas* chloroplast by fusing the coding sequence of the protein to *cis* elements (promoter/5'UTR and 3'UTR) from endogenous chloroplast genes such as *psbA, atpA* and *rbcL* (Franklin and Mayfield, 2004, Bateman and Purton, 2000). However, the steady-state amount of recombinant protein is typically less than a few percent of the total soluble protein in the cell, which is significantly less than the amount of PsbA, AtpA, or RbcL protein. Whilst greater proteolytic degradation of the recombinant protein may account in part for this difference, it is clear that the rate of synthesis is also far from optimal.

One suggestion is that *cis* elements immediately downstream of the start codon of endogenous genes are important for efficient RNA synthesis and/or for RNA stability and efficient translation. These elements would be absent from the synthetic transgene constructs. Evidence for this idea comes from reports where various *N*-terminal coding regions such as *rbcL, psbA, psbD* and *atpA,* were included in GUS reporter gene constructs, thereby creating translational fusions (Kasai et al., 2003). Each expression vector applied in this project contains a promoter/5'UTR sequence and was designed to not generate any *N*-terminal extension. Since the best yield of the reporter GUS protein was achieved from an *atpA* promoter/5'UTR construct that included additionally 25 codons of coding sequence (Kasai et al., 2003), it was therefore decided to modify the pASapI, pCSap2 and pSSapI expression vectors by including the 25 codons of the *atpA* gene (CDS25) ahead of the starting codon of the *L1* sequence.

Furthermore, the difficulties with the *L1* gene expression in the plant and bacterial cells inspired this project to introduce additional modifications within its coding sequence. Previous work on efficiency of the HPV16 L1 protein expression in the host cell lines resulted in discovering and mapping its various inhibitory elements (Collier et al., 2002). It was established that there are numerous independently acting inhibitory sections and all of them are positioned within the first 514 nucleotides of the *L1* gene. Moreover, deletion of the first 10 codons of the L1 gene appears to increase the capsomer synthesis in *E. coli* and this modification did not affect the L1 structure and effectively increased its immunogenicity (Schadlich et al., 2009a). The L1 sequence was therefore re-designed and 10 amino acids from the N-terminal end of this protein was removed.

In order to introduce the CDS sequence (ATGGCAATGCGTACTCCAGAAGAACTTAGT AATCTTATTAAAGATTTAATTGAACAATACACTCCAGAAGTGAAA) to the existing synthetic *L1* gene, the sets of primers suitable for 2-step nested PCR were designed. Figure 6-5 presents the cloning steps during the assembly of the modified *L1* expression vectors.

The original synthetic construct obtained from GeneArt functioned as a template for the initial PCR amplification and the following primers were used: to amplify the unmodified *L1* gene (*L1*-CDS25) 1_L1_CDS25.F (GAAGAACTTAGTAATCTTATT AAAGATTTAATTGAACAATACACTCCAGAAGTGAAAATGCAAGTTACATTCATTTAC) and L1_CDS25.R (CGAATTCGCATGCTTATTATAATTTACG) were used whereas to amplify the truncated L1 version (ΔN_*L1*-CDS25) 1_L1_CDS25_dN.F (GAAGAACTTAGTAATCTTATTAAAGATTTAATTGAACAATACACTCCAGAAGTGAA AATGATTACATGTTACGAAAACGACG) and L1_CDS25.R were used. The PCR product of the first amplification reaction functioned as a template for the second PCR and the following primers were used for both versions of the *L1* gene: 2_L1_CDS25.F (TGCTCTTCAATGGCAAT GCGTACTCCAGAAGAACTTAGTAATCTTATTAAAG) and L1_CDS25.R. 2_L1_CDS25.F and L1_CDS25.R were designed to create *Sap*I and *Sph*I restriction sites at the flanking ends to enable cloning the amplified *L1*-CDS25 and ΔN_*L1*-CDS25 sequences into the pASapI and pSSapI expression vectors (Figure 6-5).

Figure 6-6 presents the diagrams of the seven modified and unmodified *L1* constructs that were used in the further experiments: pASapI-*L1*, pASapI-*L1*-CDS25, pASapI- ΔN_*L1*-CDS25, pCSap2-*L1*, pSSapI-*L1*, pSSapI-*L1*-CDS25 and pSSapI-ΔN_*L1*-CDS25.

Both *L1*-CDS25 constructs contain two start codons – introduced at the beginning of CDS25 but also the L1 sequences whereas two of the ΔN_*L1*-CDS25 constructs contain only a single start codon that is placed at the beginning of the CDS25 sequence.

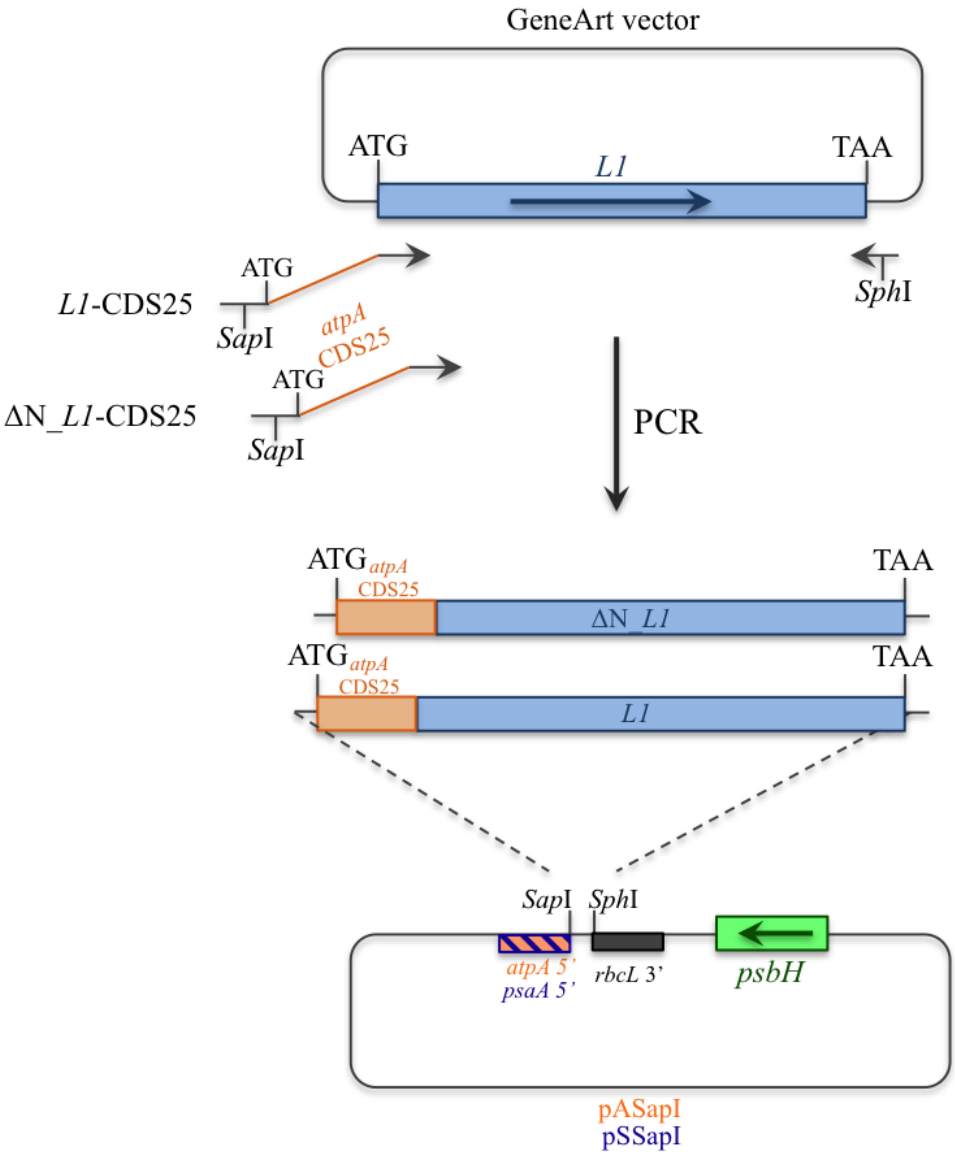


Figure 6-5 Cloning strategy of the modified L1 construct.

Two sets of forward primers introducing the first 25-codon sequence of the *atpA* gene (*atpA* CDS25) were designed: *L1*-CDS25 amplified the whole L1 sequence whereas ΔN_*L1*-CDS25 was designed to amplify the *L1* gene without the first 10 codons. PCR amplification resulted in creation of two modified *L1* genes with incorporated *atpA* CDS25 sequence and two flanking restriction sites *Sap*I and *Sph*I that enabled their cloning into pASapI, pCSap2 and pSSapI expression vectors.

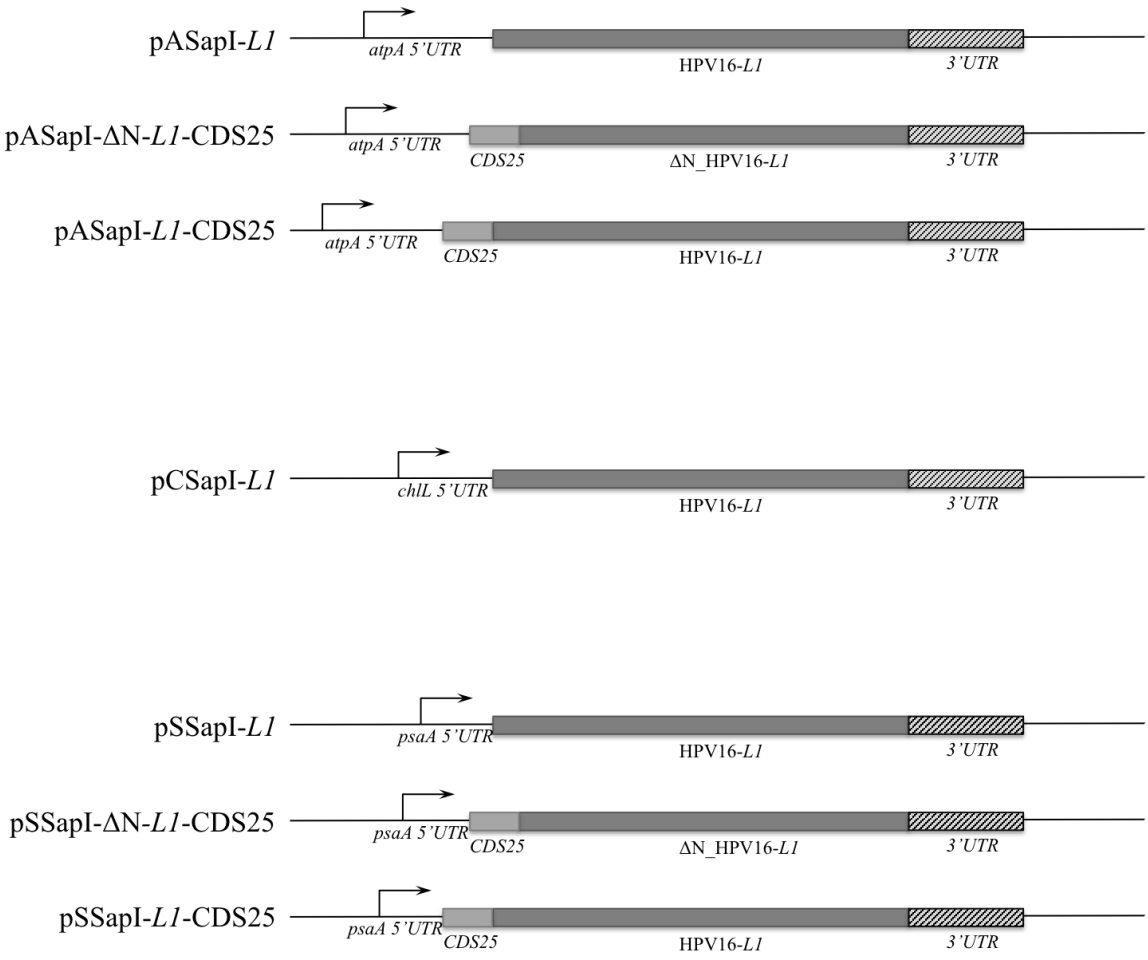


Figure 6-6 Schematic diagrams of the modified *L1* expression vectors.

Synthetic *L1* gene codon optimized for the *C. reinhardtii* chloroplast was introduced into the three vectors: pASapI (controlled by *atpA* promoter/5'UTR), pCSap2 (controlled by *chlL* promoter/5'UTR) and pSSapI (controlled by *atpA* promoter/5'UTR). Additionally pASapI-*L1* and pSSapI-*L1* constructs were further modified: supplemented with **CDS25** sequence – first 25 codons of the *atpA* gene, **ΔN** – first 10 codons of the *L1* gene were removed.

6.3.1.5. Testing the created L1 expression vectors in *E. coli*

The inhibitory sequences within the first ten codons of the L1 gene particularly influenced the protein expression in *E. coli* (Collier et al., 2002). In order to test the functionality of the seven *L1* expression vectors in *E. coli*, bacterial cells were transformed with the created constructs. Whole cell extracts were subjected to western blot analysis where samples were probed with the anti-L1 antibody and the results of this analysis are presented in Figure 6-7. *E. coli* transformed with pSSapI, pCSap2 and pASapI functioned as negative controls. Two unspecific bands of approximately 36 kDa and 70 kDa were present in all tested cell lines and the 36 kDa band was chosen to serve as a loading control. Interestingly, only the cell line transformed with pSSapI-ΔN_*L1*-CDS25 gave a distinct band at the expected 56 kDa size for L1, whereas the remaining *L1* transformants displayed a very faint presence of this protein. It is clear that removing the inhibitory region from *L1* sequence contributed towards improved efficiency of the L1 protein expression although this was observed only under the *psaA* regulatory region. Furthermore, all L1 transformants showed the presence of various unexpected bands of size between 36 kDa and 56 kDa that were absent in all three negative controls (Figure 6-7). Intriguingly, the concentration of protein of approximately 38-39 kDa appeared to be substantially higher in pASapI-*L1*, pASapI-*L1*-CDS25, pASapI-ΔN_*L1*-CDS25 and pCSap2-*L1* transgenic *E. coli* cell lines compared to pSSapI-*L1*, pSSapI-*L1*-CDS25 and pSSapI-ΔN_*L1*-CDS25.

The 36 kDa–56 kDa bands detected with the anti-L1 antibody presented in Figure 6-7 are presumably degradation intermediates of the full-length L1 protein as no such bands are present in the negative control samples. Specific proteases abundant in *E. coli* cells target various abnormal proteins for instance incomplete proteins, free subunits of multimeric complexes, misfolded or genetically engineered proteins (Goldberg, 2003). Misfolded proteins however, are the main target of proteases as the critical issue is accessibility of the recognition elements within the degraded protein sequence (Gottesman, 1996). The cellular protein folding process is assisted by molecular chaperones – proteins assisting in the folding or unfolding of the newly synthesized proteins (Feldman and Frydman, 2000) however chaperones are also frequently involved in the degradation process of the misfolded proteins (McClellan and Frydman, 2001).

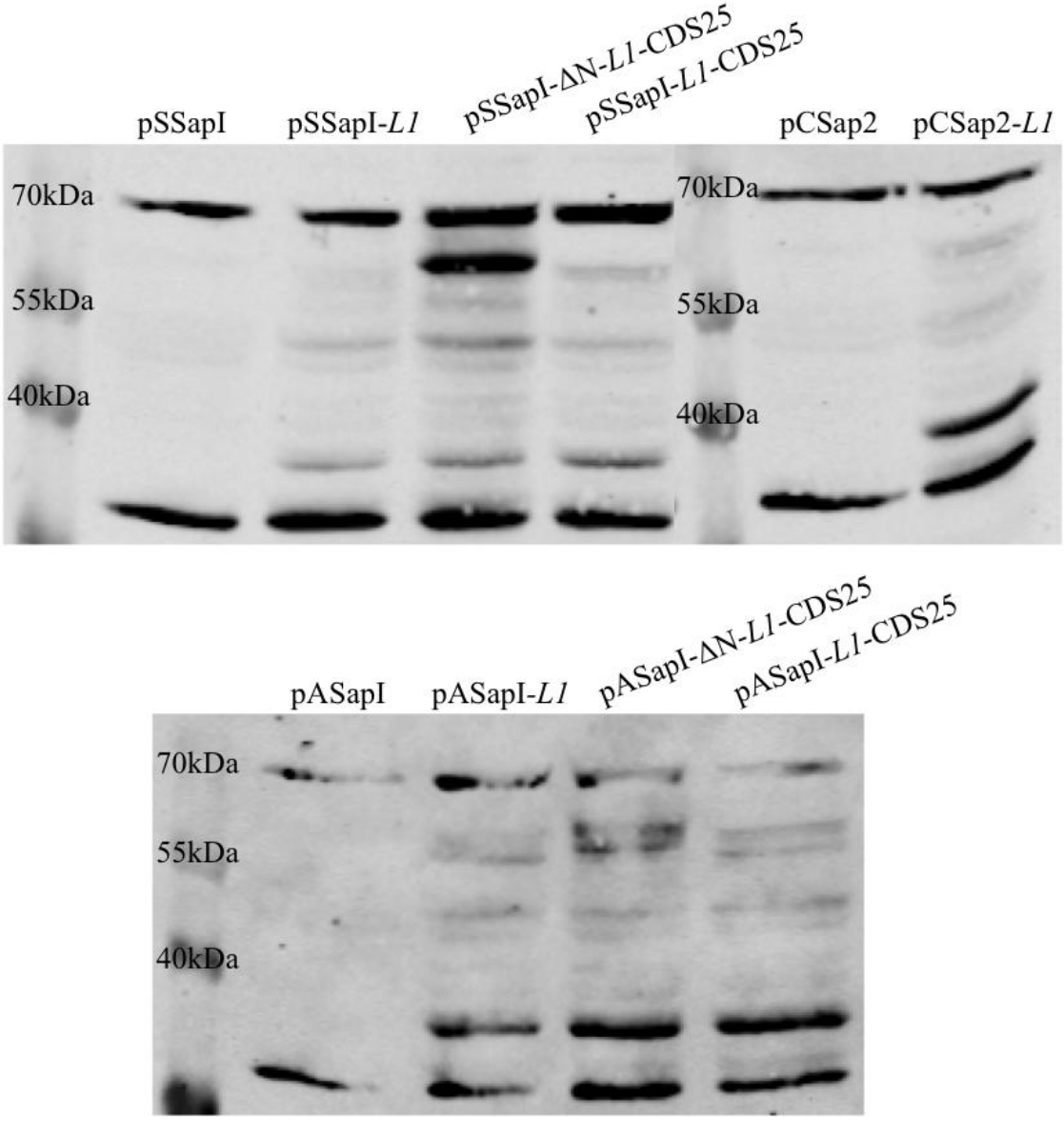


Figure 6-7 Western blot analysis of the *E. coli* cell lines transformed with the modified *L1* expression vectors.

Whole cell extracts of the ten transgenic cell lines were subjected to probing with an anti-L1 antibody. Expected size of the L1 protein is approximately 56 kDa. Unspecific bands of approximately 36 kDa functioned as loading control. Unspecific bands smaller than 56 kDa and larger than 36 kDa are presumably a result of the L1 protein degradation.

E. coli transformed with pSSapI, pCSap2 and pASapI functioned as negative controls.

It has been demonstrated in *E. coli* that proteins containing specific amino acids at the amino-terminus are rapidly degraded (Tobias et al., 1991) hence the amino acids affecting protein stability are normally very rarely present at the N-terminus (Miller et al., 1987). Similarly, a sequence of the last five amino acids at the carboxy-terminus is also crucial for the protein stability, as it has been demonstrated that polar or charged amino acids have stabilizing effects on the newly synthesized protein (Parsell et al., 1990). Analogously, the presence of specific short sequences in eukaryotic cells such as KDEL at a unique position of C- or N-terminus are also sufficient for proteolytic targeting of a correctly expressed protein (Gottesman, 1996).

Modification within the N-terminus of the *L1* sequence resulted in efficient expression of the altered L1 protein in *E. coli* under the *psaA* promoter/5'UTR. Additional changes in the *L1* sequence has been proven to further induce L1 yield in the transgenic *E. coli*, for instance deletion of 29 codons from the carboxyl-terminal resulted in 2-fold increase of L1 concentration in the bacterial cells without affecting its pentamer formation (Schadlich et al., 2009a).

6.3.1.6. Chloroplast transformation of *C. reinhardtii* with the modified L1 plasmids and analysis of the L1 transformants

The TN72 recipient cell line was transformed with the created L1 constructs (Figure 6-6) using the method where the recipient cell line was vortexed with the plasmid DNA in the presence of glass beads. Afterwards the TN72 cells were plated on agar plates containing minimal medium. Single colonies appeared approximately after 3-4 weeks and were subjected to three rounds of re-streaking to single colonies on fresh agar plates.

Several transformed colonies were isolated, one isolate per construct was subjected to PCR analysis with FLANK1 and psbH.R primers and the results are presented in Figure 6-8.a. The recipient cell line TN72 transformed with pASapI (TNEmp) functioned as a negative control. PCR amplification of the tested transformants confirmed the correct integration of the L1 expression cassettes as it resulted in PCR products of the expected size: for the negative control, TNEmp, approximately 2.3 kb, pCSap2-L1 transformant approximately 3 kb, pASapI-L1 and pASapI-ΔN_L1-CDS25 transformants approximately 3.2 kb, pASapI-L1-CDS25 transformant approximately 3.15 kb, pSSapI-

L1 and pSSapI-ΔN_L1-CDS25 transformants approximately 2.9 kb, pSSapI-L1-CDS25 transformant approximately 2.85 kb. The PCR results were confirmed by sequencing.

Subsequently, western blot analysis of the whole cell extracts of the seven transformants and also the negative control, TNEmp, was carried out and the results are presented in Figure 6-8.b and in Figure 6-8.c. All cultures prior to harvesting were maintained under continuous illumination except for the pCSap2-L1 transformant, which was cultured in the dark, as well as in the light. The pCSap2 construct was generated with the intention of creating an inducible expression vector activated in the dark and suitable for chloroplast transformation of *Chlamydomonas reinhardtii* (see Chapter 5). Western blot analysis confirmed all the transformed cell lines expressed the *L1* transgene to different degrees, as the band of approximately 56 kDa was present in all transformed cell line samples. Interestingly, two transformed cell lines pASapI-ΔN_L1-CDS25 and pSSapI-ΔN_L1-CDS25 showed the presence of protein larger than the expected 56 kDa. The approximate CDS25 size is 2 kDa whereas the modified ΔN_L1 has been estimated to be 55.3 kDa (Schadlich et al., 2009a). During the ΔN_L1-CDS25 cloning (Figure 6-5) the start codon of L1 protein was destroyed hence its translation was driven from the start of the CDS25 sequence and both L1 proteins of the ΔN_L1-CDS25 constructs are larger than unmodified L1.

Figure 6-8.c displays the expression levels of the transgenic L1 protein. Deletion of 10 amino acids from the N-terminus of L1 substantially increases its accumulation in transgenic chloroplasts particularly in the pASapI constructs where the expression of the ΔN_L1-CDS25 version was increased two-fold compared to the unmodified L1 and L1-CDS25. Additionally, the *chlL* regulatory region appears to be functional as an inducible promoter since there is clear induction of L1 protein synthesis when the transformed pCSap2-*L1 Chlamydomonas* cell line was maintained in dark. Diverse levels of transgenic L1 protein in transformed cell lines can be a reflection of either dissimilar efficiency of its expression but also its stability in chloroplasts. It has been verified that protein stability can frequently be the limiting factor in accumulation of transgenic proteins in chloroplasts (Birch-Machin et al., 2004) however, transcript stability can also substantially influence expression of transgenes within chloroplasts (Wurbs et al., 2007).

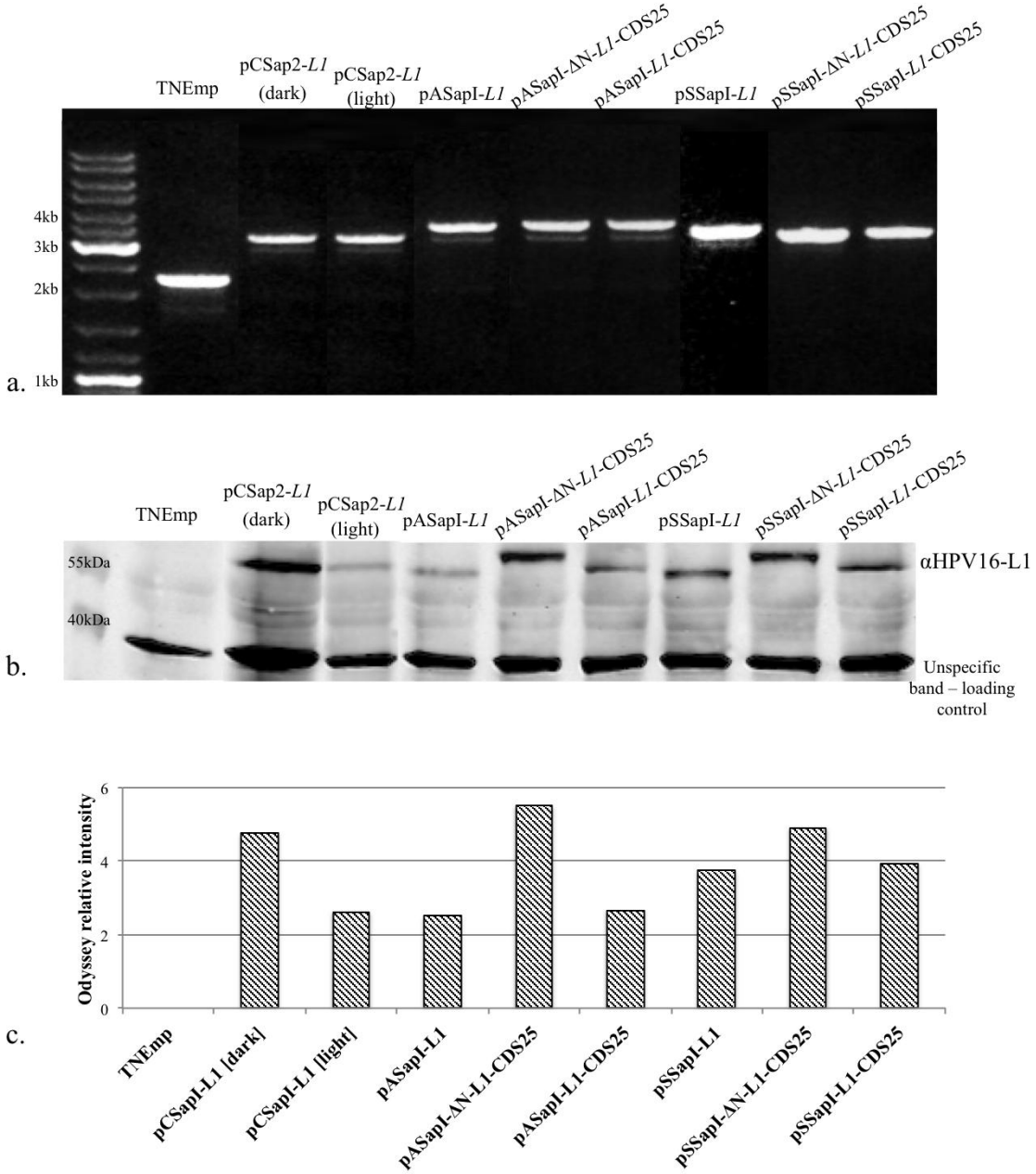


Figure 6-8 Analysis of the transformants of *C. reinhardtii* transformed with the modified *L1* expression vectors.

a. PCR analysis of the transformed cell lines with eight modified *L1* constructs with psbH.R and FLANK1 primers. The expected bands are as follows: TNEmp ~ 2.3 kb, pCSap2-*L1* ~ 3 kb, pASapI-*L1* and pASapI-ΔN_*L1*-CDS25 ~ 3.2 kb, pASapI- *L1*-CDS25 ~ 3.15 kb, pSSapI-*L1* and pSSapI-ΔN_*L1*-CDS25 ~ 2.9 kb, pSSapI-*L1*-CDS25 ~ 2.85 kb. The PCR results were confirmed by sequencing.
b. Western blot analysis of the transformed cell lines with eight modified *L1* constructs. The expected bands are of approximately 56 kDa.
c. Semi-quantitative analysis of the L1 concentration in the transformed cell lines.

Negative control **TNEmp** – TN72 recipient cell line of *C. reinhardtii* transformed with pASapI expression vector.

pCSap2-*L1* transformants were maintained either under constant illumination (light) or in dark whereas all other transformed cell lines were cultured under constant illumination.

The approach where the *L1* sequence was fused with the short sequence of an abundant protein in order to increase the efficiency of the recombinant L1 expression in the transgenic plants has been widely applied. It has been demonstrated that expression of the HPV16 L1 capsid protein in transplastomic plants was achieved only with N-terminal translational fusion of this protein with the short sequence of the plastid encoded photosynthetic proteins (Lenzi et al., 2008).

A similar idea to link two protein sequences, an antigen with an adjuvant, has recently been announced (Waheed et al., 2011a). In this report a successful expression of a fusion protein containing two pentameric proteins: an adjuvant-like heat-labile enterotoxin subunit B (LTB) of *E. coli* and modified HPV-16 L1 in which two cysteines at position 175 and 428 were replaced by serine residues was validated in tobacco plants. The created fusion protein was demonstrated to increase L1 immunogenicity but also, by combining production of two vaccine components, it has potential to reduce the administering costs of the L1 vaccine.

Creating fusion proteins for chloroplast expression in *Chlamydomonas reinhardtii* has also been applied in order to increase the level of a recombinant protein. With the aim to increase the luciferase yield in the transformed *C. reinhardtii*, a fusion protein containing a native protein-processing site from preferredoxin (preFd) positioned between the Rubisco LSU and luciferase coding regions was created and as a result the modified luciferase was accumulated in algal cells up to 33 times greater level compared to the unmodified luciferase (Muto et al., 2009). preFd functioned as a proteolytic processing site that enabled release of the exogenous protein from the endogenous fusion partner following endogenous cleavage by the chloroplast's stromal processing peptide.

6.3.2. Expression of the transgenic cyanovirin (CVN) in *C. reinhardtii* chloroplast

6.3.2.1. Construction of the chloroplast transformation vectors containing the synthetic cyanovirin-N (CVN) gene and its native version from *Nostoc sp*.

Firstly, the following four expression vectors were created: two versions of pASapI containing *atpA* promoter/5'UTR and two versions of pSSapI containing *psaA* promoter/5'UTR (Ninlayarn, 2012), including either a synthetic version of cyanovirin-N gene (sCVN) or the native cyanovirin-N gene amplified from *Nostoc sp* (oCVN) (Figure 6-9). oCVN was amplified from the pTrakCVN plasmid kindly provided by Professor Julian Ma of St George's University of London using the following primers: oCVN.F (GTGGCTCTTCGATGCTTGGTAAATTCTCCCAGACC) and oCVN.R (GCGGCATGCTTATTATTCGTATTTCAGGGTACC). Both primers were designed with the purpose of introducing an ATG codon at the 5'end and an additional TAA codon at the 3' end of the native version of cyanovirin gene. The synthetic form of the CVN gene was codon adapted for the *C. reinhardtii* chloroplast genome and synthesized commercially by GeneArt (www.geneart.com). Codon Adaptation Index (CAI) was calculated using the graphical codon usage software (http://gcua.schoedl.de/) and was approximately 0.8 as described for the L1 gene (Morton, 1998). Full sequences of both versions of cyanovirin-N, oCVN and sCVN are provided in Figure 6-9. For cloning purposes, the oCVN and sCVN genes were flanked by *Sap*I and *Sph*I restriction sites so enabling their incorporation into the pASapI and pSSapI vectors (Figure 6-10). All the created CVN expression vectors were confirmed by sequencing. Figure 6-10 presents a schematic illustration of the cloning strategy of the *L1* gene within the pASapI-*L1*, pCSap2-*L1* and pSSapI-*L1* expression vectors.

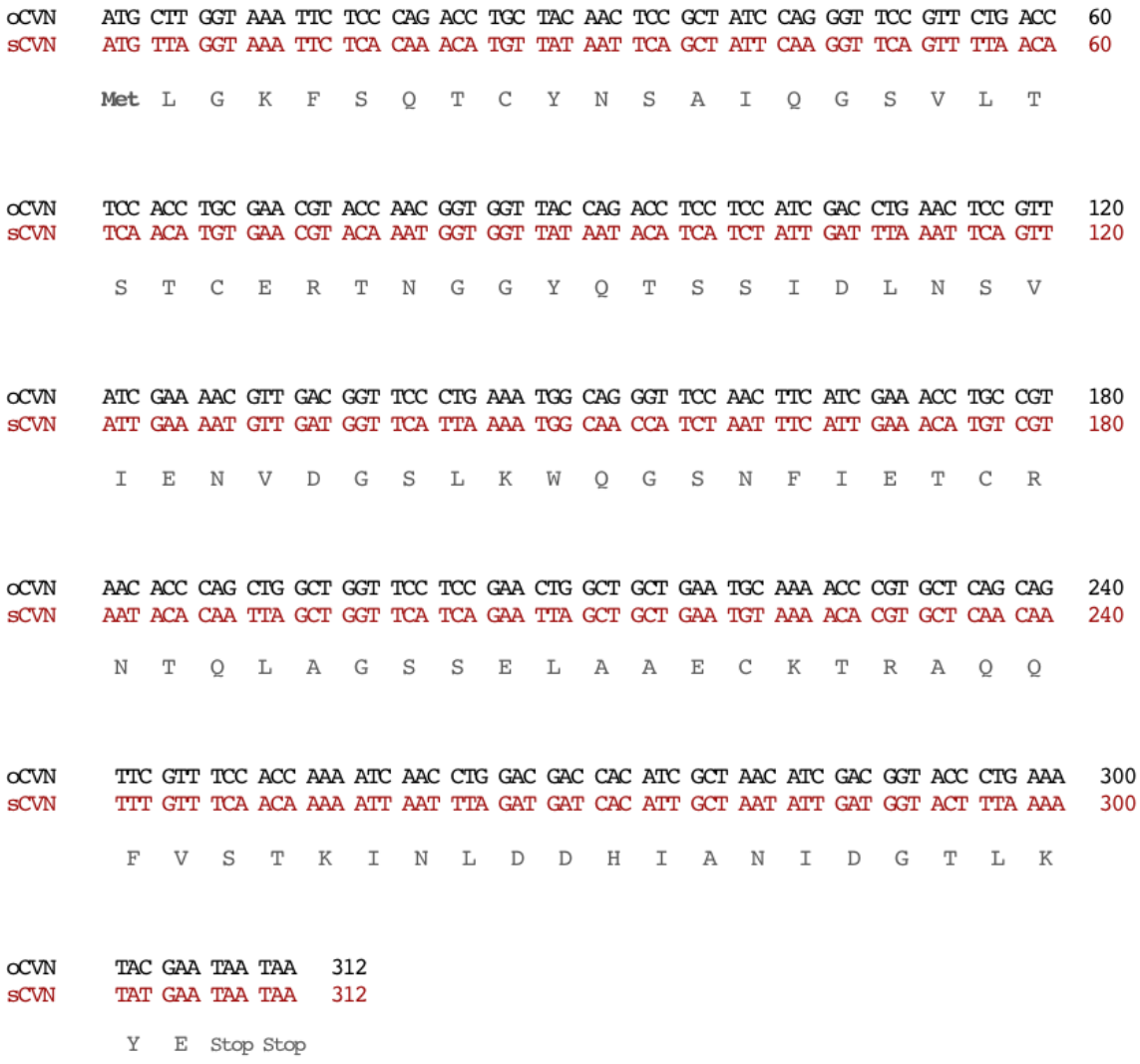


Figure 6-9 Alignment of two versions of the cyanovirin-N (CVN) sequence.

A synthetic version of CVN (sCVN) from cyanobacterium *Nostoc sp*. was synthesized based on *Chlamydomonas* chloroplast codon usage and compared to its native type (oCVN).

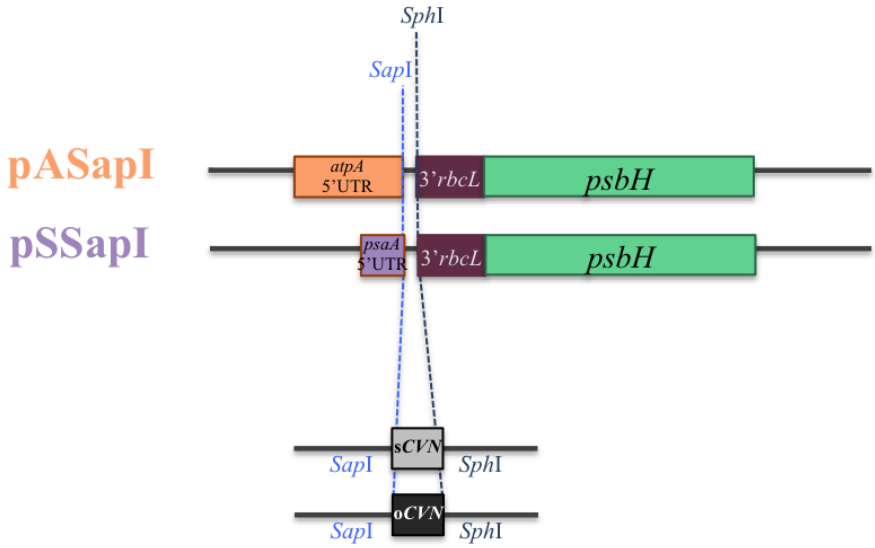


Figure 6-10 Schematic diagram of the CVN plasmids under control of two different regulatory regions.

Two types of the *CVN* gene introduced into the *Sap*I-*Sph*I site of the two vectors: pASapI, and pSSapI: synthetic *CVN* codon optimized for *C. reinhardtii* chloroplast (s*CVN*) and *CVN* amplified from *Nostoc sp* .

6.3.2.2 Expression of the CVN genes in the *C. reinhardtii* chloroplast – does codon optimization improve transgene expression?

The TN72 recipient cell line was transformed with the created CVN constructs (Figure 6-10) using the glass beads transformation method. Afterwards recipient cells were plated on agar plates containing minimal medium. Single colonies appeared approximately after 3-4 weeks and were subjected to three rounds of re-streaking on fresh agar plates.

Several transformed colonies were isolated and one isolate per construct was subjected to PCR analysis with FLANK1 and psbH.R primers and the results are presented in Figure 6-11.a. The wild type and TNEmp strains functioned as negative controls. PCR amplification of the tested transformants confirmed the correct incorporation of the CVN expression cassettes since it resulted in the creation of PCR products of expected sizes:

for the wild type approximately 1.2 kb, for the negative control, TNEmp, approximately 2.3 kb, for the pASapI-oCVN and pASapI-sCVN transformants approximately 2.5 kb, for the pSSapI-oCVN and pSSapI-sCVN transformants approximately 2.3 kb. The results were confirmed by sequencing.

Subsequently, western blot analysis of whole cell extracts was performed using the ECL detection system and the results are presented in Figure 6-11.b. An additional negative control – TN72 recipient cell line was also analysed. Cyanovirin-N is a small protein of 11 kDa and unfortunately free chlorophyll in the samples migrates in this region. As a result, the chlorophyll-protein complexes caused bleaching of the membrane probed with the anti-CVN antibody where ECL was applied as the detection system. The chlorophyll also interfears with the Odyssey® detection technique as its fluorescence obscures the signal from the secondary antibody designed for this detection system. Consequently, an additional step to remove chlorophyll, where samples were treated with acetone was applied. The final western blot is presented in Figure 6-11.c. Provisional analysis indicates at least three transformants produce recombinant cyanovirin-N at detectable levels: under the *atpA* regulatory region both versions of this gene, synthetic (pASapI-sCVN) and native (pASapI-oCVN), are well expressed whereas under *psaA* promoter/5'UTR only the synthetic version (pSSapI-sCVN) is well expressed. Unfortunately, an unspecific band of a similar size is also present in all negative controls WT, TNEmp and TN72, which makes this evaluation more challenging. Additionally, I was unable to detect purified cyanovirin-N acting as a positive control.

The expression of cyanovirin-N in chloroplasts of transgenic plants has been very challenging (Elghabi et al., 2011). Low expression of transgenic CVN in tobacco plastids was associated with two factors contributing to a similar degree: low mRNA stability and protein stability. In order to overcome this problem they fused the CVN sequence with either GFP or PlyGBS sequences as either C- or N-terminus extensions. As a result, it was concluded that stabilization of both termini are required for efficient CVN expression fused with GFP, whereas fusion of CVN with PlyGBS is essential only at the N-terminus.

To date, only a few reports of successful attempts to express cyanovirin-N in chloroplasts of transgenic plants have been reported (Sexton et al., 2006, Sexton et al., 2009, Elghabi et al., 2011). It appears that sequence modifications are required for efficient cyanovirin-N expression in plants.

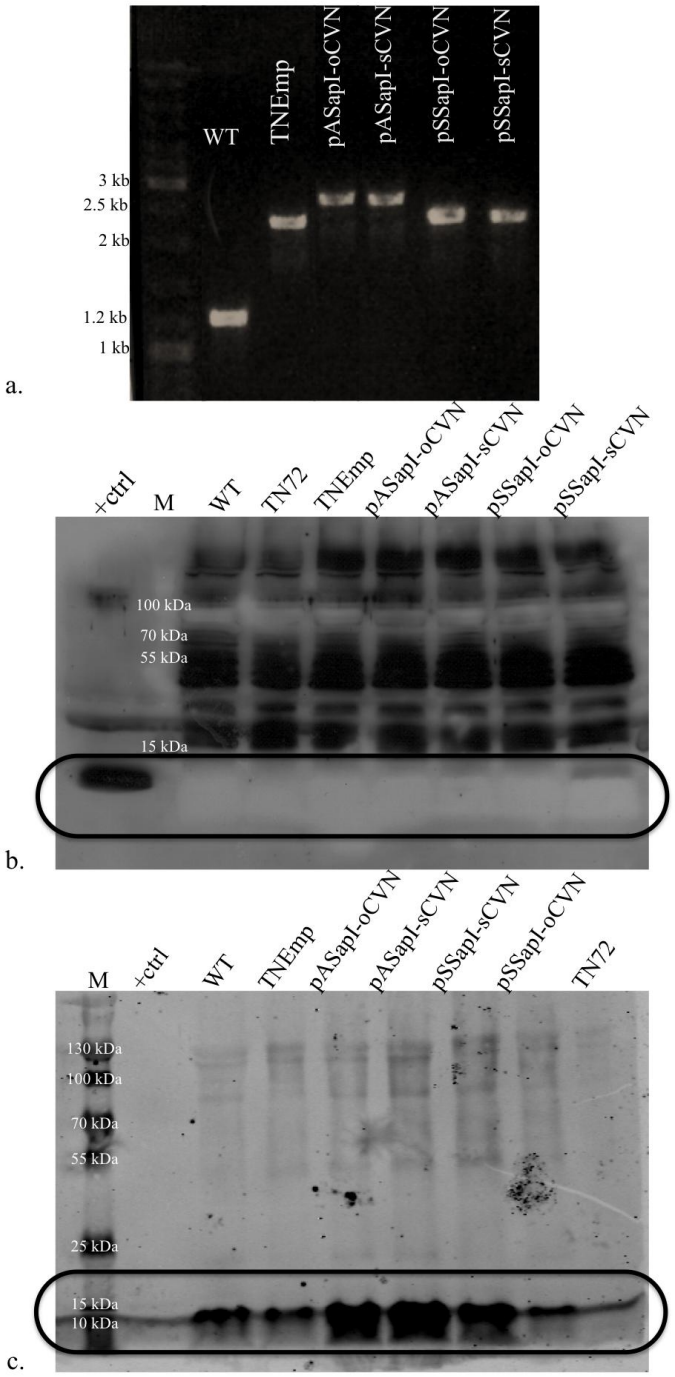


Figure 6-11 Analysis of the *C. reinhardtii* isolates transformed with cyanovirin-N (CVN).

a. PCR analysis of the transformants using FLANK1 and psbH.R primers. Expected bands are as follows: WT ~1.2 kb, TNEmp ~ 2.4 kb, pASapI-o/sCVN transformants ~2.5 kb, pSSapI-o/sCVN transformants ~2.4 kb. PCR results were confirmed by sequencing.
b. ECL-based western blot analysis of the whole cell extracts of the CVN transformants; the circled bleached part is the result of chlorophyll present in the protein samples. Expected bands are of 11 kDa.
c. Odyssey®-based western blot analysis of the whole cell extracts of the CVN transformants acetone-treated to remove chlorophyll from the protein samples. Expected bands are of 11 kDa.

+ctrl – 10ng of purified cyanovirin-N operating as positive control, 10µg of cyanovirin-N, M – protein marker, negative controls: WT – wild type, TN72 – recipient cell line, TNEmp – TN72 transformed with pASapI vector; transformant lines: pASapI-oCVN, pASapI-sCVN, pSSapI-oCVN, pSSapI-sCVN. oCVN – original gene isolated from Nostoc sp. sCVN – synthetic codon-optimised version of cyanovirin-N gene.

Provisional results obtained in this report are very promising as it suggests a possibility to use algal transformation systems are not only potentially safer and more economic, but also better adapted particularly for the cyanovirin-N production compared to plant expression systems.

6.4. Conclusions and future work

Transformation of the BST-SAME recipient cell line resulted in the creation of L1 transformants, however neither of them expressed the L1 transgene and the *L1* transcripts were not present at a detectable level (Figure 6-4). However, transformation of another recipient cell line TN72 resulted in creation of seven transformed cell lines containing the *L1* transgene and the expression of L1 protein in transformed isolates was at the detectable concentration (Figure 6-8). There were seven different constructs used in this application therefore, in the next step it is necessary to establish which regulatory region was the most efficient for L1 synthesis in algal chloroplasts.

Addition of the CDS25 sequence to the expression vector upstream of the *L1* coding sequence appears to improve the L1 protein expression and this was true for both pASapI and pSSapI constructs yet, quantitative protein analysis needs to be performed. Additionally, efficient enrichment and purification methods for unmodified and modified L1 proteins have to be established, since the concentration of the transgenic L1 protein in tested samples of *C. reinhardtii* transformants is relatively low. Moreover, further investigations on modified L1 proteins need to be performed such as assessment of their capability to firstly form VLPs and secondly induce immune response in mammalian cells.

In the second part of this project, *C. reinhardtii* transformants synthesizing detectable levels of transgenic cyanovirin-N under two types regulatory regions: *atpA* and *psaA* were successfully created. Preliminary results indicate that the first promoter/5'UTR region regulates the CVN production more proficiently than the latter. Moreover, the synthetic version, codon adapted for *C. reinhardtii* chloroplast, particularly under *psaA* regulatory region, results in higher level of recombinant cyanovirin expression compared to the native version from *Nostoc sp*. These findings however, ought to be further evaluated and quantified. Furthermore, biological activity of transgenic cyanovirin needs to be evaluated.

CHAPTER 7. GENERAL DISCUSSION

7.1 Growth optimization of *Chlorella sorokiniana* for biofuel production

Increasing energy demands and disconcerting carbon emission level observed in recent years have resulted in increased research into renewable energy sources, particularly those derived from photosynthetic organisms. In contrast to other forms of renewable energy, such as solar, wind or tidal, liquid biofuels can be used in existing engines. Currently available biodiesel obtained from oil crops is of sufficient quality yet its production competes for arable land with food crops. Also, its energy and greenhouse gas emission balance is less than desirable (Hill et al., 2006). Hence, it is to no surprise that research into alternative sources of biofuel that can be obtained from microalgae has been increasing in recent years (Chen et al., 2012, Chisti, 2008, Fishman et al., 2010, Gong and Jiang, 2011, Gouveia and Oliveira, 2009, Graham et al., 2012, Hu et al., 2008b, Lei et al., 2012, Li et al., 2008, Scott et al., 2010, Varfolomeev and Wasserman, 2011, Williams, 2007).

Various advantages of microalgae over plants as a source of biodiesel have been established, since microalgal biomass yield is significantly higher than that of plants and is characterised by a higher lipid content and lower land requirements when compared to higher plants (Brennan and Owende, 2010). Yet, many challenges need to be tackled in order to exploit biofuel production from algae as a viable idea from an economic point of view. The most important issue is the identification of the best performing strains that are not only fast growing but also capable of lipid accumulation to a high level. Therefore, the initial aim of this project was focused on the characterisation of biomass productivity of an algal strain that has been reported to have a very high growth rate – *Chlorella sorokiniana* (Sorokin and Myers, 1953b, Sorokin and Krauss, 1958, Sorokin and Krauss, 1959, Sorokin, 1967). Chapter 3 presents the compiled data on the growth properties of this species. Initially, three strains of *C. sorokiniana* (UTEX1230, H-1983 and H-1986) were compared in terms of growth rate in various conditions, such as wide temperature range (Figure 3-5), carbon source (Figure 3-8), salinity tolerance (Figure 3-10) or growth at reduced oxygen concentration (Figure 3-11). As a result, the UTEX1230 strain was established to be the most robust strain of *C. sorokiniana* since it grows at the fastest rate in the temperature range of 16.5^0C-39^0C reaching a specific growth rate μ=0.33 in

mixotrophic conditions at 35^0C. Furthermore, this strain tolerates salinity up to 1% without a significant change in its growth rate. The ability to withstand the damaging influence of increased salt concentration in the growth medium is particularly useful in scaled-up projects sited in areas with limited fresh water supply. Increased salinity also functions as a stressful environmental condition capable of inducing lipid production in microalgae (Lei et al., 2012). The presence of NaCl exceeding 1% (w/v) in the growth medium somewhat inhibits growth of *C. sorokiniana* (Figure 3-10.b). It remains to be seen whether this influences the overall lipid productivity of the strain.

Although a specific growth rate for the UTEX1230 strain of 0.33 was established in laboratory conditions at 35^0C, this value should be confirmed using scaled up algal production in photobioreactors and in open pond systems. Outdoor conditions are characterised by relatively high temperature fluctuation particularly between day and night conditions and this difference depends on the type of climate in which the algal culture is maintained. It has been demonstrated that UTEX1230 has the ability to grow in a wide range of temperature conditions ranging between 16.5^0C and 42^0C, and 35^0C has been established as the optimum (Figure 3-5). For many microalgal strains the optimal temperature range is lower and most of the time it varies between 25^0C and 30^0C (Mehlitz, 2009). The ability to tolerate higher temperature conditions gives UTEX1230 an important advantage for outdoor cultivation, particularly in areas characterised by high irradiance and elevated temperature.

In order to further exploit the potential of UTEX1230 as an efficient factory for biofuel production, six mutants with diminished chlorophyll antenna size (TAM1-6) were generated (Figure 3-12). The most promising mutants TAM2 and TAM4 are characterised by an increased chlorophyll a/b ratio (Figure 3-13) and significantly lower chlorophyll content (Figure 3-14) in comparison to the wild type. The chlorophyll a/b ratio was approximately 50% higher for TAM2 and 66% for TAM4 (Figure 3-13). Both TAM2 and TAM4 also had a substantially decreased biochemical size of the PSII antenna (Figure 3-16) in addition to its decline in functional activity (Figure 3-17). Finally, TAM2 and TAM4 were examined in terms of growth rates in various light conditions (Figure 3-19). Growth rates of both strains remained at the same rate regardless of light intensity whereas for wild type the growth rates diminished with decreasing light (Figure 3-19.b).

All TAM isolates were examined in terms of the size of the chlorophyll antenna of PSII only. Further experiments are currently being completed to assess if the PSI antenna size is also altered in the TAM strains and if such modification influences algal productivity. It has been shown however, that reducing the PSII antenna complex is more critical than PSI for overall algal biomass production (Polle et al., 2002).

The TAM mutants were created with the intention of increasing productivity of *C. sorokiniana* by increasing the efficiency of solar energy to biomass conversion. In principle, the reduction of the antenna size and total chlorophyll content results in an increased efficiency of light penetration (Melis, 2009). The optimal size of the light harvesting complex is predominantly important in outdoor cultivation systems where acclimation to imbalanced light distribution within the culture has a significant impact on cell density and biomass productivity (Beckmann et al., 2009). The results presented in chapter 3 confirmed this concept, since the specific growth rates of TAM2 and TAM4 were substantially higher in respect to wild type, particularly in cultures under medium to low irradiation (Figure 3-19). Similarly, reducing chlorophyll antenna size resulted in substantial improvement of photosynthetic efficiency in *C. reinhardtii* where such modification was achieved via RNAi technology (Mussgnug et al., 2007) or genetic engineering (Polle et al., 2003). Truncated chlorophyll antenna mutants (*tla*) of *Chlamydomonas reinhardtii* and *Dunaliella salina* have found practical application in photoproduction of hydrogen (Kosourov et al., 2011, Polle et al., 2002). As a result, a *tla* strain displaying increased efficiency of hydrogen production was achieved in cultures under high light intensity over a long period of time in sulphur and phosphorus depletion (Kosourov et al., 2011).

Current work in our research group is focused on determination of lipid productivity of *C. sorokiniana* UTEX1230 wild type. Initial data suggest that under nitrogen-depleted conditions at 25^0C, storage lipid accumulation was seen to be induced and reached a maximum of 0.13 mg of derived fatty acid methyl esters (FAMEs or biodiesel) per mg of algal dry weight (Sofie Vonlanthen, personal communication). This rate is not optimal hence, Purton's group has attempted to boost lipid productivity of UTEX1230 strain by applying novel strategies such as discussed above by reduction of chlorophyll antenna size or increased temperature and salinity.

7.2 Development of transformation methods for *Chlorella sorokiniana*

In order to further utilise any microalgal strain for biofuel production, it is necessary to develop a genetic engineering technology that would allow strain improvement beyond that possible by classic mutagenesis. To date, various algal strains have been subjected to nuclear transformation yet, chloroplast transformation technology has been developed for only a few microalgal species (see Chapter 4 for details).

The development of transformation systems for algae commenced with the chloroplast transformation of *C. reinhardtii* (Boynton et al., 1988) – the most studied microalgal species in terms of fundamental mechanisms of biological processes. Transformation methods for this species are currently widely available (Purton, 2006) yet, its growth parameters such as growth rate and resistance to elevated temperature conditions are considerably lower than those of *C. sorokiniana* UTEX1230.

The major obstacle in the development of a chloroplast transformation method for *Chlorella sorokiniana* UTEX1230 is the absence of an available genomic database for this strain. Therefore, it was necessary to design a set of suitable degenerate primers that were created based on the likely homology in coding sequences between *C. sorokiniana* and *C. vulgaris*. As a result, a 3.1 kb fragment of chloroplast DNA was successfully amplified and sequenced (Figure 4-3). The amplified fragment was used to assemble two expression vectors containing the *aadA* as a selectable marker: pU.Cs3-*aadA* and pU.Cs7.3-psbX-*aadA*. The pU.Cs3-*aadA* expression vector (Figure 4-4) contains the *aadA* gene controlled by the promoter/5'UTR of the *atpA* gene of *C. reinhardtii*, whereas in the pU.Cs7.3-psbX-*aadA* expression vector the *aadA* gene was regulated by the promoter/5'UTR of the *psbA* gene amplified from *C. sorokiniana* H-1983 (Figure 4-6). Transformation with the pU.Cs3-*aadA* expression vector resulted in recovery of a number of putative transformants. PCR amplification indicated that four isolates contained the *aadA* gene in their genome, however, Southern blot analysis showed the *aadA* was not incorporated between *clpP* and *psbB* genes as it was anticipated (Figure 4-5). Similarly, transformants created with the pU.Cs7.3-psbX-*aadA* expression vector resulted in numerous putative transformants, yet none of them contained the *aadA* incorporated at the targeted place within the chloroplast genome (Figure 4-7). The newly established *psbA* promoter/5'UTR of *C. sorokiniana* proved to be functional in *E. coli* (Figure 4-9)

however, it was not operative in the *C. reinhardtii* chloroplast (Figure 4-10 and Figure 4-11).

Since the *C. sorokiniana psbA* promoter/5'UTR was shown to be non-functional in *C. reinhardtii*, it is possible this element is not functional in *C. sorokiniana*, which could explain the lack of transformants when using the pU.Cs7.3-psbX-*aadA* expression vector. Therefore, it is necessary to isolate an alternative homologous promoter/5'UTR for further tests. The use of homologous promoter is generally considered to be more efficient than a heterologous promoter since it has been reported that the expression of transgene fused to heterologous promoters was unsuccessful in diatoms (Apt et al., 1996) whereas in *C. reinhardtii* the transformation rate was low (Hall et al., 1993) or transgene expression was unstable (Tang et al., 1995)

A high background of false positive colonies following the transformation of *C. sorokiniana* is an additional obstacle associated with the presented work. Therefore, choosing a suitable selectable marker plays a critical role in the efficient selection for true transformant lines. Consequently, a new approach was chosen for creation of photosynthesis-deficient recipient strains for chloroplast transformation of *C. sorokiniana*. 5-fluorodeoxyuridine (FdUrd) – based mutagenesis was applied with metronidazole as a selective agent for photosynthetic mutants. This approach was feasible since *C. sorokiniana* similarly to *C. reinhardtii* is able to grow heterotrophically using acetate as a carbon source and is able to develop normal chloroplasts when cultured in the dark (Somerville, 1986). As a result, four putative photosynthetic mutants - PSM315, PSM941, PSM1345, and PSM1645 were isolated and characterised in terms of growth and fluorescence and it has been verified that all of them are PSII mutants (Figure 4-12 and Figure 4-13).

Such a random mutagenesis approach frequently leads to the formation of numerous mutant phenotypes. It has been demonstrated that FdUrd mutagenesis performed in *C. reinhardtii* resulted in a range of defects such as impairments in photosynthetic pigment accumulation, PSII activity, PSI activity, and photosynthetic carbon metabolism activity (Spreitzer and Mets, 1981). The four isolated mutant lines are currently being analysed for the sensitivity towards light. Additionally, several attempts were made to determine whether the specific mutations were within the plastid genome. Following the identification of a mutated plastid gene, a chloroplast transformation method based on the strategy of restoring photosynthesis could be established. On the other hand, any nuclear

mutations resulting in the impaired photosynthesis phenotype could be potentially rescued via chloroplast genetic manipulation, since it has been demonstrated that nuclear genes have been successfully expressed in the chloroplasts of the *C. reinhardtii* nuclear mutants (Surzycki et al., 2007).

To date, only one instance of a successful transformation of *Chlorella sorokiniana* has been reported (Dawson et al., 1997). In this report, nuclear transformation with the nitrate reductase (NR) gene isolated from *Chlorella vulgaris* was performed on a nitrate reductase-deficient mutant and this resulted in stable transformants capable of growth on nitrate as the sole nitrogen source.

The final part of chapter 3 presents the attempts to develop a nuclear transformation method for *C. sorokiniana* UTEX1230. This project involved selection for spontaneous dominant mutations within an endogenous nuclear gene *RPL41* resulting in resistance to an inhibitor of cytoplasmic translation – cycloheximide (CYH). It has been established that the level of resistance to CYH (Figure 4-14) was directly proportional to the type of a point mutation at position 56 of the RPL41 amino acid sequence (Figure 4-15). The P-L variant of the *RPL41* gene conferred the highest resistance to CYH and appears to be a good candidate as a selectable marker for nuclear transformation. Therefore, several methods aiming to amplify the full *RPL41* gene sequence were applied (Figure 4-15). However, the 5' end of this gene including its promoter/5'UTR has not been identified yet. Future work involves applying alterations in the amplification methods (see Chapter 4) in order to achieve a successful recovery of the 5' region of the *RPL41* gene. Subsequently, the amplified *RPL41* gene containing the P-L mutation could be employed as an efficient selectable marker for nuclear transformation of *C. sorokiniana*.

Additionally to the development of the suitable cloning and selectable marker strategy, an approach for the optimal DNA delivery method needs to be addressed for both chloroplast and nuclear transformation procedures. In the presented work, two DNA delivery methods were tested: a glass bead-mediated technique and the biolistic particle delivery method. To date, the most frequent methods used for transformation of various *Chlorella* strains involve electroporation (Chen et al., 2001, Wang et al., 2007), incubation of the protoplasts with polyethylene glycol and dimethyl sulfoxide (Hawkins and Nakamura, 1999) or incubation with cellulase prior to using a modified polyethylene glycol method (Liu et al., 2012). Therefore, optimisation of these methods for

transformation of *C. sorokiniana* could potentially result in increased efficiency of the foreign DNA delivery to the recipient strain.

7.3 Biodiesel production in algae

Selection for the best performing strain in terms of biomass productivity and lipid content presented in the first part of this thesis has been currently the most popular approach for strain choice in biodiesel production (Hu et al., 2008b, Rodolfi et al., 2009, Larkum et al., 2012). Although adaptive evolution is required for the species survival, living cells are not genetically programmed for optimised mass production (Larkum et al., 2012). Therefore, strain selection is often supported by high-throughput selection from algal collections followed by UV mutagenesis (Deng et al., 2011b). This type of genetic modification is not considered as a "GM" technique therefore, this approach escapes the regulatory concerns regarding using genetically modified algae in outdoor conditions (Larkum et al., 2012). UV mutagenesis results in the selection of a variety of mutants, and while it may lead to isolation of strains with higher than average lipid content, they frequently have additional undesirable genomic changes. Current selection methods for microalgal strains accumulating substantial amounts of lipids are not optimal. Therefore, it is essential to use a targeted genetic engineering approach to improve metabolism of the existing strains in order to achieve economic sustainability of biodiesel. In principle, by using novel biotechnology procedures, yields of biodiesel can be improved without an increase in the amount of energy required for its production (Lee et al., 2008), and for that reason various genetic engineering approaches performed on microalgal strains have recently been under rapid development (Yan and Liao, 2009, Rismani-Yazdi et al., 2011, Imamura et al., 2012).

The existing method of biofuel production from algal oil bodies involves an esterification step of the extracted triacylglycerides (TAGs) to form fatty acid methyl esters (biodiesel). Since alkanes are the major constituents of conventional diesel, the work presented in this thesis aimed to exclude the esterification process by direct engineering of alkane-producing *C. reinhardtii*. Although the biochemical pathway of alkane biosynthesis in eukaryotic organisms able to synthesise alkanes is still poorly understood, research on two prokaryotic species, *Vibrio furnissii* (Park, 2005) and *Synechococcus elongatus* (Schirmer et al., 2010), resulted in identification of two pathways for alkane synthesis

and the work presented in this thesis aimed to genetically engineer both proposed pathways.

The first approach involved an attempt to overexpress a thioesterase gene in order to increase a pool of free fatty acids – a direct substrate in the proposed alkane synthesis pathway for *Vibrio furnissii* (Park, 2005). Therefore, a suitable thioesterase from *Cuphea hookeriana* (Jones et al., 1995) optimised for expression in chloroplasts of *C. reinhardtii* was introduced into the plastid genome of a recipient cell line. Unfortunately, no transformed colonies were isolated and an alternative strategy to overexpress a thioesterase was discussed (see Chapter 5). A new tactic would involve alteration of the fatty acid chain length by overexpressing a thioesterase that results in synthesis of medium chain fatty acids (Dehesh et al., 1996b, Radakovits et al., 2011) such as a thioesterase from *U. californica* (BTE). BTE expressed in *E. coli* resulted in increased production of fatty acids, predominantly medium chain C_{12} and C_{14} (Voelker and Davies, 1994, Lu et al., 2008). Since medium chain alkanes are predominantly found in biodiesel, this approach would aim to generate high quality biodiesel.

The second approach involved amplification of genes coding for fatty acid reductase and aldehyde decarbonylase from *Synechocystis* PCC6803 (see Chapter 5). Both genes were cloned into four expression vectors controlled by four types of promoter/5'UTR region (Figure 5-5). Unfortunately, neither of the genes was successfully expressed in the chloroplasts of the transformed cell lines individually (Figure 5-14 and Figure 5-15) or when both genes were introduced in the form of operon controlled by the same promoter/5'UTR (Figure 5-8 and Figure 5-9). All four promoter/5'UTR proved to be functional in *E. coli* and drove expression of the reductase gene as judged by western blotting (Figure 5-7).

Among the four types of applied promoter/5'UTRs, the 5' regulatory region of *chlL* was designed to test a possibility to develop a system for biofuel production using an inducible promoter (Ninlayarn, 2012). An inducible system for chloroplast gene expression has been successfully employed in the biofuel production in *C. reinhardtii* (Surzycki et al., 2007). In this report, a cytochrome C_6 promoter fused to the nucleus-encoded Nac2 gene was introduced into the nac2-26 mutant strain deficient in Nac2. This system was evaluated for sustained cycling hydrogen production as inducible system in *Chlamydomonas*.

The possible reasons for the failure of this experiment were thoroughly discussed in Chapter 5. The major concern however, was directly related to the optimisation of the coding sequences for expression in *C. reinhardtii* chloroplasts since every protein-coding sequence naturally expressed in living organisms demonstrates a bias in synonymous codon usage that reflects the genome composition bias (Morton and So, 2000). Highly expressed genes tend to be composed of a particular selection of codons complementary to the most abundant tRNAs, and this appears to be an adaptation to increase translation efficiency of highly expressed genes (Ikemura, 1985). This idea suggested that codon bias improves transformation efficiency of the foreign genes (Mayfield and Kindle, 1990, Heitzer et al., 2007, Wu et al., 2011a). On the contrary, it has been reported that plastid genes in *Chlamydomonas* characterised by high overall codon adaptation index present distinctly lower codon adaptation at the 5' end, predominantly within the first 10-20 codons (Morton and So, 2000). It has been suggested in *E. coli* that increased level of rare codons at the 5' end of genes is associated with the optimal structure of the transcript and creation of ribosome binding sites (Eyre-Walker and Bulmer, 1993) and increased level of rare codons at the 5' end plays part in the mechanism of gene regulation (Chen and Inouye, 1994). Therefore, the future work on alkane biosynthesis in algal plastids should involve the optimisation and further adaptation of the coding and regulatory sequences of the chloroplast expression vectors.

7.4 Synthesis of high value product in green algae

Traditional industrial scale production of recombinant protein-based therapeutics involves mammalian, bacterial or yeast cell lines (Walsh, 2010). However, these systems have several disadvantages, for instance high maintenance cost, safety issues, and scalability are associated with those systems (Fischer et al., 2004). Therefore, interest in research towards an alternative expression system using algal cells as bioreactors have been increasing in the recent years (Rasala et al., 2010, Franklin and Mayfield, 2004, Mayfield and Franklin, 2005).

An established system for protein expression in *C. reinhardtii* plastids developed in Purton's lab enabled a technique for production of a HPV vaccine component – major capsid protein L1 of HPV16 and a potent anti-HIV protein – cyanovirin-N (CVN). Two versions of the synthetic codon-adapted HPV16 *L1* gene were created, the unmodified and truncated version where the first 10 codons were removed. Both versions of the

HPV16 *L1* gene were further customised by including 25 codons of the *atpA* gene (CDS25) upstream of the *L1* sequence. The latter modification was integrated in order to ensure the efficient RNA synthesis and/or for RNA stability and efficient translation of the transgene. Subsequently, both types of the HPV16 *L1* gene, with or without the CDS25 sequence (Figure 6-6), were introduced into the plastid genome of a cell wall-deficient, photosynthetically-impaired recipient cell line of *C. reinhardtii*. The expression of both versions of the HPV16 *L1* gene was achieved under three different promoter/5'UTRs – *atpA*, *chlL* and *psaA* and the foreign protein synthesis was tested by western blot analysis (Figure 6-8). Interestingly, the expression of the truncated version of HPV16 L1 in *E coli* was achieved only under the *atpA* promoter/5'UTR with included CDS25 sequence, whereas the L1 protein appears to be substantially degraded in the remaining cell lines (Figure 6-7).

Furthermore, attempts were made to express another recombinant therapeutic protein – CVN in plastids of *C. reinhardtii*. The coding sequence of CVN introduced to the chloroplast genome was either directly amplified from *Nostoc sp.* or codon adapted for *C. reinhardtii* chloroplast genome. Preliminary results suggest that *C. reinhardtii* transformants are capable of generating detectable levels of CVN under two regulatory elements: *atpA* and *psaA* (Figure 6-11) yet, the *atpA* promoter/5'UTR region appears to be more efficient than *psaA*, although further verification and quantification analysis is required to confirm these findings.

The oral application of HPV16 virus-like particles has been shown to be highly immunogenic in mice (Gerber et al., 2001) and this ability opens a possibility for oral vaccination against papillomaviruses. *Chlamydomonas sp.* and *Chlorella sp.* are GRAS organisms hence, there is a possibility of the cost reduction of HPV vaccine production, since the vaccine production could potentially be accomplished on site and the processing and purification procedures could be omitted. It has been revealed that genetically modified lactic acid bacteria producing recombinant CVN have been successfully incorporated in a food product and this oral administration resulted in significantly decreased viral activity (Li et al., 2011) what appears to be a promising strategy for proof-of-concept of mucosally administered vaccines and microbicides. Additionally, the concept of a direct oral immunisation with a vaccine-producing *C. reinhardtii* strains has recently been verified (Dreesen et al., 2010). In this report, lyophilized transgenic *C. reinhardtii* cell lines expressing fibronectin-binding domain D2 of *Staphylococcus aureus*

fused to the cholera toxin B subunit acting as mucosal adjuvant (Rappuoli et al., 1999) were applied in a 5-week oral immunisation program performed in mice. As a result, algae-based immune response was seen in the treated mice. Furthermore, the algae-based vaccine production was confirmed to be stable at room temperature for 1.5 years (Dreesen et al., 2010).

Low expression and difficulties in detection of both recombinant proteins in the transformed cell lines are generally associated with transcription, translation or protein stability. Therefore, it is essential to establish which step is critical in both projects in order to increase the production efficiency and exploit the developed algal expression system as an efficient platform for the recombinant protein production. It has been established that plastid gene expression is regulated primarily at the translation level (Eberhard et al., 2002) and the translation rate of foreign proteins tend to be substantially lower when compared to the expression level of endogenous genes (Coragliotti et al., 2011). As discussed earlier in this chapter, codon adaptation plays an important role in the efficient expression of foreign proteins. However, other factors have been associated with the expression of foreign proteins in algal cells. Numerous publications from recent years review various factors affecting the transgene expression and foreign protein stability (Mayfield et al., 1995, Leon-Banares et al., 2004, Barnes et al., 2005, Beligni et al., 2004, Manuell et al., 2004, Purton, 2006, Coragliotti et al., 2011, Rasala et al., 2011) and they were also discussed in Chapter 5 and 6. Therefore, it is essential to employ all discussed tactics to ensure the efficient accumulation of the L1 and CVN proteins.

7.5 Conclusions, challenges and perspectives

C. reinhardtii and *C. sorokiniana* as non-food GRAS organisms with several important attributes are promising hosts for biofuel and human therapeutic protein production. It has been demonstrated that both species are easy to scale up (Dreesen et al., 2010, Zheng et al., 2012) and the scaled up systems do not compete with agriculture. Additionally, an essential molecular biology toolbox successfully developed for *C. reinhardtii* (Boynton et al., 1988, Debuchy et al., 1989, Walker et al., 2005c, Purton, 2006) enables a straightforward and rapid construction of transformed lines expressing efficiently foreign genes of interest. The reports describing successful transformation of *C. sorokiniana* (Dawson et al., 1997) and results presented in Chapter 4 of this thesis look very promising. However, it is clear this system requires further optimisation. Moreover, the

biomass remaining after the recovery of valuable compounds can be further utilized and processed providing an additional source of revenue.

However, before this proposed technology can be fully exploited, certain obstacles have to be overcome. Firstly, increasing the expression level of foreign genes is a crucial issue. Several problems have been addressed in this thesis on how to improve the expression level of foreign proteins and all of them should be seriously considered in the future attempts to improve algal molecular techniques. Secondly, the optimisation of the scaled-up productions at industrial level and the operational parameters such as optimal size of the photobioreactor or open ponds, mixing rates, light intensity and distribution, initial density, temperature, nutrient quantity need to be monitored and controlled, in order to achieve maximal algal productivity (Rosales-Mendoza et al., 2012). Furthermore, the optimisation of downstream processing parameters specific for each application need to be considered, since this step also affects the yield and cost of the recombinant compounds. Therefore, the development of economically viable transgenic microalgae-based bioreactors fundamentally depends on efficient cooperation between researchers dealing with upstream side of process engineering with that of the downstream processing.

REFERENCES

AARTS, M. G., KEIJZER, C. J., STIEKEMA, W. J. & PEREIRA, A. 1995. Molecular characterization of the CER1 gene of arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. *The Plant Cell,* 7**,** 2115-27.

ABE, J., HORI, S., TSUCHIKANE, Y., KITAO, N., KATO, M. & SEKIMOTO, H. 2011. Stable Nuclear Transformation of the Closterium peracerosum-strigosum-littorale Complex. *Plant and Cell Physiology,* 52**,** 1676-1685.

ABRAMSON, A. L., STEINBERG, B. M. & WINKLER, B. 1987. Laryngeal papillomatosis: clinical, histopathologic and molecular studies. *The Laryngoscope,* 97**,** 678-85.

ADAMS III, W. W., DEMMIG-ADAMS, B. & LANGE, O. L. 1993. Carotenoid Composition and Metabolism in Green and Blue-Green Algal Lichens in the Field. *Oecologia,* 94**,** 576-584.

AGUSTÍ, S. 1991. Light environment within dense algal populations: cell size influences on self-shading. *Journal of Plankton Research,* 13**,** 863-871.

AMOTZ, A. B. 2004. Industrial production of microalgal cell mass and secondary products - major industrial species: Dunaliella. *Handbook of Microalgal Culture,* 57-82.

ANDERSEN, R. A. 2004. Biology and systematics of heterokont and haptophyte algae. *American Journal of Botany,* 91**,** 1508-1522.

ANDERSSON, J., WALTERS, R. G., HORTON, P. & JANSSON, S. 2001. Antisense inhibition of the photosynthetic antenna proteins CP29 and CP26: implications for the mechanism of protective energy dissipation. *The Plant Cell,* 13**,** 1193-204.

ANEMAET, I. G., BEKKER, M. & HELLINGWERF, K. J. 2010. Algal photosynthesis as the primary driver for a sustainable development in energy, feed, and food production. *Marine Biotechnology,* 12**,** 619-29.

APEL, K. & KLOPPSTECH, K. 1978. The plastid membranes of barley (*Hordeum vulgare*). Light-induced appearance of mRNA coding for the apoprotein of the light-harvesting chlorophyll a/b protein. *European Journal of Biochemistry / FEBS,* 85**,** 581-8.

APT, K. E., KROTH-PANCIC, P. G. & GROSSMAN, A. R. 1996. Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Molecular & General Genetics : MGG,* 252**,** 572-9.

ARNOLD, K. E. & MURRAY, S. N. 1980. Relationships between irradiance and photosynthesis for marine benthic green algae (Chlorophyta) of differing morphologies. *Journal of Experimental Marine Biology and Ecology,* 43**,** 183-192.

ASLESON, C. M. & LEFEBVRE, P. A. 1998. Genetic analysis of flagellar length control in *Chlamydomonas reinhardtii*: a new long-flagella locus and extragenic suppressor mutations. *Genetics,* 148**,** 693-702.

BAILEY, S., HORTON, P. & WALTERS, R. G. 2004. Acclimation of *Arabidopsis thaliana* to the light environment: the relationship between photosynthetic function and chloroplast composition. *Planta,* 218**,** 793-802.

BALFOORT, H. W., SNOEK, J., SMITHS, J. R. M., BREEDVELD, L. W., HOFSTRAAT, J. W. & RINGELBERG, J. 1992. Automatic identification of algae: neural network analysis of flow cytometric data. *Journal of Plankton Research,* 14**,** 575-589.

BALZARINI, J. & VAN DAMME, L. 2007. Microbicide drug candidates to prevent HIV infection. *Lancet,* 369**,** 787-97.

BALZARINI, J., VAN LAETHEM, K., PEUMANS, W. J., VAN DAMME, E. J., BOLMSTEDT, A., GAGO, F. & SCHOLS, D. 2006. Mutational pathways, resistance profile, and side effects of cyanovirin relative to human immunodeficiency virus type 1 strains with N-glycan deletions in their gp120 envelopes. *Journal of Virology,* 80**,** 8411-21.

BARKAN, A. & GOLDSCHMIDT-CLERMONT, M. 2000. Participation of nuclear genes in chloroplast gene expression. *Biochimie,* 82**,** 559-572.

BARNES, D., FRANKLIN, S., SCHULTZ, J., HENRY, R., BROWN, E., CORAGLIOTTI, A. & MAYFIELD, S. P. 2005. Contribution of 5'- and 3'-untranslated regions of plastid mRNAs to the expression of *Chlamydomonas reinhardtii* chloroplast genes. *Molecular Genetics and Genomics : MGG,* 274**,** 625-36.

BARONE, P., ZHANG, X. H. & WIDHOLM, J. M. 2009. Tobacco plastid transformation using the feedback-insensitive anthranilate synthase [alpha]-subunit of tobacco (*ASA2*) as a new selectable marker. *Journal of Experimental Botany,* 60**,** 3195-202.

BARRIENTOS, L. G., O'KEEFE, B. R., BRAY, M., SANCHEZ, A., GRONENBORN, A. M. & BOYD, M. R. 2003. Cyanovirin-N binds to the viral surface glycoprotein, GP1,2 and inhibits infectivity of Ebola virus. *Antiviral Research,* 58**,** 47-56.

BASOVA, M. M. 2005. Fatty acid composition of lipids in microalgae. *International Journal on Algae,* 7**,** 33-57.

BATEMAN, J. M. & PURTON, S. 2000. Tools for chloroplast transformation in *Chlamydomonas*: expression vectors and a new dominant selectable marker. *Molecular Genetics and Genomics,* 263**,** 404-10.

BAZAN, S. B., DE ALENCAR MUNIZ CHAVES, A., AIRES, K. A., CIANCIARULLO, A. M., GARCEA, R. L. & HO, P. L. 2009. Expression and characterization of HPV-16 L1 capsid protein in Pichia pastoris. *Archives of Virology,* 154**,** 1609-17.

BECKER, W. 2007. Microalgae in Human and Animal Nutrition. *Handbook of Microalgal Culture.* Blackwell Publishing Ltd.

BECKMANN, J., LEHR, F., FINAZZI, G., HANKAMER, B., POSTEN, C., WOBBE, L. & KRUSE, O. 2009. Improvement of light to biomass conversion by de-regulation of light-harvesting protein translation in Chlamydomonas reinhardtii. *Journal of Biotechnology,* 142**,** 70-7.

BEER, L. L., BOYD, E. S., PETERS, J. W. & POSEWITZ, M. C. 2009. Engineering algae for biohydrogen and biofuel production. *Current Opinion in Biotechnology,* 20**,** 264-271.

BELIGNI, M. V., YAMAGUCHI, K. & MAYFIELD, S. P. 2004. The translational apparatus of *Chlamydomonas reinhardtii* chloroplast. *Photosynthesis Research,* 82**,** 315-25.

BELKOURA, M., BENIDER, A. & DAUTA, A. 1997. Effects of temperature, light intensity and growth phase on the biochemical composition of Chlorella sorokiniana Shihira & Krauss. *Annales De Limnologie-International Journal of Limnology,* 33**,** 3-11.

BELKOURA, M., BENIDER, A., EL ANTARI, A. & DAUTA, A. 2000. Effect of environmental conditions on the fatty acid composition of the green alga Chlorella sorokiniana Shihira et Krauss. *Archiv fuer Hydrobiologie Supplement***,** 93-101.

BENSON, A. A. & CALVIN, M. 1950. Carbon dioxide fixation by green plants *Annual Review of Plant Physiology and Plant Molecular Biology,* 1**,** 25-42.

BETTERLE, N., BALLOTTARI, M., HIENERWADEL, R., DALL'OSTO, L. & BASSI, R. 2010. Dynamics of zeaxanthin binding to the photosystem II monomeric antenna protein Lhcb6 (CP24) and modulation of its photoprotection properties. *Archives of Biochemistry and Biophysics,* 504**,** 67-77.

BIEMELT, S., SONNEWALD, U., GALMBACHER, P., WILLMITZER, L. & MULLER, M. 2003. Production of human papillomavirus type 16 virus-like particles in transgenic plants. *Journal of Virology,* 77**,** 9211-20.

BIRCH-MACHIN, I., NEWELL, C. A., HIBBERD, J. M. & GRAY, J. C. 2004. Accumulation of rotavirus VP6 protein in chloroplasts of transplastomic tobacco is limited by protein stability. *Plant Biotechnology Journal,* 2**,** 261-70.

BOLMSTEDT, A. J., O'KEEFE, B. R., SHENOY, S. R., MCMAHON, J. B. & BOYD, M. R. 2001. Cyanovirin-N defines a new class of antiviral agent targeting N-linked, high-mannose glycans in an oligosaccharide-specific manner. *Molecular Pharmacology,* 59**,** 949-54.

BONAVENTURA, C. & MYERS, J. 1969. Fluorescence and oxygen evolution from *Chlorella pyrenoidosa*. *Biochimica Et Biophysica Acta,* 189**,** 366-83.

BONENTE, G., HOWES, B. D., CAFFARRI, S., SMULEVICH, G. & BASSI, R. 2008. Interactions between the photosystem II subunit PsbS and xanthophylls studied in vivo and in vitro. *The Journal of Biological Chemistry,* 283**,** 8434-45.

BONENTE, G., PIPPA, S., CASTELLANO, S., BASSI, R. & BALLOTTARI, M. 2012. Acclimation of *Chlamydomonas reinhardtii* to Different Growth Irradiances. *The Journal of Biological Chemistry,* 287**,** 5833-5847.

BORDEL, S., GUIEYSSE, B. & MUNOZ, R. 2009. Mechanistic Model for the Reclamation of Industrial Wastewaters Using Algal-Bacterial Photobioreactors. *Environmental Science & Technology,* 43**,** 3200-3207.

BOSCH, F. X., LORINCZ, A., MUNOZ, N., MEIJER, C. J. & SHAH, K. V. 2002. The causal relation between human papillomavirus and cervical cancer. *Journal of Clinical Pathology,* 55**,** 244-65.

BOSTOCK, J., MCANDREW, B., RICHARDS, R., JAUNCEY, K., TELFER, T., LORENZEN, K., LITTLE, D., ROSS, L., HANDISYDE, N., GATWARD, I. & CORNER, R. 2010. Aquaculture: global status and trends. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences,* 365**,** 2897-912.

BOUDREAU, E. & TURMEL, M. 1995. Gene rearrangements in *Chlamydomonas* chloroplast DNAs are accounted for by inversions and by the expansion/contraction of the inverted repeat. *Plant Molecular Biology,* 27**,** 351-364.

BOYD, M. R., GUSTAFSON, K. R., MCMAHON, J. B., SHOEMAKER, R. H., O'KEEFE, B. R., MORI, T., GULAKOWSKI, R. J., WU, L., RIVERA, M. I., LAURENCOT, C. M., CURRENS, M. J., CARDELLINA, J. H., 2ND, BUCKHEIT, R. W., JR., NARA, P. L., PANNELL, L. K., SOWDER, R. C., 2ND & HENDERSON, L. E. 1997. Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrobial Agents and Chemotherapy,* 41**,** 1521-30.

BOYNTON, J. E., GILLHAM, N. W., HARRIS, E. H., HOSLER, J. P., JOHNSON, A. M., JONES, A. R., RANDOLPH-ANDERSON, B. L., ROBERTSON, D., KLEIN, T. M., SHARK, K. B. & ET, A. 1988. Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science,* 240**,** 1534-1538.

BRAHMA, S. K., HARGRAVES, P. E., HOWARD, W. F. & NELSON, W. H. 1983. A Resonance Raman Method for the Rapid Detection and Identification of Algae in Water. *Applied Spectroscopy,* 37**,** 55-58.

BRENNAN, L. & OWENDE, P. 2010. Biofuels from microalgae, a review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews,* 14**,** 557-577.

BROWN, A. C., KNIGHTS, B. A. & CONWAY, E. 1969. Hydrocarbon content and its relationship to physiological state in the green alga *Botryococcus braunii*. *Phytochemistry,* 8**,** 543-547.

BUCHHEIM, M. A. & CHAPMAN, R. L. 1991. Phylogeny of the colonial green flagellates - a study of 18S and 26S ribosomal-RNA sequence data. *Biosystems,* 25**,** 85-100.

BUFFA, V., STIEH, D., MAMHOOD, N., HU, Q. X., FLETCHER, P. & SHATTOCK, R. J. 2009. Cyanovirin-N potently inhibits human immunodeficiency virus type 1 infection in cellular and cervical explant models. *Journal of General Virology,* 90**,** 234-243.

BULEON, A., GALLANT, D. J., BOUCHET, B., MOUILLE, G., D'HULST, C., KOSSMANN, J. & BALL, S. 1997. Starches from A to C (Chlamydomonas reinhardtii as a Model Microbial System to Investigate the Biosynthesis of the Plant Amylopectin Crystal). *Plant Physiology,* 115**,** 949-957.

BURD, E. M. 2003. Human papillomavirus and cervical cancer. *Clinical Microbiology Reviews,* 16**,** 1-17.

BURJA, A. M., TAMAGNINI, P., BUSTARD, M. T. & WRIGHT, P. C. 2001. Identification of the green alga, *Chlorella vulgaris* (SDC1) using cyanobacteria derived 16S rDNA primers: targeting the chloroplast. *FEMS Microbiology Letters,* 202**,** 195-203.

BURTON, S. G. 2001. Development of bioreactors for application of biocatalysts in biotransformations and bioremediation. *Pure and Applied Chemistry,* 73**,** 77-83.

CAHOON, A. B. & TIMKO, M. P. 2000. Yellow-in-the-dark mutants of *Chlamydomonas* lack the CHLL subunit of light-independent protochlorophyllide reductase. *The Plant Cell,* 12**,** 559-68.

CALLEJA-MACIAS, I. E., KALANTARI, M., ALLAN, B., WILLIAMSON, A. L., CHUNG, L. P., COLLINS, R. J., ZUNA, R. E., DUNN, S. T., ORTIZ-LOPEZ, R., BARRERA-SALDANA, H. A., CUBIE, H. A., CUSCHIERI, K., VILLA, L. L. & BERNARD, H. U. 2005. Papillomavirus subtypes are natural and old taxa: phylogeny of human papillomavirus types 44 and 55 and 68a and -b. *Journal of virology,* 79**,** 6565-9.

CARAMUJO, M.-J., BOSCHKER, H. T. S. & ADMIRAAL, W. 2008. Fatty acid profiles of algae mark the development and composition of harpacticoid copepods. *Freshwater Biology,* 53**,** 77-90.

CARNICAS, E., JIMENEZ, C. & NIELL, F. X. 1999. Effects of changes of irradiance on the pigment composition of *Gracilaria tenuistipitata* var. liui Zhang et Xia. *Journal of Photochemistry and Photobiology B: Biology,* 50**,** 149-158.

CARRER, H., HOCKENBERRY, T. N., SVAB, Z. & MALIGA, P. 1993. Kanamycin resistance as a selectable marker for plastid transformation in tobacco. *Molecular & General Genetics,* 241**,** 49-56.

CERUTTI, H., JOHNSON, A. M., GILLHAM, N. W. & BOYNTON, J. E. 1997. Epigenetic silencing of a foreign gene in nuclear transformants of *Chlamydomonas*. *The Plant Cell,* 9**,** 925-45.

CHADER, S., HACENE, H. & AGATHOS, S. N. Study of hydrogen production by three strains of Chlorella isolated from the soil in the Algerian Sahara. 2nd International Workshop on Hydrogen Energy from Renewable Origin, Oct 27-29 2007 Ghardaia, ALGERIA. Pergamon-Elsevier Science Ltd, 4941-4946.

CHEESBROUGH, T. M. & KOLATTUKUDY, P. E. 1984. Alkane biosynthesis by decarbonylation of aldehydes catalyzed by a particulate preparation from *Pisum sativum*. *Proceedings of the National Academy of Sciences of the United States of America,* 81**,** 6613-7.

CHEESBROUGH, T. M. & KOLATTUKUDY, P. E. 1988. Microsomal preparation from an animal tissue catalyzes release of carbon monoxide from a fatty aldehyde to generate an alkane. *The Journal of Biological Chemistry,* 263**,** 2738-43.

CHEN, G. T. & INOUYE, M. 1994. Role of the AGA/AGG codons, the rarest codons in global gene expression in *Escherichia coli*. *Genes & Development,* 8**,** 2641-52.

CHEN, L., LIU, T., ZHANG, W., CHEN, X. & WANG, J. 2012. Biodiesel production from algae oil high in free fatty acids by two-step catalytic conversion. *Bioresource Technology,* 111**,** 208-14.

CHEN, Y., WANG, Y. & SUN, Y. 2001. Highly efficient expression of rabbit neutrophil peptide-1 gene in *Chlorella ellipsoidea* cells. *Current Genetics,* 39**,** 365–70.

CHENG, Q., DAY, A., DOWSON-DAY, M., SHEN, G. F. & DIXON, R. 2005. The *Klebsiella pneumoniae* nitrogenase Fe protein gene (nifH) functionally substitutes for the chlL gene in *Chlamydomonas reinhardtii*. *Biochemical and Biophysical Research Communications,* 329**,** 966-75.

CHISTI, Y. 2007. Biodiesel from microalgae. *Biotechnology Advances,* 25**,** 294-306.

CHISTI, Y. 2008. Biodiesel from microalgae beats bioethanol. *Trends in Biotechnology,* 26**,** 126-31.

CHOQUET, Y., RAHIRE, M., GIRARDBASCOU, J., ERICKSON, J. & ROCHAIX, J. D. 1992. A Chloroplast Gene Is Required for the Light-Independent Accumulation of Chlorophyll in Chlamydomonas-Reinhardtii. *Embo Journal,* 11**,** 1697-1704.

CHOQUET, Y. & WOLLMAN, F.-A. 2002. Translational regulations as specific traits of chloroplast gene expression. *FEBS Letters,* 529**,** 39-42.

CHRISTENSON, L. & SIMS, R. 2011. Production and harvesting of microalgae for wastewater treatment, biofuels, and bioproducts. *Biotechnology Advances,* 29**,** 686-702.

COHEN, J. 2009. Beyond Thailand: Making Sense of a Qualified AIDS Vaccine "Success". *Science,* 326**,** 652-653.

COLL, J. M. 2006. Methodologies for transferring DNA into eukaryotic microalgae. *Spanish Journal of Agricultural Research,* 4**,** 316-330.

COLLET, P., HELIAS, A., LARDON, L., RAS, M., GOY, R. A. & STEYER, J. P. 2011. Life-cycle assessment of microalgae culture coupled to biogas production. *Bioresource Technology,* 102**,** 207-14.

COLLIER, B., OBERG, D., ZHAO, X. & SCHWARTZ, S. 2002. Specific inactivation of inhibitory sequences in the 5' end of the Human Papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells. *Journal of Virology,* 76**,** 2739-52.

CORAGLIOTTI, A. T., BELIGNI, M. V., FRANKLIN, S. E. & MAYFIELD, S. P. 2011. Molecular factors affecting the accumulation of recombinant proteins in the *Chlamydomonas reinhardtii* chloroplast. *Molecular Biotechnology,* 48**,** 60-75.

CORDERO, B. F., COUSO, I., LEON, R., RODRIGUEZ, H. & VARGAS, M. A. 2011. Enhancement of carotenoids biosynthesis in *Chlamydomonas reinhardtii* by nuclear transformation using a phytoene synthase gene isolated from *Chlorella zofingiensis*. *Applied Microbiology and Biotechnology,* 91**,** 341-351.

CRAMER, M. & MYERS, J. 1948. NITRATE REDUCTION AND ASSIMILATION IN CHLORELLA. *Journal of General Physiology,* 32**,** 93-102.

CUARESMA, M., JANSSEN, M., VILCHEZ, C. & WIJFFELS, R. H. 2009. Productivity of *Chlorella sorokiniana* in a short light-path (SLP) panel photobioreactor under high irradiance. *Biotechnology and Bioengineering,* 104**,** 352-9.

CUI, Y., JIANG, P., WANG, J., LI, F., CHEN, Y., ZHENG, G. & QIN, S. 2012. Genetic transformation of Platymonas (Tetraselmis) subcordiformis (Prasinophyceae, Chlorophyta) using particle bombardment and glass-bead agitation. *Chinese Journal of Oceanology and Limnology,* 30**,** 471-475.

CULP, T. D. & CHRISTENSEN, N. D. 2004. Kinetics of in vitro adsorption and entry of papillomavirus virions. *Virology,* 319**,** 152-61.

CUMMINS, J. 2004. *Pharm Crops for Vaccines and Therapeutic Antibodies* [Online]. Available: http://www.biosafety-info.net/article.php?aid=160.

DALE, M. P. & CAUSTON, D. R. 1992. Use of the Chlorophyll a/b Ratio as a Bioassay for the Light Environment of a Plant. *Functional Ecology,* 6**,** 190-196.

DALL'OSTO, L., CAZZANIGA, S., NORTH, H., MARION-POLL, A. & BASSI, R. 2007. The Arabidopsis aba4-1 mutant reveals a specific function for neoxanthin in protection against photooxidative stress. *The Plant Cell,* 19**,** 1048-64.

DANIELL, H., MUTHUKUMAR, B. & LEE, S. B. 2001. Marker free transgenic plants: engineering the chloroplast genome without the use of antibiotic selection. *Current Genetics,* 39**,** 109-16.

DAVIES, D. & PLASKITT, A. 1971. Genetic and structural analyses of cell-wall formation in *Chlamydomonas reinhardtii*. *Genetics Research (Cambridge),* 33-43.

DAVISON, P. A., HUNTER, C. N. & HORTON, P. 2002. Overexpression of [beta]-carotene hydroxylase enhances stress tolerance in Arabidopsis. *Nature,* 418**,** 203-206.

DAWSON, H. N., BURLINGAME, R. & CANNONS, A. C. 1997. Stable Transformation of *Chlorella*: Rescue of Nitrate Reductase-Deficient Mutants with the Nitrate Reductase Gene. *Current Microbiology,* 35**,** 356-62.

DAY, A. & GOLDSCHMIDT-CLERMONT, M. 2011. The chloroplast transformation toolbox: selectable markers and marker removal. *Plant Biotechnology Journal,* 9**,** 540-53.

DAY, P. M., LOWY, D. R. & SCHILLER, J. T. 2003. Papillomaviruses infect cells via a clathrin-dependent pathway. *Virology,* 307**,** 1-11.

DE BIANCHI, S., BETTERLE, N., KOURIL, R., CAZZANIGA, S., BOEKEMA, E., BASSI, R. & DALL'OSTO, L. 2011. Arabidopsis mutants deleted in the light-harvesting protein Lhcb4 have a disrupted photosystem II macrostructure and are defective in photoprotection. *The Plant Cell,* 23**,** 2659-79.

DE-BASHAN, L. E., TREJO, A., HUSS, V. A., HERNANDEZ, J. P. & BASHAN, Y. 2008. *Chlorella sorokiniana* UTEX2805, a heat and intense, sunlight-tolerant microalga with potential for removing ammonium from wastewater. *Bioresource Technology,* 99**,** 4980-9.

DEBUCHY, R., PURTON, S. & ROCHAIX, J. D. 1989. The argininosuccinate lyase gene of *Chlamydomonas reinhardtii:* an important tool for nuclear transformation and for correlating the genetic and molecular maps of the ARG7 locus. *The EMBO Journal,* 8**,** 2803-9.

DEHESH, K., EDWARDS, P., FILLATTI, J., SLABAUGH, M. & BYRNE, J. 1998. KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea sp*. is a medium chain specific condensing enzyme. *The Plant Journal,* 15**,** 383-390.

DEHESH, K., EDWARDS, P., HAYES, T., CRANMER, A. M. & FILLATTI, J. 1996a. Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. *Plant Physiology,* 110**,** 203-210.

DEHESH, K., JONES, A., KNUTZON, D. S. & VOELKER, T. A. 1996b. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. *The Plant Journal,* 9**,** 167-172.

DEHESH, K., TAI, H., EDWARDS, P., BYRNE, J. & JAWORSKI, J. G. 2001. Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis. *Plant Physiology,* 125**,** 1103-14.

DELL, G., WILKINSON, K. W., TRANTER, R., PARISH, J., LEO BRADY, R. & GASTON, K. 2003. Comparison of the Structure and DNA-binding Properties of the E2 Proteins from an Oncogenic and a Non-oncogenic Human Papillomavirus. *Journal of Molecular Biology,* 334**,** 979-991.

DELWICHE, C. F. 1999. Tracing the Thread of Plastid Diversity through the Tapestry of Life. *The American Naturalist,* 154**,** S164-S177.

DEMMIG-ADAMS, B. & ADAMS III, W. W. 1992. Photoprotection and other responses of plants to high light stress. . *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.,* 43**,** 599-626.

DEMMIG-ADAMS, B. & ADAMS III, W. W. 2002. Antioxidants in photosynthesis and human nutrition. *Science,* 298**,** 2149-53.

DEMMIG-ADAMS, B., GILMORE, A. M. & ADAMS, W. W., 3RD 1996. Carotenoids 3: in vivo function of carotenoids in higher plants. *FASEB Journal* 10**,** 403-12.

DEMMIG-ADAMS, B., WINTER, K., KRUGER, A. & CZYGAN, F.-C. 1989. Zeaxanthin and the Induction and Relaxation Kinetics of the Dissipation of Excess Excitation Energy in Leaves in 2% O2, 0% CO2. *Plant Physiology,* 90**,** 887-893.

DENG, X., FEI, X. & LI, Y. 2011a. The effects of nutritional restriction on neutral lipid accumulation in *Chlamydomonas* and *Chlorella*. *African Journal of Microbiology Research,* 5**,** 260-270.

DENG, X., LI, Y. & FEI, X. 2011b. Effects of selective medium on lipid accumulation of chlorellas and screening of high lipid mutants through ultraviolet mutagenesis. *African Journal of Agricultural Research,* 6**,** 3768-3774.

DENNIS, M. & KOLATTUKUDY, P. E. 1992. A cobalt-porphyrin enzyme converts a fatty aldehyde to a hydrocarbon and CO. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 5306-10.

DENNIS, M. W. & KOLATTUKUDY, P. E. 1991. Alkane biosynthesis by decarbonylation of aldehyde catalyzed by a microsomal preparation from *Botryococcus braunii*. *Archives of Biochemistry and Biophysics,* 287**,** 268-75.

DENT, R. M., HAGLUND, C. M., CHIN, B. L., KOBAYASHI, M. C. & NIYOGI, K. K. 2005. Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. *Plant Physiology,* 137**,** 545-56.

DERELLE, E., FERRAZ, C., ROMBAUTS, S., ROUZÉ, P., WORDEN, A. Z., ROBBENS, S., PARTENSKY, F., DEGROEVE, S., ECHEYNIÉ, S., COOKE, R., SAEYS, Y., WUYTS, J., JABBARI, K., BOWLER, C., PANAUD, O., PIÉGU, B., BALL, S. G., RAL, J. P., BOUGET, F. Y., PIGANEAU, G., DE BAETS, B., PICARD, A., DELSENY, M., DEMAILLE, J., VAN DE PEER, Y. & MOREAU, H. 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 11647-11652.

DISMUKES, G. C., CARRIERI, D., BENNETTE, N., ANANYEV, G. M. & POSEWITZ, M. C. 2008. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology,* 19**,** 235-40.

DOEBBE, A., RUPPRECHT, J., BECKMANN, J., MUSSGNUG, J. H., HALLMANN, A., HANKAMER, B. & KRUSE, O. 2007. Functional integration of the HUP1 hexose symporter gene into the genome of C. reinhardtii: Impacts on biological H2 production. *Journal of Biotechnology,* 131**,** 27-33.

DOETSCH, N. A., FAVREAU, M. R., KUSCUOGLU, N., THOMPSON, M. D. & HALLICK, R. B. 2001. Chloroplast transformation in *Euglena gracilis*: splicing of a group III twintron transcribed from a transgenic psbK operon. *Current Genetics,* 39**,** 49-60.

DOORBAR, J. 2006. Molecular biology of human papillomavirus infection and cervical cancer. *Clinical Science,* 110**,** 525-41.

DORR, R. & HUSS, V. A. R. 1990. Characterization of nuclear-DNA in 12 species of Chlorella (Chlorococcales, Chlorophyta) by DNA reassociation. *Biosystems,* 24**,** 145-155.

DRECHSEL, O. & BOCK, R. 2011. Selection of Shine-Dalgarno sequences in plastids. *Nucleic Acids Research,* 39**,** 1427-1438.

DREESEN, I. A., CHARPIN-EL HAMRI, G. & FUSSENEGGER, M. 2010. Heat-stable oral alga-based vaccine protects mice from *Staphylococcus aureus* infection. *Journal of Biotechnology,* 145**,** 273-80.

DUFOURMANTEL, N., DUBALD, M., MATRINGE, M., CANARD, H., GARCON, F., JOB, C., KAY, E., WISNIEWSKI, J. P., FERULLO, J. M., PELISSIER, B., SAILLAND, A. & TISSOT, G. 2007. Generation and characterization of soybean and marker-free tobacco plastid transformants over-expressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide tolerance. *Plant Biotechnology Journal,* 5**,** 118-33.

DUNAHAY, T. G. 1993. Transformation of *Chlamydomonas reinhardtii* with silicon carbide whiskers. *Biotechniques,* 15**,** 452-5, 457-8, 460.

DUNAHAY, T. G., JARVIS, E. E. & ROESSLER, P. G. 1995. Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. *Journal of Phycology,* 31**,** 1004-1012.

EBENEZER, V., MEDLIN, L. K. & KI, J. S. 2012. Molecular detection, quantification, and diversity evaluation of microalgae. *Marine Biotechnology,* 14**,** 129-42.

EBERHARD, S., DRAPIER, D. & WOLLMAN, F. A. 2002. Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. *The Plant Journal,* 31**,** 149-60.

EICHLER-STAHLBERG, A., WEISHEIT, W., RUECKER, O. & HEITZER, M. 2009. Strategies to facilitate transgene expression in Chlamydomonas reinhardtii. *Planta,* 229**,** 873-83.

ELGHABI, Z., KARCHER, D., ZHOU, F., RUF, S. & BOCK, R. 2011. Optimization of the expression of the HIV fusion inhibitor cyanovirin-N from the tobacco plastid genome. *Plant Biotechnology Journal,* 9**,** 599-608.

ELLIS, B. E. 1977. Degradation of phenolic compounds by fresh-water algae. *Plant Science Letters,* 8**,** 213-216.

ENY, D. M. 1950. Respiration studies on Chlorella growth experiments with acid intermediates *Plant Physiology,* 25**,** 478-495.

ERICKSON, J. M., RAHIRE, M., BENNOUN, P., DELEPELAIRE, P., DINER, B. & ROCHAIX, J. D. 1984. Herbicide resistance in *Chlamydomonas reinhardtii* results from a mutation in the chloroplast gene for the 32-kilodalton protein of photosystem II. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* 81**,** 3617-3621.

ETTL, H. 1976. [About the progress of chloroplast division in *Chlamydomonas* (author's transl)]. *Protoplasma,* 88**,** 75-84.

EYRE-WALKER, A. & BULMER, M. 1993. Reduced synonymous substitution rate at the start of enterobacterial genes. *Nucleic Acids Research,* 21**,** 4599-603.

EYSTER, C. 1978. Nutrient concentration requirements for Chlorella sorokiniana. *Ohio Journal of Science,* 78**,** 79-81.

FALKOWSKI, P. 1994. The role of phytoplankton photosynthesis in global biogeochemical cycles. *Photosynthesis Research,* 39**,** 235-258.

FARGO, D. C., ZHANG, M., GILLHAM, N. W. & BOYNTON, J. E. 1998. Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. *Molecular & General Genetics,* 257**,** 271-282.

FELDMAN, D. E. & FRYDMAN, J. 2000. Protein folding in vivo: the importance of molecular chaperones. *Current Opinion in Structural Biology,* 10**,** 26-33.

FENG, C. & JOHNS, M. R. 1991. EFFECT OF C/N RATIO AND AERATION ON THE FATTY-ACID COMPOSITION OF HETEROTROPHIC CHLORELLA-SOROKINIANA. *Journal of Applied Phycology,* 3**,** 203-209.

FENG, S., XUE, L., LIU, H. & LU, P. 2008. Improvement of efficiency of genetic transformation for Dunaliella salina by glass beads method. *Molecular Biology Reports,* 36**,** 1433-1439.

FERNANDEZ-SAN MILLAN, A., ORTIGOSA, S. M., HERVAS-STUBBS, S., CORRAL-MARTINEZ, P., SEGUI-SIMARRO, J. M., GAETAN, J., COURSAGET, P. & VERAMENDI, J. 2008. Human papillomavirus L1 protein expressed in tobacco chloroplasts self-assembles into virus-like particles that are highly immunogenic. *Plant Biotechnology Journal,* 6**,** 427-41.

FIGUEROA, F., SALLES, S., AGUILERA, J., JIMENEZ, C., MERCADO, J., VINEGLA, B., FLORES-MOYA, A. & ALTAMIRANO, M. 1997. Effects of solar radiation on photoinhibition and pigmentation in the red alga *Porphyra leucosticta*. *Marine Ecology Progress Series,* 151**,** 81-90.

FINAZZI, G., JOHNSON, G. N., DALL'OSTO, L., JOLIOT, P., WOLLMAN, F. A. & BASSI, R. 2004. A zeaxanthin-independent nonphotochemical quenching mechanism localized in the photosystem II core complex. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 12375-80.

FIORE, A., DALL'OSTO, L., CAZZANIGA, S., DIRETTO, G., GIULIANO, G. & BASSI, R. 2012. A quadruple mutant of *Arabidopsis* reveals a beta-carotene hydroxylation activity for LUT1/CYP97C1 and a regulatory role of xanthophylls on determination of the PSI/PSII ratio. *BMC Plant Biology,* 12**,** 50.

FISCHER, H., ROBL, I., SUMPER, M. & KRÖGER, N. 1999. Targeting and covalent modification of cell wall and membrane proteins heterologously expressed in the diatom *Cylindrotheca fusiformis* (*Bacillariophyceae*). *Journal of Phycology,* 35**,** 113-120.

FISCHER, R., STOGER, E., SCHILLBERG, S., CHRISTOU, P. & TWYMAN, R. M. 2004. Plant-based production of biopharmaceuticals. *Current Opinion in Plant Biology,* 7**,** 152-8.

FISHMAN, D., MAJUMDAR, R., MORELLO, J., PATE, R. & YANG, J. 2010. National Algal Biofuels Technology Roadmap. *In:* ENERGY, U. S. D. O. (ed.) *National Algal Biofuels Workshop. December 9-10, 2008.* Maryland, USA.

FLORIN, L., SAPP, C., STREECK, R. E. & SAPP, M. 2002. Assembly and translocation of papillomavirus capsid proteins. *Journal of Virology,* 76**,** 10009-14.

FRANKLIN, S., NGO, B., EFUET, E. & MAYFIELD, S. P. 2002. Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. *The Plant Journal,* 30**,** 733-744.

FRANKLIN, S. E. & MAYFIELD, S. P. 2004. Prospects for molecular farming in the green alga *Chlamydomonas reinhardtii*. *Current Opinion in Plant Biology,* 7**,** 159-165.

FRIEDL, T. 1997. The evolution of the Green Algae. *Plant Systematics and Evolution***,** 87-101.

FURNAS, M. J. 1990. INSITU GROWTH-RATES OF MARINE-PHYTOPLANKTON - APPROACHES TO MEASUREMENT, COMMUNITY AND SPECIES GROWTH-RATES. *Journal of Plankton Research,* 12**,** 1117-1151.

GALLOWAY, R. & METS, L. 1989. Positive selection for sensitivity of *Chlamydomonas reinhardtii* to Photosystem II-inhibiting herbicides. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 975**,** 66-71.

GATENBY, A. A., ROTHSTEIN, S. J. & BRADLEY, D. 1988. Using bacteria to analyze sequences involved in chloroplast gene-expression. *Photosynthesis Research,* 19**,** 7-22.

GEIJTENBEEK, T. B. H., KWON, D. S., TORENSMA, R., VAN VLIET, S. J., VAN DUIJNHOVEN, G. C. F., MIDDEL, J., CORNELISSEN, I. L. M. H. A., NOTTET, H. S. L. M., KEWALRAMANI, V. N., LITTMAN, D. R., FIGDOR, C. G. & VAN KOOYK, Y. 2000. DC-SIGN, a Dendritic Cell Specific HIV-1-Binding Protein that Enhances trans-Infection of T Cells. *Cell,* 100**,** 587-597.

GERBER, S., LANE, C., BROWN, D. M., LORD, E., DILORENZO, M., CLEMENTS, J. D., RYBICKI, E., WILLIAMSON, A. L. & ROSE, R. C. 2001. Human papillomavirus virus-like particles are efficient oral immunogens when coadministered with *Escherichia coli* heat-labile enterotoxin mutant R192G or CpG DNA. *Journal of Virology,* 75**,** 4752-60.

GERMAIN, V., RYLOTT, E. L., LARSON, T. R., SHERSON, S. M., BECHTOLD, N., CARDE, J. P., BRYCE, J. H., GRAHAM, I. A. & SMITH, S. M. 2001. Requirement for 3-ketoacyl-CoA thiolase-2 in peroxisome development, fatty acid beta-oxidation and breakdown of triacylglycerol in lipid bodies of *Arabidopsis* seedlings. *The Plant Journal,* 28**,** 1-12.

GHASEMI, Y., RASOUL-AMINI, S. & FOTOOH-ABADI, E. 2011. The biotransformation, biodegradation, and bioremediation of organic compounds by microalgae. *Journal of Phycology,* 47**,** 969-980.

GILMORE, A. M. & YAMAMOTO, H. Y. 1991. Zeaxanthin Formation and Energy-Dependent Fluorescence Quenching in Pea Chloroplasts under Artificially Mediated Linear and Cyclic Electron Transport. *Plant Physiology,* 96**,** 635-43.

GIOMARELLI, B., PROVVEDI, R., MEACCI, F., MAGGI, T., MEDAGLINI, D., POZZI, G., MORI, T., MCMAHON, J. B., GARDELLA, R. & BOYD, M. R. 2002. The microbicide cyanovirin-N expressed on the surface of commensal bacterium *Streptococcus gordonii* captures HIV-1. *AIDS,* 16**,** 1351-6.

GOLDBERG, A. L. 2003. Protein degradation and protection against misfolded or damaged proteins. *Nature,* 426**,** 895-9.

GOLDSCHMIDT-CLERMONT, M. 1991. Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of *Chlamydomonas*. *Nucleic Acids Research,* 19**,** 4083-9.

GOLDSCHMIDT-CLERMONT, M., RAHIRE, M. & ROCHAIX, J. D. 2008. Redundant cis-acting determinants of 3' processing and RNA stability in the chloroplast rbcL mRNA of *Chlamydomonas*. *The Plant Journal,* 53**,** 566-77.

GONG, Y., GUO, X., WAN, X., LIANG, Z. & JIANG, M. 2011a. Characterization of a novel thioesterase (PtTE) from *Phaeodactylum tricornutum*. *Journal of Basic Microbiology,* 51**,** 666-672.

GONG, Y., HU, H., GAO, Y., XU, X. & GAO, H. 2011b. Microalgae as platforms for production of recombinant proteins and valuable compounds: progress and prospects. *Journal of Industrial Microbiology & Biotechnology,* 38**,** 1879-90.

GONG, Y. & JIANG, M. 2011. Biodiesel production with microalgae as feedstock: from strains to biodiesel. *Biotechnology Letters,* 33**,** 1269-84.

GOTTESMAN, S. 1996. Proteases and their targets in Escherichia coli. *Annual Review of Genetics,* 30**,** 465-506.

GOUVEIA, L. & OLIVEIRA, A. C. 2009. Microalgae as a raw material for biofuels production. *Journal of Industrial Microbiology & Biotechnology,* 36**,** 269-274.

GRAHAM, J. M., GRAHAM, L. E., ZULKIXY, S. B., PXEGER, B. F., HOOVER, S. W. & YOSHITANI, J. 2012. Freshwater diatoms as a source of lipids for biofuels. *Journal of Industrial Microbiology & Biotechnology,* 39**,** 419-428.

GRAHAM, L. E. & WILCOX, L. 2000. Algae (vol 36, pg 788, 2000). *Journal of Phycology,* 36**,** 975-975.

GRANT, D. M., GILLHAM, N. W. & BOYNTON, J. E. 1980. Inheritance of chloroplast DNA in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America,* 77**,** 6067-6071.

GREEN, B. R. 2011. Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal,* 66**,** 34-44.

GRIFFITHS, M. & HARRISON, S. 2009. Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *Journal of Applied Phycology,* 21**,** 493-507.

GROSS, R. E. & DUGGER, W. M. 1969. Photoinhibition of growth of a yellow and a colorless form of *Chlamydomonas reinhardtii Photochemistry and Photobiology,* 10**,** 243-250.
GROSSMAN, A. R., HARRIS, E. E., HAUSER, C., LEFEBVRE, P. A., MARTINEZ, D., ROKHSAR, D., SHRAGER, J., SILFLOW, C. D., STERN, D., VALLON, O. & ZHANG, Z. 2003. *Chlamydomonas reinhardtii* at the crossroads of genomics. *Eukaryotic Cell,* 2**,** 1137-50.
GUMPEL, N. J. & PURTON, S. 1994. Playing tag with *Chlamydomonas*. *Trends in Cell Biology,* 4**,** 299-301.
GUMPEL, N. J., ROCHAIX, J. D. & PURTON, S. 1994. Studies on homologous recombination in the green alga *Chlamydomonas reinhardtii*. *Current Genetics,* 26**,** 438-42.
GUSCHINA, I. A. & HARWOOD, J. L. 2006. Lipids and lipid metabolism in eukaryotic algae. *Progress in Lipid Research,* 45**,** 160-86.
HALL, L. M., TAYLOR, K. B. & JONES, D. D. 1993. Expression of a foreign gene in *Chlamydomonas reinhardtii*. *Gene,* 124**,** 75-81.
HALLAHAN, B. J., PURTON, S., IVISON, A., WRIGHT, D. & EVANS, M. C. W. 1995. Analysis of the proposed Fe-S-X binding region of Photosystem 1 by site directed mutation of PsaA in *Chlamydomonas reinhardtii*. *Photosynthesis Research,* 46**,** 257-264.
HALLMANN, A. 2007. Algal Transgenics and Biotechnology. *Algae,* 1**,** 81-98.
HALLMANN, A. & WODNIOK, S. 2006. Swapped green algal promoters: aphVIII-based gene constructs with *Chlamydomonas* flanking sequences work as dominant selectable markers in *Volvox* and vice versa. *Plant Cell Reports,* 25**,** 582-91.
HAMBOURGER, M., MOORE, G. F., KRAMER, D. M., GUST, D., MOORE, A. L. & MOORE, T. A. 2009. Biology and technology for photochemical fuel production. *Chemical Society Reviews,* 38**,** 25-35.
HAN, X., LI, Y., SUN, X., WEI, X., SUN, Y. & WANG, Y. 2005. Studies on the heterotrophic cultivation of transgenic *Chlorella* containing the rabbit defensin gene. *Process Biochemistry,* 40**,** 3055-3060.
HARPER, A. L., VON GESJEN, S. E., LINFORD, A. S., PETERSON, M. P., FAIRCLOTH, R. S., THISSEN, M. M. & BRUSSLAN, J. A. 2004. Chlorophyllide a Oxygenase mRNA and Protein Levels Correlate with the Chlorophyll a/b Ratio in *Arabidopsis thaliana*. *Photosynthesis Research,* 79**,** 149-59.
HARRIS, E. H. 1989. *The Chlamydomonas Sourcebook,* San Diego, Academic Press.
HARRIS, E. H. 2001. *Clamydomonas* as a model organism *Annual Review of Plant Physiology and Plant Molecular Biology,* 52**,** 363-406.
HARRIS, E. H., BURKHART, B. D., GILLHAM, N. W. & BOYNTON, J. E. 1989. Antibiotic-resistance mutations in the chloroplast 16S and 23S ribosomal RNA genes of *Chlamydomonas reinhardtii* - correlation of genetic and physical maps of the chloroplast genome. *Genetics,* 123**,** 281-292.
HASNAIN, S. E., MANAVATHU, E. K. & LEUNG, W. C. 1985. DNA-mediated transformation of *Chlamydomonas reinhardi* cells: use of aminoglycoside 3'-phosphotransferase as a selectable marker. *Molecular and Cellular Biology,* 5**,** 3647-50.
HAVAUX, M., DALL'OSTO, L. & BASSI, R. 2007. Zeaxanthin has enhanced antioxidant capacity with respect to all other xanthophylls in Arabidopsis leaves and functions independent of binding to PSII antennae. *Plant Physiology,* 145**,** 1506-20.
HAWKINS, R. L. & NAKAMURA, M. 1999. Expression of Human Growth Hormone by the Eukaryotic Alga, *Chlorella*. *Current Microbiology,* 38**,** 335-341.
HE, D.-M., QIAN, K.-X., SHEN, G.-F., ZHANG, Z.-F., LI, Y.-N., SU, Z.-L. & SHAO, H.-B. 2007. Recombination and expression of classical swine fever virus (CSFV) structural protein E2 gene in Chlamydomonas reinhardtii chroloplasts. *Colloids and Surfaces B: Biointerfaces,* 55**,** 26-30.
HEATH, R. L. 1979. A Constitutive Enzyme System for Glucose Transport by *Chlorella sorokiniana*. *Plant Physiology,* 64**,** 224-7.
HEIFETZ, P. B. 2000. Genetic engineering of the chloroplast. *Biochimie,* 82**,** 655-666.
HEISER, W. 1992. Optimization of Biolistic Transformation using the Helium-driven PDS-1000/He system. *BIO-RAD US/EG Bulletin***,** 8-8.
HEITZER, M., ECKERT, A., FUHRMANN, M. & GRIESBECK, C. 2007. Influence of codon bias on the expression of foreign genes in microalgae. *Transgenic Microalgae as Green Cell Factories,* 616**,** 46-53.
HELLE, F., WYCHOWSKI, C., VU-DAC, N., GUSTAFSON, K. R., VOISSET, C. & DUBUISSON, J. 2006. Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans. *The Journal of Biological Chemistry,* 281**,** 25177-83.
HEMAISWARYA, S., RAJA, R., KUMAR, R. R., GANESAN, V. & ANBAZHAGAN, C. 2011. Microalgae: a sustainable feed source for aquaculture. *World Journal of Microbiology & Biotechnology,* 27**,** 1737-1746.
HEMSCHEMEIER, A. & HAPPE, T. 2011. Alternative photosynthetic electron transport pathways during anaerobiosis in the green alga *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1807**,** 919-926.
HEREDIA-ARROYO, T., WEI, W., RUAN, R. & HU, B. 2011. Mixotrophic cultivation of Chlorella vulgaris and its potential application for the oil accumulation from non-sugar materials. *Biomass and Bioenergy,* 35**,** 2245-2253.
HESS, J. L. & TOLBERT, N. E. 1967. Changes in Chlorophyll a/b Ratio and Products of CO(2) Fixation by Algae Grown in Blue or Red Light. *Plant Physiology,* 42**,** 1123-30.
HILL, J., NELSON, E., TILMAN, D., POLASKY, S. & TIFFANY, D. 2006. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 11206-10.
HIRSCHBERG, J. 2001. Carotenoid biosynthesis in flowering plants. *Current Opinion in Plant Biology,* 4**,** 210-8.

HOANG, L. T., LIANG, J. S., CARRERAS, I. & SIPE, J. D. 1995. Novel SAA gene products in Mus caroli mice defined by rapid amplification of complementary DNA ends (3′ and 5′ RACE). *Amyloid-International Journal of Experimental and Clinical Investigation,* 2**,** 75-82.

HORTENSTEINER, S., CHINNER, J., MATILE, P., THOMAS, H. & DONNISON, I. S. 2000. Chlorophyll breakdown in *Chlorella protothecoides*: characterization of degreening and cloning of degreening-related genes. *Plant Molecular Biology,* 42**,** 439-50.

HU, Q., SOMMERFELD, M., JARVIS, E., GHIRARDI, M., POSEWITZ, M., SEIBERT, M. & DARZINS, A. 2008a. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal,* 54**,** 621-639.

HU, Q., SOMMERFELD, M., JARVIS, E., GHIRARDI, M., POSEWITZ, M., SEIBERT, M. & DARZINS, A. 2008b. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal,* 54**,** 621-39.

HUANG, F. C., KLAUS, S. M., HERZ, S., ZOU, Z., KOOP, H. U. & GOLDS, T. J. 2002. Efficient plastid transformation in tobacco using the aphA-6 gene and kanamycin selection. *Molecular Genetics and Genomics,* 268**,** 19-27.

HUSS, V. A. R. & JAHNKE, K. D. 1994. DNA-base modification in Chlorella. *European Journal of Phycology,* 29**,** 87-92.

IAMTHAM, S. & DAY, A. 2000. Removal of antibiotic resistance genes from transgenic tobacco plastids. *Nature Biotechnology,* 18**,** 1172-6.

IKEMURA, T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution,* 2**,** 13-34.

ILLMAN, A. M., SCRAGG, A. H. & SHALES, S. W. 2000. Increase in *Chlorella* strains calorific values when grown in low nitrogen medium. *Enzyme and Microbial Technology,* 27**,** 631-635.

IMAMURA, S., HAGIWARA, D., SUZUKI, F., KURANO, N. & HARAYAMA, S. 2012. Genetic transformation of *Pseudochoricystis ellipsoidea*, an aliphatic hydrocarbon-producing green alga. *Journal of General and Applied Microbiology,* 58**,** 1-10.

INOUE, K., DREYFUSS, B. W., KINDLE, K. L., STERN, D. B., MERCHANT, S. & SODEINDE, O. A. 1997. Ccs1, a nuclear gene required for the post-translational assembly of chloroplast c-type cytochromes. *The Journal of Biological Chemistry,* 272**,** 31747-31754.

IWAMOTO, H. 2007. Industrial Production of Microalgal Cell-Mass and Secondary Products - Major Industrial Species: Chlorella. *Handbook of Microalgal Culture.* Blackwell Publishing Ltd.

JAMES, O. O., MAITY, S., USMAN, L. A., AJANAKU, K. O., AJANI, O. O., SIYANBOLA, T. O., SAHU, S. & CHAUBEY, R. 2010. Towards the conversion of carbohydrate biomass feedstocks to biofuels via hydroxylmethylfurfural. *Energy & Environmental Science,* 3**,** 1833-1850.

JARVIS, E. E. & BROWN, L. M. 1991. Transient expression of firefly luciferase in protoplasts of the green alga *Chlorella ellipsoidea*. *Current Genetics,* 19**,** 317-321.

JENSON, A. B., GEYER, S., SUNDBERG, J. P. & GHIM, S.-J. 2001. Human Papillomavirus and Skin Cancer. *Journal of Investigative Dermatology Symposium* 6**,** 203-206.

JOHANNINGMEIER, U., SOPP, G., BRAUNER, M., ALTENFELD, U., ORAWSKI, G. & OETTMEIER, W. 2000. Herbicide resistance and supersensitivity in Ala(250) or Ala(251) mutants of the D1 protein in Chlamydomonas reinhardtii. *Pesticide Biochemistry and Physiology,* 66**,** 9-19.

JOHNSON, M. P. & RUBAN, A. V. 2009. Photoprotective energy dissipation in higher plants involves alteration of the excited state energy of the emitting chlorophyll(s) in the light harvesting antenna II (LHCII). *The Journal of Biological Chemistry,* 284**,** 23592-601.

JONES, A., DAVIES, H. M. & VOELKER, T. A. 1995. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. *The Plant Cell,* 7**,** 359-71.

JOSUE, J. S. & FRANK, H. A. 2002. Direct determination of the S-1 excited-state energies of xanthophylls by low-temperature fluorescence spectroscopy. *Journal of Physical Chemistry A,* 106**,** 4815-4824.

KALSCHEUER, R., STOLTING, T. & STEINBUCHEL, A. 2006. Microdiesel: Escherichia coli engineered for fuel production. *Microbiology,* 152**,** 2529-36.

KAMIYA, A. & KOWALLIK, W. 1987. Photoinhibition of Glucose Uptake in *Chlorella*. *Plant and Cell Physiology,* 28**,** 611-619.

KAN, K. S. & THORNBER, J. P. 1976. The Light-harvesting Chlorophyll a/b-Protein Complex of Chlamydomonas reinhardii. *Plant Physiology,* 57**,** 47-52.

KASAI, S., YOSHIMURA, S., ISHIKURA, K., TAKAOKA, Y., KOBAYASHI, K., KATO, K. & SHINMYO, A. 2003. Effect of coding regions on chloroplast gene expression in *Chlamydomonas reinhardtii*. *Journal of Bioscience and Bioengineering,* 95**,** 276-282.

KATHIRESAN, S., CHANDRASHEKAR, A., RAVISHANKAR, G. A. & SARADA, R. 2009. Agrobacterium-mediated transformation in the green alga *Haematococcus pluvialis* (*Chlorophyceae, Volvocales*). *Journal of Phycology,* 45**,** 642-649.

KAWAI, S., MURAO, S., MOCHIZUKI, M., SHIBUYA, I., YANO, K. & TAKAGI, M. 1992. Drastic alteration of cycloheximide sensitivity by substitution of one amino acid in the L41 ribosomal protein of yeasts. *Journal of Bacteriology,* 174**,** 254-262.

KEPLER, T. B. & PERELSON, A. S. 1995. Modeling and optimisation of populations subject to time-dependent mutation. . *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 8219-8223.

KESSLER, E. 1953. [The mechanism of nitrate reduction by green algae. II. Comparative physiological studies.]. *Archives of Microbiology,* 19**,** 438-57.

KESSLER, E. 1978. Physiological and biochemical contributions to the taxonomy of the genus *Chlorella*. *Archives of Microbiology,* 119**,** 13-16.

KESSLER, E. 1985. Upper limits of temperature for growth in Chlorella (Chlorophyceae) *Plant Systematics and Evolution,* 151**,** 67-71.
KILIAN, O., BENEMANN, C. S. E., NIYOGI, K. K. & VICK, B. 2011. High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis sp*. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 21265-21269.
KIM, D. H., KIM, Y. T., CHO, J. J., BAE, J. H., HUR, S. B., HWANG, I. & CHOI, T. J. 2002. Stable integration and functional expression of flounder growth hormone gene in transformed microalga, *Chlorella ellipsoidea*. *Marine Biotechnology,* 4**,** 63-73.
KINDLE, K. L. 1990. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 1228-1232.
KINDLE, K. L., RICHARDS, K. L. & STERN, D. B. 1991. Engineering the chloroplast genome: techniques and capabilities for chloroplast transformation in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America,* 88**,** 1721-5.
KIRNBAUER, R., BOOY, F., CHENG, N., LOWY, D. R. & SCHILLER, J. T. 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 12180-4.
KLAUS, D., OHLROGGE, J. B., NEUHAUS, H. E. & DORMANN, P. 2004. Increased fatty acid production in potato by engineering of acetyl-CoA carboxylase. *Planta,* 219**,** 389-96.
KLAUS, S. M., HUANG, F. C., EIBL, C., KOOP, H. U. & GOLDS, T. J. 2003. Rapid and proven production of transplastomic tobacco plants by restoration of pigmentation and photosynthesis. *The Plant Journal,* 35**,** 811-21.
KLEIMAN, R. 1999. *Chemistry of New Industrial Oilseed Crops,* Portland, OR, Timber Press.
KODE, V., MUDD, E. A., IAMTHAM, S. & DAY, A. 2006. Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. *The Plant Journal,* 46**,** 901-9.
KOHL, T. O., HITZEROTH, II, CHRISTENSEN, N. D. & RYBICKI, E. P. 2007. Expression of HPV-11 L1 protein in transgenic *Arabidopsis thaliana* and *Nicotiana tabacum*. *BMC Biotechnology,* 7**,** 56-56.
KOSOUROV, S. N., GHIRARDI, M. L. & SEIBERT, M. 2011. A truncated antenna mutant of Chlamydomonas reinhardtii can produce more hydrogen than the parental strain. *International Journal of Hydrogen Energy,* 36**,** 2044-2048.
KOTZE, L., SMITH, J. J., DEN HAAN, R., VAN ZYL, W. H. & GOERGENS, J. F. 2011. Expression of human papillomavirus type 16 (HPV16) L1 protein in Pichia pastoris. *African Journal of Biotechnology,* 10**,** 214-219.
KRUSE, O., RUPPRECHT, J., MUSSGNUG, J. H., DISMUKES, G. C. & HANKAMER, B. 2005. Photosynthesis: a blueprint for solar energy capture and biohydrogen production technologies. *Photochemical & Photobiological Sciences : Official Journal of the European Photochemistry Association and the European Society for Photobiology,* 4**,** 957-70.
KUMAR, M. S., MIAO, Z. H. & WYATT, S. K. 2010. Influence of nutrient loads, feeding frequency and inoculum source on growth of Chlorella vulgaris in digested piggery effluent culture medium. *Bioresource Technology,* 101**,** 6012-8.
KUMAR, S., DHINGRA, A. & DANIELL, H. 2004a. Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. *Plant Molecular Biology,* 56**,** 203-16.
KUMAR, S. V., MISQUITTA, R. W., REDDY, V. S., RAO, B. J. & RAJAM, M. V. 2004b. Genetic transformation of the green alga - *Chlamydomonas reinhardtii* by *Agrobacterium tumefaciens. Plant Science,* 166**,** 731-738.
KUNZ, W. F. 1972. Response of the alga Chlorella sorokiniana to 60 Co gamma radiation. *Nature,* 236**,** 178-9.
KURATA, T., KAWABATA-AWAI, C., SAKURADANI, E., SHIMIZU, S., OKADA, K. & WADA, T. 2003. The YORE-YORE gene regulates multiple aspects of epidermal cell differentiation in *Arabidopsis*. *The Plant Journal,* 36**,** 55-66.
KUSNADI, A. R., NIKOLOV, Z. L. & HOWARD, J. A. 1997. Production of recombinant proteins in transgenic plants: Practical considerations. *Biotechnology and Bioengineering,* 56**,** 473-84.
LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.
LALUCAT, J., IMPERIAL, J. & PARES, R. 1984. Utilization of light for the assimilation of organic matter in *Chlorella sp*. VJ79. *Biotechnology and Bioengineering,* 26**,** 677-81.
LAPIDOT, M., RAVEH, D., SIVAN, A., ARAD, S. M. & SHAPIRA, M. 2002. Stable Chloroplast Transformation of the Unicellular Red Alga *Porphyridium* Species. *Plant Physiology,* 129**,** 7-12.
LARKUM, A. W. D., DOUGLAS, S. E. & RAVEN, J. A. 2003. *Photosynthesis in algae,* Kluwer Academic.
LARKUM, A. W. D., ROSS, I. L., KRUSE, O. & HANKAMER, B. 2012. Selection, breeding and engineering of microalgae for bioenergy and biofuel production. *Trends in Biotechnology,* 30**,** 198-205.
LEE, D., CHEN, A. & NAIR, R. 2008. Genetically engineered crops for biofuel production: regulatory perspectives. *Biotechnology & Genetic Engineering Reviews,* 25**,** 331-61.
LEELAVATHI, S. & REDDY, V. S. 2003. Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. *Molecular Breeding,* 11**,** 49-58.
LEI, A., CHEN, H., SHEN, G., HU, Z., CHEN, L. & WANG, J. 2012. Expression of fatty acid synthesis genes and fatty acid accumulation in *Haematococcus pluvialis* under different stressors. *Biotechnology for Biofuels,* 5.
LENZ, P., DAY, P. M., PANG, Y. Y., FRYE, S. A., JENSEN, P. N., LOWY, D. R. & SCHILLER, J. T. 2001. Papillomavirus-like particles induce acute activation of dendritic cells. *Journal of Immunology,* 166**,** 5346-55.
LENZI, P., SCOTTI, N., ALAGNA, F., TORNESELLO, M. L., POMPA, A., VITALE, A., STRADIS, A., MONTI, L., GRILLO, S., BUONAGURO, F. M., MALIGA, P. & CARDI, T. 2008. Translational fusion of chloroplast-

expressed human papillomavirus type 16 L1 capsid protein enhances antigen accumulation in transplastomic tobacco. *Transgenic Research,* 17**,** 1091-1102.
LEON-BANARES, R., GONZALEZ-BALLESTER, D., GALVAN, A. & FERNANDEZ, E. 2004. Transgenic microalgae as green cell-factories. *Trends in Biotechnology,* 22**,** 45-52.
LERCHE, K. & HALLMANN, A. 2009. Stable nuclear transformation of *Gonium pectorale*. *BMC Biotechnology,* 9**,** 64.
LI, M., PATTON, D. L., COSGROVE-SWEENEY, Y., RATNER, D., ROHAN, L. C., COLE, A. M., TARWATER, P. M., GUPTA, P. & RAMRATNAM, B. 2011. Incorporation of the HIV-1 Microbicide Cyanovirin-N in a Food Product. *Jaids-Journal of Acquired Immune Deficiency Syndromes,* 58**,** 379-384.
LI, W., RUF, S. & BOCK, R. 2010. Chloramphenicol acetyltransferase as selectable marker for plastid transformation. *Plant Molecular Biology*.
LI, Y., HORSMAN, M., WU, N., LAN, C. Q. & DUBOIS-CALERO, N. 2008. Biofuels from microalgae. *Biotechnology Progress,* 24**,** 815-20.
LIANG, S., LIU, X., CHEN, F. & CHEN, Z. 2004. Current microalgal health food R & D activities in China

Asian Pacific Phycology in the 21st Century: Prospects and Challenges. *In:* ANG, P. O. (ed.). Springer Netherlands.
LIANG, Y., SARKANY, N. & CUI, Y. 2009. Biomass and lipid productivities of Chlorella vulgaris under autotrophic, heterotrophic and mixotrophic growth conditions. *Biotechnology Letters,* 31**,** 1043-1049.
LICHTENTHALER, H. K. & BUSCHMANN, C. 2001. Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS Spectroscopy. *Current Protocols in Food Analytical Chemistry.* John Wiley & Sons, Inc.
LIN, H. & GOODENOUGH, U. W. 2007. Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, MID and MTD1. *Genetics,* 176**,** 913-925.
LINDAHL, M., YANG, D. H. & ANDERSSON, B. 1995. Regulatory proteolysis of the major light-harvesting chlorophyll a/b protein of photosystem II by a light-induced membrane-associated enzymic system. *European Journal of Biochemistry / FEBS,* 231**,** 503-9.
LINDBERG, P., PARK, S. & MELIS, A. 2010. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. *Metabolic Engineering,* 12**,** 70-79.
LIU, H. L., LI, W. S., LEI, T., ZHENG, J., ZHANG, Z., YAN, X. F., WANG, Z. Z., WANG, Y. L. & SI, L. S. 2005. Expression of human papillomavirus type 16 L1 protein in transgenic tobacco plants. *Acta Biochimica et Biophysica Sinica (Shanghai),* 37**,** 153-8.
LIU, L., WANG, Y., ZHANG, Y., CHEN, X., ZHANG, P. & MA, S. 2012. Development of a New Method for Genetic Transformation of the Green Alga *Chlorella ellipsoidea*. *Molecular Biotechnology***,** 1-9.
LIU, Y.-G. & CHEN, Y. 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques,* 43**,** 649-+.
LIU, Z., YAN, H., WANG, K., KUANG, T., ZHANG, J., GUI, L., AN, X. & CHANG, W. 2004. Crystal structure of spinach major light-harvesting complex at 2.72Å resolution. *Nature,* 428**,** 287-292.
LOPEZ-RODAS, V., AGRELO, M., CARRILLO, E., FERRERO, L. M., LARRAURI, A., MARTIN-OTERO, L. & COSTAS, E. 2001. Resistance of microalgae to modern water contaminants as the result of rare spontaneous mutations. *European Journal of Phycology,* 36**,** 179-190.
LOPPES, R. 1968. Ethyl methanesulfonate: an effective mutagen in *Chlamydomonas reinhardi*. *Molecular & General Genetics,* 102**,** 229-31.
LU, X., VORA, H. & KHOSLA, C. 2008. Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metabolic Engineering,* 10**,** 333-9.
LU, Y. M., LI, J., XUE, L. X., YAN, H. X., YUAN, H. J. & WANG, C. 2011. A duplicated carbonic anhydrase 1 (DCA1) promoter mediates the nitrate reductase gene switch of *Dunaliella salina*. *Journal of Applied Phycology,* 23**,** 673-680.
LUMBRERAS, V. & PURTON, S. 1998. Recent advances in Chlamydomonas transgenics. *Protist,* 149**,** 23-27.
LURIA, S. E. & DELBRUCK, M. 1943. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics,* 28**,** 491-511.
MACLEAN, J., KOEKEMOER, M., OLIVIER, A. J., STEWART, D., HITZEROTH, II, RADEMACHER, T., FISCHER, R., WILLIAMSON, A. L. & RYBICKI, E. P. 2007. Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *Journal of General Virology,* 88**,** 1460-9.
MADOKA, Y., TOMIZAWA, K., MIZOI, J., NISHIDA, I., NAGANO, Y. & SASAKI, Y. 2002. Chloroplast transformation with modified accD operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco. *Plant and Cell Physiology,* 43**,** 1518-25.
MAHALINGAM, R. & FEDOROFF, N. 2001. Screening insertion libraries for mutations in many genes simultaneously using DNA microarrays. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 7420-7425.
MALKIN, S., ARMOND, P. A., MOONEY, H. A. & FORK, D. C. 1981. Photosystem II Photosynthetic Unit Sizes from Fluorescence Induction in Leaves : CORRELATION TO PHOTOSYNTHETIC CAPACITY. *Plant Physiology,* 67**,** 570-9.
MANNING, J. E., WOLSTENHOLME, D. R., RYAN, R. S., HUNTER, J. A. & RICHARDS, O. C. 1971. Circular chloroplast DNA from Euglena gracilis. *Proceedings of the National Academy of Sciences of the United States of America,* 68**,** 1169-73.
MANUELL, A., BELIGNI, M. V., YAMAGUCHI, K. & MAYFIELD, S. P. 2004. Regulation of chloroplast translation: interactions of RNA elements, RNA-binding proteins and the plastid ribosome. *Biochemical Society Transactions,* 32**,** 601-5.

MANUELL, A. L., BELIGNI, M. V., ELDER, J. H., SIEFKER, D. T., TRAN, M., WEBER, A., MCDONALD, T. L. & MAYFIELD, S. P. 2007. Robust expression of a bioactive mammalian protein in *Chlamydomonas* chloroplast. *Plant Biotechnology Journal,* 5**,** 402-12.

MARES-PERLMAN, J. A. & ERDMAN, J. W. 2002. Can lutein protect against chronic disease? A multidisciplinary approach involving basic science and epidemiology to weigh evidence and design analytic strategies. Introduction. *Journal of Nutrition,* 132**,** 517S-517S.

MARTIN, W. & KOWALLIK, K. V. 1999. Annotated English translation of Mereschkowsky's 1905 paper "Uber Natur und Ursprung der Chromatophoren im Pflanzenreiche". *European Journal of Phycology,* 34**,** 287-295.

MARTIN, W. & SCHNARRENBERGER, C. 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Current Genetics,* 32**,** 1-18.

MATA, T. M., MARTINS, A. N. A. & CAETANO, N. S. 2010. Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews,* 14**,** 217-232.

MATSUKAWA, R., HOTTA, M., MASUDA, Y., CHIHARA, M. & KARUBE, I. 2000. Antioxidants from carbon dioxide fixing Chlorella sorokiniana. *Journal of Applied Phycology,* 12**,** 263-267.

MATTOX, K. R. 1984. *Classification of the green algae: a concept based on comparative cytology.,* London, Academic Press.

MAUL, J. E., LILLY, J. W., CUI, L., DEPAMPHILIS, C. W., MILLER, W., HARRIS, E. H. & STERN, D. B. 2002. The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *The Plant Cell,* 14**,** 2659-79.

MAYFIELD, S. P. & FRANKLIN, S. E. 2005. Expression of human antibodies in eukaryotic micro-algae. *Vaccine,* 23**,** 1828-32.

MAYFIELD, S. P., FRANKLIN, S. E. & LERNER, R. A. 2003. Expression and assembly of a fully active antibody in algae. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 438-42.

MAYFIELD, S. P. & KINDLE, K. L. 1990. Stable nuclear transformation of *Chlamydomonas reinhardtii* by using a *C. reinhardtii* gene as the selectable marker. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 2087-2091.

MAYFIELD, S. P., MANUELL, A. L., CHEN, S., WU, J., TRAN, M., SIEFKER, D., MUTO, M. & MARIN-NAVARRO, J. 2007. *Chlamydomonas reinhardtii* chloroplasts as protein factories. *Current Opinion in Biotechnology,* 18**,** 126-133.

MAYFIELD, S. P., YOHN, C. B., COHEN, A. & DANON, A. 1995. Regulation of chloroplast gene expression. *Annual Review of Plant Physiology and Plant Molecular Biology,* 46**,** 147-166.

MCCLELLAN, A. J. & FRYDMAN, J. 2001. Molecular chaperones and the art of recognizing a lost cause. *Nature Cell Biology,* 3**,** E51-3.

MCCUSKER, J. H. & HABER, J. E. 1988. Cycloheximide-resistant temperature-sensitive lethal mutations of Saccharomyces cerevisiae. *Genetics,* 119**,** 303-303.

MEGHARAJ, M., MADHAVI, D. R., SREENIVASULU, C., UMAMAHESWARI, A. & VENKATESWARLU, K. 1994. Biodegradation of methyl parathion by soil isolates of microalgae and cyanobacteria. *Bulletin of Environmental Contamination and Toxicology,* 53**,** 292-297.

MEGHARAJ, M., VENKATESWARLU, K. & RAO, A. S. 1987. Metabolism of monocrotophos and quinalphos by algae isolated from soil. *Bulletin of Environmental Contamination and Toxicology,* 39**,** 251-256.

MEHLITZ, T. H. 2009. *Temperature Influence and Heat Management Requirements of Microalgae Cultivation in Photobioreactors.* MSc, California Polytechnic State University (CalPoly) San Luis.

MELIS, A. 1999. Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage ? *Trends in Plant Science,* 4**,** 130-135.

MELIS, A. 2009. Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency. *Plant Science,* 177**,** 272-280.

MELIS, A., NEIDHARDT, J. & BENEMANN, J. 1998. *Dunaliella salina* (*Chlorophyta*) with small chlorophyll antenna sizes exhibit higher photosynthetic productivities and photon use efficiencies than normally pigmented cells. *Journal of Applied Phycology,* 10**,** 515-525.

MELIS, A., ZHANG, L., FORESTIER, M., GHIRARDI, M. L. & SEIBERT, M. 2000. Sustained photobiological hydrogen gas production upon reversible inactivation of oxygen evolution in the green alga *Chlamydomonas reinhardtii*. *Plant Physiology,* 122**,** 127-135.

MERCHANT, S. S., PROCHNIK, S. E., VALLON, O., HARRIS, E. H., KARPOWICZ, S. J., WITMAN, G. B., TERRY, A., SALAMOV, A., FRITZ-LAYLIN, L. K., MARECHAL-DROUARD, L., MARSHALL, W. F., QU, L. H., NELSON, D. R., SANDERFOOT, A. A., SPALDING, M. H., KAPITONOV, V. V., REN, Q., FERRIS, P., LINDQUIST, E., SHAPIRO, H., LUCAS, S. M., GRIMWOOD, J., SCHMUTZ, J., CARDOL, P., CERUTTI, H., CHANFREAU, G., CHEN, C. L., COGNAT, V., CROFT, M. T., DENT, R., DUTCHER, S., FERNANDEZ, E., FUKUZAWA, H., GONZALEZ-BALLESTER, D., GONZALEZ-HALPHEN, D., HALLMANN, A., HANIKENNE, M., HIPPLER, M., INWOOD, W., JABBARI, K., KALANON, M., KURAS, R., LEFEBVRE, P. A., LEMAIRE, S. D., LOBANOV, A. V., LOHR, M., MANUELL, A., MEIER, I., METS, L., MITTAG, M., MITTELMEIER, T., MORONEY, J. V., MOSELEY, J., NAPOLI, C., NEDELCU, A. M., NIYOGI, K., NOVOSELOV, S. V., PAULSEN, I. T., PAZOUR, G., PURTON, S., RAL, J. P., RIANO-PACHON, D. M., RIEKHOF, W., RYMARQUIS, L., SCHRODA, M., STERN, D., UMEN, J., WILLOWS, R., WILSON, N., ZIMMER, S. L., ALLMER, J., BALK, J., BISOVA, K., CHEN, C. J., ELIAS, M., GENDLER, K., HAUSER, C., LAMB, M. R., LEDFORD, H., LONG, J. C., MINAGAWA, J., PAGE, M. D., PAN, J., POOTAKHAM, W., ROJE, S., ROSE, A., STAHLBERG, E., TERAUCHI, A. M., YANG, P., BALL, S., BOWLER, C., DIECKMANN, C. L., GLADYSHEV, V. N., GREEN, P., JORGENSEN, R.,

MAYFIELD, S., MUELLER-ROEBER, B., RAJAMANI, S., SAYRE, R. T., BROKSTEIN, P., et al. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science,* 318**,** 245-50.

MERGENHAGEN, D. & MERGENHAGEN, E. 1989. The expression of a circadian rhythm in two strains of Chlamydomonas reinhardii in space. *Advances in Space Research : the Official Journal of the Committee on Space Research (COSPAR),* 9**,** 261-70.

METZGER, P., BERKALOFF, C., CASADEVALL, E. & COUTE, A. 1985. Alkadiene- and botryococcene-producing races of wild strains of *Botryococcus braunii*. *Phytochemistry,* 24**,** 2305-2312.

METZGER, P. & LARGEAU, C. 2005. Botryococcus braunii: a rich source for hydrocarbons and related ether lipids. *Applied Microbiology and Biotechnology,* 66**,** 486-496.

METZGER, P., RAGER, M. N. & LARGEAU, C. 2007. Polyacetals based on polymethylsqualene diols, precursors of algaenan in Botryococcus braunii race B. *Organic Geochemistry,* 38**,** 566-581.

MICHELET, L., LEFEBVRE-LEGENDRE, L., BURR, S. E., ROCHAIX, J. D. & GOLDSCHMIDT-CLERMONT, M. 2011. Enhanced chloroplast transgene expression in a nuclear mutant of *Chlamydomonas*. *Plant Biotechnology Journal,* 9**,** 565-74.

MICHINAKA, Y., SHIMAUCHI, T., AKI, T., NAKAJIMA, T., KAWAMOTO, S., SHIGETA, S., SUZUKI, O. & ONO, K. 2003. Extracellular secretion of free fatty acids by disruption of a fatty acyl-CoA synthetase gene in *Saccharomyces cerevisiae*. *Journal of Bioscience and Bioengineering,* 95**,** 435-40.

MILLER, C. G., STRAUCH, K. L., KUKRAL, A. M., MILLER, J. L., WINGFIELD, P. T., MAZZEI, G. J., WERLEN, R. C., GRABER, P. & MOVVA, N. R. 1987. N-terminal methionine-specific peptidase in Salmonella typhimurium. *Proceedings of the National Academy of Sciences of the United States of America,* 84**,** 2718-22.

MINKO, I., HOLLOWAY, S. P., NIKAIDO, S., CARTER, M., ODOM, O. W., JOHNSON, C. H. & HERRIN, D. L. 1999. Renilla luciferase as a vital reporter for chloroplast gene expression in *Chlamydomonas*. *Molecular & General Genetics,* 262**,** 421-425.

MINODA, A., SAKAGAMI, R., YAGISAWA, F., KUROIWA, T. & TANAKA, K. 2004. Improvement of Culture Conditions and Evidence for Nuclear Transformation by Homologous Recombination in a Red Alga, Cyanidioschyzon merolae 10D. *Plant and Cell Physiology,* 45**,** 667-671.

MISON, D. & CURLING, J. 2000. The industrial production costs of recombinant therapeutic proteins expressed in transgenic corn. *Biopharm-the Applied Technologies of Biopharmaceutical Development,* 13**,** 48-54.

MIYAGAWA, Y., TAMOI, M. & SHIGEOKA, S. 2001. Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nature Biotechnology,* 19**,** 965-969.

MIYAGAWA-YAMAGUCHI, A., OKAMI, T., KIRA, N., YAMAGUCHI, H., OHNISHI, K. & ADACHI, M. 2011. Stable nuclear transformation of the diatom *Chaetoceros sp*. *Phycological Research,* 59**,** 113-119.

MOAZED, D. & NOLLER, H. F. 1987. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature,* 327**,** 389-94.

MOELLERING, E. R., MILLER, R. & BENNING, C. 2009. Molecular Genetics of Lipid Metabolism in the Model Green Alga Chlamydomonas reinhardtii. *In:* WADA, H. & MURATA, N. (eds.) *Lipids in Photosynthesis: Essential and Regulatory Functions.*

MORI, T., BARRIENTOS, L. G., HAN, Z., GRONENBORN, A. M., TURPIN, J. A. & BOYD, M. R. 2002. Functional homologs of cyanovirin-N amenable to mass production in prokaryotic and eukaryotic hosts. *Protein Expression and Purification,* 26**,** 42-9.

MORI, T., GUSTAFSON, K. R., PANNELL, L. K., SHOEMAKER, R. H., WU, L., MCMAHON, J. B. & BOYD, M. R. 1998. Recombinant production of cyanovirin-N, a potent human immunodeficiency virus-inactivating protein derived from a cultured cyanobacterium. *Protein Expression and Purification,* 12**,** 151-8.

MOROWVAT, M. H., RASOUL-AMINI, S. & GHASEMI, Y. 2010. *Chlamydomonas* as a "new" organism for biodiesel production. *Bioresource Technology,* 101**,** 2059-2062.

MORTON, B. R. 1998. Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages. *Journal of Molecular Evolution,* 46**,** 449-59.

MORTON, B. R. & SO, B. G. 2000. Codon usage in plastid genes is correlated with context, position within the gene, and amino acid content. *Journal of Molecular Evolution,* 50**,** 184-193.

MOUGET, J.-L., DAKHAMA, A., LAVOIE, M. C. & DE LA NO√°E, J. L. 1995. Algal growth enhancement by bacteria: Is consumption of photosynthetic oxygen involved? *FEMS Microbiology Ecology,* 18**,** 35-43.

MOZZO, M., MANTELLI, M., PASSARINI, F., CAFFARRI, S., CROCE, R. & BASSI, R. 2010. Functional analysis of Photosystem I light-harvesting complexes (Lhca) gene products of *Chlamydomonas reinhardtii*. *Biochimica Et Biophysica Acta,* 1797**,** 212-21.

MULLER, J., FRIEDL, T., HEPPERLE, D., LORENZ, M. & DAY, J. G. 2005. Distinction between multiple isolates of Chlorella vulgaris (Chlorophyta, Trebouxiophyceae) and testing for conspecificity using amplified fragment length polymorphism and its rDNA sequences. *Journal of Phycology,* 41**,** 1236-1247.

MULLER, M., ZHOU, J., REED, T. D., RITTMULLER, C., BURGER, A., GABELSBERGER, J., BRASPENNING, J. & GISSMANN, L. 1997. Chimeric papillomavirus-like particles. *Virology,* 234**,** 93-111.

MULLER, P., LI, X. P. & NIYOGI, K. K. 2001. Non-photochemical quenching. A response to excess light energy. *Plant Physiology,* 125**,** 1558-66.

MULLINEAUX, C. W. & EMLYN-JONES, D. 2005. State transitions: an example of acclimation to low-light stress. *Journal of Experimental Botany,* 56**,** 389-393.

MURATA, N. 1969. Control of excitation transfer in photosynthesis. I. Light-induced change of chlorophyll a fluorescence in *Porphyridium cruentum*. *Biochimica Et Biophysica Acta,* 172**,** 242-51.

MUSSGNUG, J. H., THOMAS-HALL, S., RUPPRECHT, J., FOO, A., KLASSEN, V., MCDOWALL, A., SCHENK, P. M., KRUSE, O. & HANKAMER, B. 2007. Engineering photosynthetic light capture: impacts on improved solar energy to biomass conversion. *Plant Biotechnology Journal,* 5**,** 802-14.

MUTO, M., HENRY, R. E. & MAYFIELD, S. P. 2009. Accumulation and processing of a recombinant protein designed as a cleavable fusion to the endogenous Rubisco LSU protein in *Chlamydomonas* chloroplast. *BMC Biotechnology,* 9**,** 26.

MUTOH, E., OHTA, A. & TAKAGI, M. 1998. Studies on cycloheximide-sensitive and cycloheximide-resistant ribosomes in the yeast Candida maltosa. *Gene,* 224**,** 9-15.

MYERS, J. 1946. Oxidative assimilation in relation to photosynthesis in Chlorella *American Journal of Botany,* 33**,** 838-838.

MYERS, J. & CRAMER, M. L. 1947. RECONSIDERATION OF THE PHOTOSYNTHETIC MECHANISM IN CHLORELLA. *Science,* 105**,** 552-552.

NAKADA, T., MISAWA, K. & NOZAKI, H. 2008. Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on exhaustive 18S rRNA phylogenetic analyses. *Molecular Phylogenetics and Evolution,* 48**,** 281-291.

NELSON, J. A., SAVEREIDE, P. B. & LEFEBVRE, P. A. 1994. The CRY1 gene in Chlamydomonas reinhardtii: structure and use as a dominant selectable marker for nuclear transformation. *Molecular and Cellular Biology,* 14**,** 4011-4019.

NEWMAN, S. M., BOYNTON, J. E., GILLHAM, N. W., RANDOLPH-ANDERSON, B. L., JOHNSON, A. M. & HARRIS, E. H. 1990. Transformation of chloroplast ribosomal RNA genes in *Chlamydomonas*: molecular and genetic characterization of integration events. *Genetics,* 126**,** 875-88.

NEWMAN, S. M., HARRIS, E. H., JOHNSON, A. M., BOYNTON, J. E. & GILLHAM, N. W. 1992. Nonrandom distribution of chloroplast recombination events in Chlamydomonas reinhardtii: evidence for a hotspot and an adjacent cold region. *Genetics,* 132**,** 413-29.

NICKELSEN, J., FLEISCHMANN, M., BOUDREAU, E., RAHIRE, M. & ROCHAIX, J. D. 1999. Identification of cis-acting RNA leader elements required for chloroplast *psbD* gene expression in *Chlamydomonas*. *The Plant Cell,* 11**,** 957-70.

NICOLAS, P., FREYSSINET, G. & NIGON, V. 1980. Effect of Light on Glucose Utilization by Euglena gracilis. *Plant Physiology,* 65**,** 631-4.

NINLAYARN, T. 2012. *Chloroplast engineering in the microlga Chlamydomonas reinhardtii: molecular tools and applications.* Docroral Dissertation, University College London.

NOZAKI, H., OHTA, N., TAKANO, H. & WATANABE, M. M. 1999. Reexamination of phylogenetic relationships within the colonial Volvocales (Chlorophyta): An analysis of atpB and rbcL gene sequences. *Journal of Phycology,* 35**,** 104-112.

NUGENT, G. D., TEN HAVE, M., VAN DER GULIK, A., DIX, P. J., UIJTEWAAL, B. A. & MORDHORST, A. P. 2005. Plastid transformants of tomato selected using mutations affecting ribosome structure. *Plant Cell Reports,* 24**,** 341-9.

NUGENT, J. M. & PALMER, J. D. 1991. RNA-mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. *Cell,* 66**,** 473-81.

O'KEEFE, B. R., SMEE, D. F., TURPIN, J. A., SAUCEDO, C. J., GUSTAFSON, K. R., MORI, T., BLAKESLEE, D., BUCKHEIT, R. & BOYD, M. R. 2003. Potent anti-influenza activity of cyanovirin-N and interactions with viral hemagglutinin. *Antimicrobial Agents and Chemotherapy,* 47**,** 2518-25.

OCHMAN, H., GERBER, A. S. & HARTL, D. L. 1988. Genetic applications of an inverse polymerase chain-reaction. *Genetics,* 120**,** 621-623.

OHLROGGE, J. B. & JAWORSKI, J. G. 1997. Regulation of Fatty Acid Synthesis. *Annual Review of Plant Physiology and Plant Molecular Biology,* 48**,** 109-136.

OKULLO, S. J. & REYNES, F. 2011. Can reserve additions in mature crude oil provinces attenuate peak oil? *Energy,* 36**,** 5755-5764.

OSWALD, W. J. & GOLUEKE, C. G. 1960. Biological Transformation of Solar Energy. *In:* WAYNE, W. U. (ed.) *Advances in Applied Microbiology.* Academic Press.

PARK, M. O. 2005. New pathway for long-chain n-alkane synthesis via 1-alcohol in *Vibrio furnissii* M1. *Journal of Bacteriology,* 187**,** 1426-9.

PARK, Y.-I., CHOW, W. S., ANDERSON, J. M. & HURRY, V. M. 1996. Differential susceptibility of Photosystem II to light stress in light-acclimated pea leaves depends on the capacity for photochemical and non-radiative dissipation of light. *Plant Science,* 115**,** 137-149.

PARKIN, D. M. & BRAY, F. 2006. Chapter 2: The burden of HPV-related cancers. *Vaccine,* 24 Suppl 3**,** S3/11-25.

PARSELL, D. A., SILBER, K. R. & SAUER, R. T. 1990. Carboxy-terminal determinants of intracellular protein degradation. *Genes & Development,* 4**,** 277-86.

PARSONS, T. R. & TAKAHASHI, M. 1973. Environmental Control of Phytoplankton Cell Size. *Limnology & Oceanography,* 18**,** 511-515.

PASSWATER, R. & SOLOMON, N. 1997. Algae: The next generation of superfoods. *The Experts Optimal Health Journal,* 1**,** 2-10.

PATTERSON, B. K., LANDAY, A., ANDERSSON, J., BROWN, C., BEHBAHANI, H., JIYAMAPA, D., BURKI, Z., STANISLAWSKI, D., CZERNIEWSKI, M. A. & GARCIA, P. 1998. Repertoire of chemokine receptor expression in the female genital tract: implications for human immunodeficiency virus transmission. *The American Journal of Pathology,* 153**,** 481-90.

PAZ DE LA ROSA, G., MONROY-GARCIA, A., MORA-GARCIA MDE, L., PENA, C. G., HERNANDEZ-MONTES, J., WEISS-STEIDER, B. & GOMEZ-LIM, M. A. 2009. An HPV 16 L1-based chimeric human papilloma virus-like particles containing a string of epitopes produced in plants is able to elicit humoral and cytotoxic T-cell activity in mice. *Virology Journal,* 6**,** 2-2.

PECH-PACHECO, J. L., CRISTOBAL, G., ALVAREZ-BORREGO, J. & COHEN, L. 2001. Automatic system for phytoplanktonic algae identification. . *Limnetica,* 1**,** 143-158.

PEELER, T. C., STEPHENSON, M. B., EINSPAHR, K. J. & THOMPSON, G. A. 1989. Lipid Characterization of an Enriched Plasma Membrane Fraction of Dunaliella salina Grown in Media of Varying Salinity. *Plant Physiology,* 89**,** 970-6.

PERALES-VELA, H. V., PENA-CASTRO, J. M. & CANIZARES-VILLANUEVA, R. O. 2006. Heavy metal detoxification in eukaryotic microalgae. *Chemosphere,* 64**,** 1-10.

PETSCH, D. & ANSPACH, F. B. 2000. Endotoxin removal from protein solutions. *Journal of Biotechnology,* 76**,** 97-119.

PHILLIP, D., RUBAN, A. V., HORTON, P., ASATO, A. & YOUNG, A. J. 1996. Quenching of chlorophyll fluorescence in the major light-harvesting complex of photosystem II: a systematic study of the effect of carotenoid structure. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 1492-7.

PIECHULLA, B. & RIESSELMANN, S. 1990. Effect of Temperature Alterations on the Diurnal Expression Pattern of the Chlorophyll a/b Binding Proteins in Tomato Seedlings. *Plant Physiology,* 94**,** 1903-6.

POLLARD, M. R., ANDERSON, L., FAN, C., HAWKINS, D. J. & DAVIES, H. M. 1991. A specific acyl-ACP thioesterase implicated in medium chain fatty acid production in immature cotyledons of *Umbellularia californica*. *Archives of Biochemistry and Biophysics,* 284**,** 306-312.

POLLE, J. E., KANAKAGIRI, S. D. & MELIS, A. 2003. tla1, a DNA insertional transformant of the green alga *Chlamydomonas reinhardtii* with a truncated light-harvesting chlorophyll antenna size. *Planta,* 217**,** 49-59.

POLLE, J. E. W., BENEMANN, J. R., TANAKA, A. & MELIS, A. 2000. Photosynthetic solar conversion efficiency in wild type and truncated Chl antenna tranformant (Chl b-less mutant) of the green alga *Chlamydomonas reinhardtii*. *Proceedings of the 2000 Hydrogen Program Review***,** 20.

POLLE, J. E. W., KANAKAGIRI, S., JIN, E., MASUDA, T. & MELIS, A. 2002. Truncated chlorophyll antenna size of the photosystems - a practical method to improve microalgal productivity and hydrogen production in mass culture. *International Journal of Hydrogen Energy,* 27**,** 1257-1264.

POPE, M. & HAASE, A. T. 2003. Transmission, acute HIV-1 infection and the quest for strategies to prevent infection. *Nature Medicine,* 9**,** 847-52.

PORRA, R. J., THOMPSON, W. A. & KRIEDEMANN, P. E. 1989. Determination of accurate extinction coefficients and simultaneous-equations for assaying chlorophyll-a and chlorophyll-b extracted with 4 different solvents - verification of the concentration of chlorophyll standards by atomic-absorption spectroscopy. *Biochimica Et Biophysica Acta,* 975**,** 384-394.

POTTER, C. J. & LUO, L. 2010. Splinkerette PCR for mapping transposable elements in Drosophila. *PloS one,* 5**,** e10168.

POULSEN, N., CHESLEY, P. M. & KRÖGER, N. 2006. Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (*Bacillariophyceae*). *Journal of Phycology,* 42**,** 1059-1065.

POWELL, E. E., MAPIOUR, M. L., EVITTS, R. W. & HILL, G. A. 2009. Growth kinetics of *Chlorella vulgaris* and its use as a cathodic half cell. *Bioresource Technology,* 100**,** 269-74.

PURTON, S. 2006. Tools and Techniques for Chloroplast Transformation. *In:* LEON, R. & GAVAN (eds.) *Transgenic Microalgae as Green Cell Factories.* Landes Bioscience.

PURTON, S. & ROCHAIX, J. D. 1994. Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. *Plant Molecular Biology,* 24**,** 533-7.

RACE, H. L., HERRMANN, R. G. & MARTIN, W. 1999. Why have organelles retained genomes? *Trends in Genetics,* 15**,** 364-370.

RADAKOVITS, R., EDUAFO, P. M. & POSEWITZ, M. C. 2011. Genetic engineering of fatty acid chain length in *Phaeodactylum tricornutum*. *Metabolic Engineering,* 13**,** 89-95.

RADAKOVITS, R., JINKERSON, R. E., DARZINS, A. & POSEWITZ, M. C. 2010. Genetic engineering of algae for enhanced biofuel production. *Eukaryotic Cell,* 9**,** 486-501.

RAMEL, F., BIRTIC, S., CUINE, S., TRIANTAPHYLIDES, C., RAVANAT, J. L. & HAVAUX, M. 2012. Chemical quenching of singlet oxygen by carotenoids in plants. *Plant Physiology,* 158**,** 1267-78.

RAMIREZ, J. J. & CORBACHO, M. 2005. Population dynamics of Botryococcus braunii Kutzing 1849 in a shallow tropical eutrophic lake. *Archiv fuer Hydrobiologie Supplement,* 156**,** 129-143.

RANDOLPH-ANDERSON, B. L., SATO, R., JOHNSON, A. M., HARRIS, E. H., HAUSER, C. R., OEDA, K., ISHIGE, F., NISHIO, S., GILLHAM, N. W. & BOYNTON, J. E. 1998. Isolation and characterization of a mutant protoporphyrinogen oxidase gene from Chlamydomonas reinhardtii conferring resistance to porphyric herbicides. *Plant Molecular Biology,* 38**,** 839-859.

RAO, A. R., DAYANANDA, C., SARADA, R., SHAMALA, T. R. & RAVISHANKAR, G. A. 2007. Effect of salinity on growth of green alga Botryococcus braunii and its constituents. *Bioresource Technology,* 98**,** 560-564.

RAPLEY, R. 2000. *The Nucleic Acid Protocols Handbook,* Humana Press.

RAPPUOLI, R., PIZZA, M., DOUCE, G. & DOUGAN, G. 1999. Structure and mucosal adjuvanticity of cholera and *Escherichia coli* heat-labile enterotoxins. *Immunology Today,* 20**,** 493-500.

RASALA, B. A., MUTO, M., LEE, P. A., JAGER, M., CARDOSO, R. M. F., BEHNKE, C. A., KIRK, P., HOKANSON, C. A., CREA, R., MENDEZ, M. & MAYFIELD, S. P. 2010. Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of Chlamydomonas reinhardtii. *Plant Biotechnology Journal,* 8**,** 719-733.

RASALA, B. A., MUTO, M., SULLIVAN, J. & MAYFIELD, S. P. 2011. Improved heterologous protein expression in the chloroplast of Chlamydomonas reinhardtii through promoter and 5 ' untranslated region optimization. *Plant Biotechnology Journal,* 9**,** 674-683.

REDDING, K., MACMILLAN, F., LEIBL, W., BRETTEL, K., HANLEY, J., RUTHERFORD, A. W., BRETON, J. & ROCHAIX, J. D. 1998. A systematic survey of conserved histidines in the core subunits of Photosystem I by site-directed mutagenesis reveals the likely axial ligands of P700. *The EMBO Journal,* 17**,** 50-60.

REDDY, P. S., MAHANTY, S., KAUL, T., NAIR, S., SOPORY, S. K. & REDDY, M. K. 2008. A high-throughput genome-walking method and its use for cloning unknown flanking sequences. *Analytical Biochemistry,* 381**,** 248-253.

REES, D., NOCTOR, G., RUBAN, A. V., CROFTS, J., YOUNG, A. & HORTON, P. 1992. pH dependent chlorophyll fluorescence quenching in spinach thylakoids from light treated or dark adapted leaves. *Photosynthesis Research,* 31**,** 11-19.

REISER, S. & SOMERVILLE, C. 1997. Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase. *Journal of Bacteriology,* 179**,** 2969-75.

REMACLE, C., CARDOL, P., COOSEMANS, N., GAISNE, M. & BONNEFOY, N. 2006. High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 4771-4776.

REMACLE, C., CLINE, S., BOUTAFFALA, L., GABILLY, S., LAROSA, V., BARBIERI, M. R., COOSEMANS, N. & HAMEL, P. P. 2009. The ARG9 gene encodes the plastid-resident N-acetyl ornithine aminotransferase in the green alga *Chlamydomonas reinhardtii*. *Eukaryotic Cell,* 8**,** 1460-3.

REYNOSO, G. T. & DE GAMBOA, B. A. 1982. Salt tolerance in the freshwater algae Chlamydomon as reinhardii: Effect of proline and taurine. *Comparative Biochemistry and Physiology Part A: Physiology,* 73**,** 95-99.

RIESSELMANN, S. & PIECHULLA, B. 1990. Effect of dark phases and temperature on the chlorophyll a/b binding protein mRNA level oscillations in tomato seedlings. *Plant Molecular Biology,* 14**,** 605-16.

RISMANI-YAZDI, H., HAZNEDAROGLU, B. Z., BIBBY, K. & PECCIA, J. 2011. Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: pathway description and gene discovery for production of next-generation biofuels. *BMC Genomics,* 12**,** 148.

ROCHAIX, J., GOLDSCHMIDT-CLERMONT, M., MERCHANT, S. & TIMKO, M. 2004. Pigment Biosynthesis: Chlorophylls, Heme, and Carotenoids

The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas. *In:* ROCHAIX, J., GOLDSCHMIDT-CLERMONT, M. & MERCHANT, S. (eds.). Springer Netherlands.

ROCHAIX, J. D. 1980. Restriction fragments from *Chlamydomonas* chloroplast DNA. *Methods in Enzymology,* 65**,** 785-95.

ROCHAIX, J. D. 1996. Post-transcriptional regulation of chloroplast gene expression in *Chlamydomonas reinhardtii*. *Plant Molecular Biology,* 32**,** 327-41.

ROCHAIX, J. D. 2001. Posttranscriptional control of chloroplast gene expression. From RNA to photosynthetic complex. *Plant Physiology,* 125**,** 142-4.

ROCHAIX, J. D. & VAN DILLEWIJN, J. 1982. Transformation of the green alga *Chlamydomonas reinhardii* with yeast DNA. *Nature,* 296**,** 70-72.

RODGERS, M. R., FLANIGAN, D. J. & JAKUBOWSKI, W. 1995. Identification of algae which interfere with the detection of Giardia cysts and Cryptosporidium oocysts and a method for alleviating this interference. *Applied and Environmental Microbiology,* 61**,** 3759-63.

RODOLFI, L., CHINI ZITTELLI, G., BASSI, N., PADOVANI, G., BIONDI, N., BONINI, G. & TREDICI, M. R. 2009. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering,* 102**,** 100-12.

ROESLER, K., SHINTANI, D., SAVAGE, L., BODDUPALLI, S. & OHLROGGE, J. 1997. Targeting of the Arabidopsis homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. *Plant Physiology,* 113**,** 75-81.

ROESSLER, P. G., BLEIBAUM, J. L., THOMPSON, G. A. & OHLROGGE, J. B. 1994. Characteristics of the gene that encodes acetyl-CoA carboxylase in the diatom Cyclotella cryptica. *Annals of the New York Academy of Sciences,* 721**,** 250-6.

ROJO, F. 2005. Specificity at the end of the tunnel: understanding substrate length discrimination by the AlkB alkane hydroxylase. *Journal of Bacteriology,* 187**,** 19-22.

ROSALES-MENDOZA, S., PAZ-MALDONADO, L. & SORIA-GUERRA, R. 2012. *Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: current status and perspectives. *Plant Cell Reports,* 31**,** 479-494.

ROSENBERG, J., OYLER, J., WILKINSON, L. & BETENBAUGH, M. 2008. A green light for engineered algae: redirecting metabolism to fuel a biotechnology revolution. *Current Opinion in Biotechnology,* 19**,** 430-6.

RUBAN, A. V., PHILLIP, D., YOUNG, A. J. & HORTON, P. 1997. Carotenoid-dependent oligomerization of the major chlorophyll a/b light harvesting complex of photosystem II of plants. *Biochemistry,* 36**,** 7855-9.

RUBAN, A. V., WALTERS, R. G. & HORTON, P. 1992. The molecular mechanism of the control of excitation energy dissipation in chloroplast membranes. Inhibition of delta pH-dependent quenching of chlorophyll fluorescence by dicyclohexylcarbodiimide. *FEBS Letters,* 309**,** 175-9.

RUDOLF, M. P., FAUSCH, S. C., DA SILVA, D. M. & KAST, W. M. 2001. Human dendritic cells are activated by chimeric human papillomavirus type-16 virus-like particles and induce epitope-specific human T cell responses in vitro. *Journal of Immunology,* 166**,** 5917-24.

RUHLMAN, T., VERMA, D., SAMSON, N. & DANIELL, H. 2010. The Role of Heterologous Chloroplast Sequence Elements in Transgene Integration and Expression. *Plant Physiology,* 152**,** 2088-2104.

SAMBROOK, J. & RUSSELL, D. W. 2001. *Molecular cloning. A laboratory manual.* , New York, Cold Spring Harbor Laboratory Press.

SAMEJIMA, H. & MYERS, J. 1958. On the heterotrophic growth of *Chlorella pyrenoidosa*. *Journal of General Microbiology,* 18**,** 107-17.

SANFORD, J. C., SMITH, F. D. & RUSSELL, J. A. 1993. Optimizing the biolistic process for different biological applications. *Methods in Enzymology,* 217**,** 483-509.
SANSAWA, H., TAKAHASHI, M., TSUCHIKURA, S. & ENDO, H. 2006. Effect of *Chlorella* and its fractions on blood pressure, cerebral stroke lesions, and life-span in stroke-prone spontaneously hypertensive rats. *Journal of Nutritional Science and Vitaminology (Tokyo),* 52**,** 457-66.
SANTABARBARA, S., REIFSCHNEIDER, K., JASAITIS, A., GU, F., AGOSTINI, G., CARBONERA, D., RAPPAPORT, F. & REDDING, K. E. 2010. Interquinone electron transfer in photosystem I as evidenced by altering the hydrogen bond strength to the phylloquinone(s). *Journal of Physical Chemistry. B,* 114**,** 9300-12.
SAPP, M., FLIGGE, C., PETZAK, I., HARRIS, J. R. & STREECK, R. E. 1998. Papillomavirus Assembly Requires Trimerization of the Major Capsid Protein by Disulfides between Two Highly Conserved Cysteines. *Journal of Virology,* 72**,** 6186-6189.
SASIDHARAN, A. & GNANAM, A. 1990. Immunological identification and quantification of light harvesting chlorophyll a/b protein of Chlorella protothecoides. *Indian Journal of Biochemistry & Biophysics,* 27**,** 1-4.
SCAIFE, M. A., BURJA, A. M. & WRIGHT, P. C. 2009. Characterization of Cyanobacterial beta-Carotene Ketolase and Hydroxylase Genes in Escherichia coli, and Their Application for Astaxanthin Biosynthesis. *Biotechnology and Bioengineering,* 103**,** 944-955.
SCHADLICH, L., SENGER, T., GERLACH, B., MUCKE, N., KLEIN, C., BRAVO, I. G., MULLER, M. & GISSMANN, L. 2009a. Analysis of Modified Human Papillomavirus Type 16 L1 Capsomeres: the Ability To Assemble into Larger Particles Correlates with Higher Immunogenicity. *Journal of Virology,* 83**,** 7690-7705.
SCHADLICH, L., SENGER, T., KIRSCHNING, C. J., MULLER, M. & GISSMANN, L. 2009b. Refining HPV 16 L1 purification from *E. coli*: reducing endotoxin contaminations and their impact on immunogenicity. *Vaccine,* 27**,** 1511-22.
SCHAEFER, D. G. 2002. A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Annual Review of Plant Biology,* 53**,** 477-501.
SCHENK, P., THOMAS-HALL, S., STEPHENS, E., MARX, U., MUSSGNUG, J., POSTEN, C., KRUSE, O. & HANKAMER, B. 2008. Second Generation Biofuels: High-Efficiency Microalgae for Biodiesel Production. *BioEnergy Research,* 1**,** 20-43.
SCHIEDLMEIER, B., SCHMITT, R., MULLER, W., KIRK, M. M., GRUBER, H., MAGES, W. & KIRK, D. L. 1994. Nuclear transformation of Volvox carteri *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 5080-5084.
SCHIRMER, A., RUDE, M. A., LI, X., POPOVA, E. & DEL CARDAYRE, S. B. 2010. Microbial biosynthesis of alkanes. *Science,* 329**,** 559-62.
SCHMIDT, G. W., MATLIN, K. S. & CHUA, N. H. 1977. A rapid procedure for selective enrichment of photosynthetic electron transport mutants. *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 610-4.
SCOTT, S. A., DAVEY, M. P., DENNIS, J. S., HORST, I., HOWE, C. J., LEA-SMITH, D. J. & SMITH, A. G. 2010. Biodiesel from algae: challenges and prospects. *Current Opinion in Biotechnology,* 21**,** 277-86.
SEARCHINGER, T., HEIMLICH, R., HOUGHTON, R. A., DONG, F., ELOBEID, A., FABIOSA, J., TOKGOZ, S., HAYES, D. & YU, T. H. 2008. Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science,* 319**,** 1238-40.
SELLWOOD, C., SLABAS, A. R. & RAWSTHORNE, S. 2000. Effects of manipulating expression of acetyl-CoA carboxylase I in *Brassica napus* L. embryos. *Biochemical Society Transactions,* 28**,** 598-600.
SERINO, G. & MALIGA, P. 1997. A negative selection scheme based on the expression of cytosine deaminase in plastids. *The Plant Journal,* 12**,** 697-701.
SEXTON, A., DRAKE, P. M., MAHMOOD, N., HARMAN, S. J., SHATTOCK, R. J. & MA, J. K. C. 2006. Transgenic plant production of Cyanovirin-N, an HIV microbicide. *FASEB Journal,* 20**,** 356-358.
SEXTON, A., HARMAN, S., SHATTOCK, R. J. & MA, J. K. C. 2009. Design, expression, and characterization of a multivalent, combination HIV microbicide. *FASEB Journal,* 23**,** 3590-3600.
SHEEHAN, J. 1998. *A look back at the US Department of Energy's Aquatic Species Program: Biodiesel from algae*.
SHENOY, S. R., O'KEEFE, B. R., BOLMSTEDT, A. J., CARTNER, L. K. & BOYD, M. R. 2001. Selective interactions of the human immunodeficiency virus-inactivating protein cyanovirin-N with high-mannose oligosaccharides on gp120 and other glycoproteins. *The Journal of Pharmacology and Experimental Therapeutics,* 297**,** 704-10.
SHI, X.-M., ZHANG, X.-W. & CHEN, F. 2000. Heterotrophic production of biomass and lutein by *Chlorella protothecoides* on various nitrogen sources. *Enzyme and Microbial Technology,* 27**,** 312-318.
SHI, X. L., LEPERE, C., SCANLAN, D. J. & VAULOT, D. 2011. Plastid 16S rRNA Gene Diversity among Eukaryotic Picophytoplankton Sorted by Flow Cytometry from the South Pacific Ocean. *PloS one,* 6.
SHIBATA, S. & HAYAKAWA, K. 2009. Bioavailability of Lutein in Chlorella Powder: A Single Ingestion of Chlorella Powder Raises Serum Lutein Concentrations in Healthy Human Volunteers. *Food Science and Technology Research,* 15**,** 449-452.
SHIMIZU, Y. 1996. Microalgal metabolites: a new perspective. *Annual Review of Microbiology,* 50**,** 431-65.
SHIMOGAWARA, K., FUJIWARA, S., GROSSMAN, A. & USUDA, H. 1998. High-efficiency transformation of Chlamydomonas reinhardtii by electroporation. *Genetics,* 148**,** 1821-1828.
SHINE, J. & DALGARNO, L. 1975. DETERMINANT OF CISTRON SPECIFICITY IN BACTERIAL RIBOSOMES. *Nature,* 254**,** 34-38.
SIAUT, M., CUINE, S., CAGNON, C., FESSLER, B., NGUYEN, M., CARRIER, P., BEYLY, A., BEISSON, F., TRIANTAPHYLIDES, C., LI-BEISSON, Y. & PELTIER, G. 2011. Oil accumulation in the model green

alga *Chlamydomonas reinhardtii*: characterization, variability between common laboratory strains and relationship with starch reserves. *BMC Biotechnology,* 11**,** 7.
SIGMUND, C. D., ETTAYEBI, M. & MORGAN, E. A. 1984. Antibiotic resistance mutations in 16S and 23S ribosomal RNA genes of Escherichia coli. *Nucleic Acids Research,* 12**,** 4653-63.
SILFLOW, C. D. & YOUNGBLOM, J. 1986. *Chlamydomonas reinhardtii* tubulin gene structure. *Annals of the New York Academy of Sciences,* 466**,** 18-30.
SIM, S. J., AN, J. Y. & KIM, B. W. 2001. Two-phase extraction culture of *Botryococcus braunii* producing long-chain unsaturated hydrocarbons. *Biotechnology Letters,* 23**,** 201-205.
SINESHCHEKOV, O. A., JUNG, K. H. & SPUDICH, J. L. 2002. Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 8689-8694.
SITTE, P. 1993. Symbiogenetic evolution of complex cells and complex plastids. *European Journal of Protistology,* 29**,** 131-143.
SIZOVA, I. A., LAPINA, T. V., FROLOVA, O. N., ALEXANDROVA, N. N., AKOPIANTS, K. E. & DANILENKO, V. N. 1996. Stable nuclear transformation of Chlamydomonas reinhardtii with a Streptomyces rimosus gene as the selective marker. *Gene,* 181**,** 13-18.
SOMANCHI, A., BARNES, D. & MAYFIELD, S. P. 2005. A nuclear gene of *Chlamydomonas reinhardtii*, *Tba1,* encodes a putative oxidoreductase required for translation of the chloroplast *psbA* mRNA. *The Plant Journal,* 42**,** 341-52.
SOMERVILLE, C. R. 1986. Analysis of Photosynthesis with Mutants of Higher Plants and Algae. *Annual Review of Plant Physiology,* 37**,** 467-506.
SORKHOH, N., AL-HASAN, R., RADWAN, S. & HOPNER, T. 1992. Self-cleaning of the Gulf. *Nature,* 359**,** 109-109.
SOROKIN, C. 1967. New High-Temperature *Chlorella*. *Science,* 158**,** 1204-&-1204-&.
SOROKIN, C. 1971. Calefaction and Phytoplankton. *BioScience,* 21**,** 1153-1159.
SOROKIN, C. & KRAUSS, R. W. 1958. The Effects of Light Intensity on the Growth Rates of Green Algae. *Plant Physiology,* 33**,** 109-13.
SOROKIN, C. & KRAUSS, R. W. 1959. Maximum Growth Rates of *Chlorella* in Steady-State and in Synchronized Cultures. *Proceedings of the National Academy of Sciences of the United States of America,* 45**,** 1740-4.
SOROKIN, C. & KRAUSS, R. W. 1962. Effects of Temperature & Illuminance on *Chlorella* Growth Uncoupled From Cell Division. *Plant Physiology,* 37**,** 37-42.
SOROKIN, C. & MYERS, J. 1953a. A high-temperature strain of Chlorella. *Science,* 117**,** 330-1.
SOROKIN, C. & MYERS, J. 1953b. A High-Temperature Strain of *Chlorella*. *Science,* 117**,** 330-331.
SPECHT, E., MIYAKE-STONER, S. & MAYFIELD, S. 2010. Micro-algae come of age as a platform for recombinant protein production. *Biotechnology Letters,* 32**,** 1373-83.
SPOEHR, H. A. & MILNER, H. W. 1949. The Chemical Composition of *Chlorella*; Effect of Environmental Conditions. *Plant Physiology,* 24**,** 120-49.
SPOLAORE, P., JOANNIS-CASSAN, C., DURAN, E. & ISAMBERT, A. 2006. Commercial applications of microalgae. *Journal of Bioscience and Bioengineering,* 101**,** 87-96.
SPREITZER, R. J. & METS, L. 1981. Photosynthesis-deficient Mutants of *Chlamydomonas* reinhardii with Associated Light-sensitive Phenotypes. *Plant Physiology,* 67**,** 565-9.
SRIRANGAN, K., PYNE, M. E. & PERRY CHOU, C. 2011. Biochemical and genetic engineering strategies to enhance hydrogen production in photosynthetic algae and cyanobacteria. *Bioresource Technology,* 102**,** 8589-8604.
STAMPACCHIA, O., GIRARD-BASCOU, J., ZANASCO, J. L., ZERGES, W., BENNOUN, P. & ROCHAIX, J. D. 1997. A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in *Chlamydomonas*. *The Plant Cell,* 9**,** 773-82.
STARR, R. C. & ZEIKUS, J. A. 1993. UTEX— the culture collection of algae at the University of Texas at Austin 1993, List of cultures. . *Journal of Phycology,* 29**,** 1-106.
STEINBRENNER, J. & SANDMANN, G. 2006. Transformation of the Green Alga *Haematococcus pluvialis* with a Phytoene Desaturase for Accelerated Astaxanthin Biosynthesis. *Applied and Environmental Microbiology,* 72**,** 7477-7484.
STENGEL, D. B., CONNAN, S. & POPPER, Z. A. 2011. Algal chemodiversity and bioactivity: sources of natural variability and implications for commercial application. *Biotechnology advances,* 29**,** 483-501.
STEPHENSON, P. G., MOORE, C. M., TERRY, M. J., ZUBKOV, M. V. & BIBBY, T. S. 2011. Improving photosynthesis for algal biofuels: toward a green revolution. *Trends in Biotechnology,* 29**,** 615-623.
STEVENS, D. R., ATTEIA, A., FRANZ√©N, L. G. & PURTON, S. 2001. Cycloheximide resistance conferred by novel mutations in ribosomal protein L41 of Chlamydomonas reinhardtii. *Molecular & General Genetics: MGG,* 264**,** 790-795.
STOLZ, P. & OBERMAYER, B. 2005. *Manufacturing microalgae for skin care,* Carol Stream, IL, ETATS-UNIS, Allured.
SUKENIK, A., BENNETT, J. & FALKOWSKI, P. 1987. Light-saturated photosynthesis - Limitation by electron transport or carbon fixation? *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 891**,** 205-215.
SUN, M., QIAN, K., SU, N., CHANG, H., LIU, J. & SHEN, G. 2003. Foot-and-mouth disease virus VP1 protein fused with cholera toxin B subunit expressed in *Chlamydomonas reinhardtii* chloroplast. *Biotechnology Letters,* 25**,** 1087-92.
SURZYCKI, R., COURNAC, L., PELTIER, G. & ROCHAIX, J. D. 2007. Potential for hydrogen production with inducible chloroplast gene expression in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 17548-53.

SURZYCKI, R., GREENHAM, K., KITAYAMA, K., DIBAL, F., WAGNER, R., ROCHAIX, J. D., AJAM, T. & SURZYCKI, S. 2009. Factors effecting expression of vaccines in microalgae. *Biologicals,* 37**,** 133-8.

SUSHCHIK, N. N., KALACHEVA, G. S., ZHILA, N. O., GLADYSHEV, M. I. & VOLOVA, T. G. 2003. A temperature dependence of the intra- and extracellular fatty-acid composition of green algae and cyanobacterium. *Russian Journal of Plant Physiology,* 50**,** 374-380.

SVAB, Z., HAJDUKIEWICZ, P. & MALIGA, P. 1990. Stable transformation of plastids in higher plants. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 8526-30.

SVAB, Z. & MALIGA, P. 1993. High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 913-7.

SVAROVSKY, S., BOROVKOV, A. & SYKES, K. 2008. Cationic gold microparticles for biolistic delivery of nucleic acids. *Biotechniques,* 45**,** 535-40.

SYRETT, P. J. 1951. The effect of cyanide on the respiration and the oxidative assimilation of glucose by Chlorella vulgaris *Annals of Botany,* 15**,** 473-492.

SZAREWSKI, A. 2010. HPV vaccine: Cervarix. *Expert Opin Biol Ther,* 3**,** 477-87.

TAKEDA, H. 1993. Taxonomical assignment of chlorococal algae from their cell wall composition. *Phytochemistry,* 34**,** 1053-1055.

TAKEDA, H. & HIROKAWA, T. 1979. Studies on the cell wall of Chlorella II. Mode of increase of glucosamine in the cell wall during the synchronous growth of Chlorella ellipsoidea. *Plant and Cell Physiology,* 20**,** 3.

TANG, D. K., QIAO, S. Y. & WU, M. 1995. Insertion mutagenesis of *Chlamydomonas reinhardtii* by electroporation and heterologous DNA. *Biochemistry and Molecular Biology International,* 36**,** 1025-35.

TE, M. R., LOHUIS & MILLER, D. J. 1998. Genetic transformation of dinoflagellates (*Amphidinium* and *Symbiodinium*): expression of GUS in microalgae using heterologous promoter constructs. *The Plant Journal,* 13**,** 427-435.

TENG, C., QIN, S., LIU, J., YU, D., LIANG, C. & TSENG, C. 2002. Transient expression of *lacZ* in bombarded unicellular green alga *Haematococcus pluvialis*. *Journal of Applied Phycology,* 14**,** 497-500.

THONES, N., HERREINER, A., SCHADLICH, L., PIUKO, K. & MULLER, M. 2008. A direct comparison of human papillomavirus type 16 L1 particles reveals a lower immunogenicity of capsomeres than viruslike particles with respect to the induced antibody response. *Journal of Virology,* 82**,** 5472-85.

TIWARI, V., SHUKLA, S. Y. & SHUKLA, D. 2009. A sugar binding protein cyanovirin-N blocks herpes simplex virus type-1 entry and cell fusion. *Antiviral Research,* 84**,** 67-75.

TOBIAS, J. W., SHRADER, T. E., ROCAP, G. & VARSHAVSKY, A. 1991. The N-end rule in bacteria. *Science,* 254**,** 1374-7.

TRAN, M., ZHOU, B., PETTERSSON, P. L., GONZALEZ, M. J. & MAYFIELD, S. P. 2009. Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnology and Bioengineering,* 104**,** 663-73.

TREMOLIERES, A., DUBERTRET, G., PAULSEN, H., HOBE, S. & DEMANDRE, C. 1995. *In vivo manipulation of lipid composition in mutants of Chlamydomonas reinhardtii: A tool to study the roles of lipids in the biogenesis of the photosynthetic membrane*.

UNAIDS. 2008. *UNAIDS Report on the Global AIDS Epidemic 2010* [Online].

VALDERRAMA, L. T., DEL CAMPO, C. M., RODRIGUEZ, C. M., DE- BASHAN, L. E. & BASHAN, Y. 2002. Treatment of recalcitrant wastewater from ethanol and citric acid production using the microalga *Chlorella vulgaris* and the macrophyte *Lemna minuscula*. *Water Research,* 36**,** 4185-92.

VARFOLOMEEV, S. D. & WASSERMAN, L. A. 2011. Microalgae as Source of Biofuel, Food, Fodder, and Medicines. *Applied Biochemistry and Microbiology,* 47**,** 789-807.

VARSANI, A., WILLIAMSON, A. L., ROSE, R. C., JAFFER, M. & RYBICKI, E. P. 2003. Expression of Human papillomavirus type 16 major capsid protein in transgenic Nicotiana tabacum cv. Xanthi. *Archives of Virology,* 148**,** 1771-86.

VARSANO, T., WOLF, S. G. & PICK, U. 2006. A chlorophyll a/b-binding protein homolog that is induced by iron deficiency is associated with enlarged photosystem I units in the eucaryotic alga *Dunaliella salina*. *The Journal of Biological Chemistry,* 281**,** 10305-15.

VEAZEY, R. S., MARX, P. A. & LACKNER, A. A. 2003. Vaginal CD4+ T cells express high levels of CCR5 and are rapidly depleted in simian immunodeficiency virus infection. *Journal of Infectious Diseases,* 187**,** 769-76.

VENKATARAMAN, L. V. 1997. Spirulina platensis (Arthrospira): Physiology, Cell Biology and Biotechnologym, edited by Avigad Vonshak. *Journal of Applied Phycology,* 9**,** 295-296.

VERMA, D. & DANIELL, H. 2007. Chloroplast vector systems for biotechnology applications. *Plant Physiology,* 145**,** 1129-43.

VOELKER, T. A. & DAVIES, H. M. 1994. Alteration of the specificity and regulation of fatty-acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. *Journal of Bacteriology,* 176**,** 7320-7327.

VOELKER, T. A., WORRELL, A. C., ANDERSON, L., BLEIBAUM, J., FAN, C., HAWKINS, D. J., RADKE, S. E. & DAVIES, H. M. 1992. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. *Science,* 257**,** 72-74.

VOIGT, J., HINKELMANN, B. & HARRIS, E. H. 1997. Production of cell wall polypeptides by different cell wall mutants of the unicellular green alga Chlamydomonas reinhardtii. *Microbiological Research,* 152**,** 189-198.

VOLKER HUSS, C. F., ELKE C. HARTMANN, MONIKA HIRMER, ANNETTE KLOBOUCEK, BARBARA M. SEIDEL, PETRA WENZELER, ERICH KESSLER 1999. BIOCHEMICAL TAXONOMY AND MOLECULAR PHYLOGENY OF THE GENUS CHLORELLA SENSU LATO (CHLOROPHYTA). *Journal of Phycology,* 35**,** 587-598.

VOLOVA, T. G., KALACHEVA, G. S., ZHILO, N. O. & PLOTNIKOV, V. F. 1998. Physiological and biochemical properties of the alga Botryococcus braunii. *Russian Journal of Plant Physiology,* 45**,** 775-779.

WAGENER, K. 1981. Mariculture on land,ÄîA system for biofuel farming in coastal deserts. *Biomass,* 1**,** 145-158.

WAGENER, K. 1983. Mass cultures of marine algae for energy farming in coastal deserts. *International Journal of Biometeorology,* 27**,** 227-233.

WAHEED, M. T., THONES, N., MULLER, M., HASSAN, S. W., GOTTSCHAMEL, J., LOSSL, E., KAUL, H. P. & LOSSL, A. G. 2011a. Plastid expression of a double-pentameric vaccine candidate containing human papillomavirus-16 L1 antigen fused with LTB as adjuvant: transplastomic plants show pleiotropic phenotypes. *Plant Biotechnology Journal,* 9**,** 651-660.

WAHEED, M. T., THONES, N., MULLER, M., HASSAN, S. W., RAZAVI, N. M., LOSSL, E., KAUL, H. P. & LOSSL, A. G. 2011b. Transplastomic expression of a modified human papillomavirus L1 protein leading to the assembly of capsomeres in tobacco: a step towards cost-effective second-generation vaccines. *Transgenic Research*.

WAKASUGI, T., NAGAI, T., KAPOOR, M., SUGITA, M., ITO, M., ITO, S., TSUDZUKI, J., NAKASHIMA, K., TSUDZUKI, T., SUZUKI, Y., HAMADA, A., OHTA, T., INAMURA, A., YOSHINAGA, K. & SUGIURA, M. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 5967-72.

WALKER, T. L., BECKER, D. K. & COLLET, C. 2005a. Characterisation of the *Dunaliella tertiolecta RbcS* genes and their promoter activity in *Chlamydomonas reinhardtii*. *Plant Cell Reports,* 23**,** 727-35.

WALKER, T. L., COLLET, C. & PURTON, S. 2005b. Algal transgenics in the genomic era *Journal of Phycology,* 41**,** 1077-1093.

WALKER, T. L., PURTON, S., BECKER, D. K. & COLLET, C. 2005c. Microalgae as bioreactors. *Plant Cell Reports,* 24**,** 629-41.

WALSH, G. 2010. Biopharmaceutical benchmarks 2010. *Nature Biotechnology,* 28**,** 917-924.

WALTERS, R. G. & HORTON, P. 1991. Resolution of components of non-photochemical chlorophyll fluorescence quenching in barley leaves. *Photosynthesis Research,* 27**,** 121-133.

WANG, C., WANG, Y., SU, Q. & GAO, W. 2007. Transient expression of the GUS gene in a unicellular marine green alga, *Chlorella sp.* MACC/C95, via electroporation. . *Biotechnology and Bioprocess Engineering,* 12**,** 180-183.

WANG, Q., GRIFFIN, H., SOUTHERN, S., JACKSON, D., MARTIN, A., MCINTOSH, P., DAVY, C., MASTERSON, P. J., WALKER, P. A., LASKEY, P., OMARY, M. B. & DOORBAR, J. 2004. Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *Journal of Virology,* 78**,** 821-33.

WANG, X., BRANDSMA, M., TREMBLAY, R., MAXWELL, D., JEVNIKAR, A. M., HUNER, N. & MA, S. 2008. A novel expression platform for the production of diabetes-associated autoantigen human glutamic acid decarboxylase (hGAD65). *BMC Biotechnol,* 8**,** 87.

WANG, Z. T., ULLRICH, N., JOO, S., WAFFENSCHMIDT, S. & GOODENOUGH, U. 2009. Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*. *Eukaryotic Cell,* 8**,** 1856-68.

WARZECHA, H., MASON, H. S., LANE, C., TRYGGVESSON, A., RYBICKI, E., WILLIAMSON, A.-L., CLEMENTS, J. D. & ROSE, R. C. 2003. Oral Immunogenicity of Human Papillomavirus-Like Particles Expressed in Potato. *Journal of Virology,* 77**,** 8702-8711.

WIDJAJA, A., CHIEN, C.-C. & JU, Y.-H. 2009. Study of increasing lipid production from fresh water microalgae Chlorella vulgaris. *Journal of the Taiwan Institute of Chemical Engineers,* 40**,** 13-20.

WILHELM, C. & JAKOB, T. 2011. From photons to biomass and biofuels: evaluation of different strategies for the improvement of algal biotechnology based on comparative energy balances. *Applied Microbiology and Biotechnology,* 92**,** 909-919.

WILLIAMS, P. J. 2007. Biofuel: microalgae cut the social and ecological costs. *Nature,* 450**,** 478-478.

WILSKI, S., JOHANNINGMEIER, U., HERTEL, S. & OETTMEIER, W. 2006. Herbicide binding in various mutants of the photosystem II D1 protein of *Chlamydomonas reinhardtii*. *Pesticide Biochemistry and Physiology,* 84**,** 157-164.

WOERTZ, I., FEFFER, A., LUNDQUIST, T. & NELSON, Y. 2009. Algae Grown on Dairy and Municipal Wastewater for Simultaneous Nutrient Removal and Lipid Production for Biofuel Feedstock. *Journal of Environmental Engineering,* 135**,** 1115-1122.

WOESSNER, J. P. & GOODENOUGH, U. W. 1992. Zygote and vegetative cell wall proteins in Chlamydomonas reinhardtii share a common epitope, (SerPro)X. *Plant Science,* 83**,** 65-76.

WORK, V. H., D'ADAMO, S., RADAKOVITS, R., JINKERSON, R. E. & POSEWITZ, M. C. 2011. Improving photosynthesis and metabolic networks for the competitive production of phototroph-derived biofuels. *Current Opinion in Biotechnology*.

WU, H. L., HSEU, R. S. & LIN, L. P. 2001. Identification of *Chlorella spp.* isolates using ribosomal DNA sequences. *Botanical Bulletin of Academia Sinica,* 42**,** 115-121.

WU, S., XU, L., HUANG, R. & WANG, Q. 2011a. Improved biohydrogen production with an expression of codon-optimized *hemH* and *lba* genes in the chloroplast of *Chlamydomonas reinhardtii*. *Bioresource Technology,* 102**,** 2610-6.

WU, S. X., XU, L. L., HUANG, R. & WANG, Q. X. 2011b. Improved biohydrogen production with an expression of codon-optimized hemH and lba genes in the chloroplast of *Chlamydomonas reinhardtii*. *Bioresource Technology,* 102**,** 2610-2616.

WU-SCHARF, D., JEONG, B., ZHANG, C. & CERUTTI, H. 2000. Transgene and transposon silencing in Chlamydomonas reinhardtii by a DEAH-box RNA helicase. *Science,* 290**,** 1159-62.
WURBS, D., RUF, S. & BOCK, R. 2007. Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. *The Plant Journal,* 49**,** 276-288.
WURTZ, E. A., SEARS, B. B., RABERT, D. K., SHEPHERD, H. S., GILLHAM, N. W. & BOYNTON, J. E. 1979. Specific increase in chloroplast gene-mutations following growth of *Chlamydomonas* in 5-fluorodeoxyuridine. *Molecular & General Genetics,* 170**,** 235-242.
YAGISAWA, F., NISHIDA, K., OKANO, Y., MINODA, A., TANAKA, K. & KUROIWA, T. 2004. Isolation of cycloheximide-resistant mutants of *Cyanidioschyzon merolae*. *Cytologia,* 69**,** 97-100.
YAMAGUCHI, K. 1996. Recent advances in microalgal bioscience in Japan, with special reference to utilization of biomass and metabolites: a review. *Journal of Applied Phycology,* 8**,** 487-502.
YAMASAKI, Y., YOKOSE, T., NISHIKAWA, T., KIM, D., JIANG, Z., YAMAGUCHI, K. & ODA, T. 2012. Effects of alginate oligosaccharide mixtures on the growth and fatty acid composition of the green alga *Chlamydomonas reinhardtii*. *Journal of Bioscience and Bioengineering,* 113**,** 112-116.
YAN, Y. & LIAO, J. C. 2009. Engineering metabolic systems for production of advanced fuels. *Journal of Industrial Microbiology & Biotechnology,* 36**,** 471-9.
YANG, Z., LI, Y., FENG, C., DONG, L. I., ZHIFANG, Z., YANXIN, L. I. U., DEXIAN, Z., YONG, W. & GUIFANG, S. 2006a. Expression of human soluble TRAIL in Chlamydomonas reinhardtii chloroplast. *Chinese Science Bulletin,* 51**,** 1703-1709.
YANG, Z. Q., LI, Y. N., CHEN, F., LI, D., ZHANG, Z. F., LIU, Y. X., ZHENG, D. X., WANG, Y. & SHEN, G. F. 2006b. Expression of human soluble TRAIL in *Chlamydomonas reinhardtii* chloroplast. *Chinese Science Bulletin,* 51**,** 1703-1709.
YE, G. N., COLBURN, S. M., XU, C. W., HAJDUKIEWICZ, P. T. & STAUB, J. M. 2003. Persistence of unselected transgenic DNA during a plastid transformation and segregation approach to herbicide resistance. *Plant Physiology,* 133**,** 402-10.
YU, H., LIU, Z.-T., LV, R. & ZHANG, W.-Q. 2010. Antiviral Activity of Recombinant Cyanovirin-N against HSV-1. *Virologica Sinica,* 25**,** 432-439.
ZAJAC, G. 2008. Influence of FAME addition to diesel fuel on exhaust fumes opacity of diesel engine. *International Agrophysics,* 22**,** 179-183.
ZAMALEEVA, A. I., SHARIPOVA, I. R., SHAMAGSUMOVA, R. V., IVANOV, A. N., EVTUGYN, G. A., ISHMUCHAMETOVA, D. G. & FAKHRULLIN, R. F. 2011. A whole-cell amperometric herbicide biosensor based on magnetically functionalised microalgae and screen-printed electrodes. *Analytical Methods,* 3**,** 509-513.
ZASLAVSKAIA, L. A., LIPPMEIER, J. C., KROTH, P. G., GROSSMAN, A. R. & APT, K. E. 2000. Transformation of the diatom *Phaeodactylum tricornutum* (*Bacillariophyceae*) with a variety of selectable marker and reporter genes. *Journal of Phycology,* 36**,** 379-386.
ZERGES, W. 2000. Translation in chloroplasts. *Biochimie,* 82**,** 583-601.
ZERGES, W., GIRARD-BASCOU, J. & ROCHAIX, J. D. 1997. Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in *Chlamydomonas reinhardtii*. *Molecular and Cellular Biology,* 17**,** 3440-8.
ZHANG, H., WANG, W., LI, Y., YANG, W. & SHEN, G. 2011a. Mixotrophic cultivation of Botryococcus braunii. *Biomass and Bioenergy,* 35**,** 1710-1715.
ZHANG, S. 2007. *Identification of Novel "yellow-in-the-dark" Mutants in Chlamydomonas reinhardtii*, University of Virginia.
ZHANG, X., LI, M., AGRAWAL, A. & SAN, K.-Y. 2011b. Efficient free fatty acid production in *Escherichia coli* using plant acyl-ACP thioesterases. *Metabolic Engineering,* 13**,** 713-722.
ZHANG, Y. & FROHMAN, M. A. 1997. Using rapid amplification of cDNA ends (RACE) to obtain full-length cDNAs. *Methods in Molecular Biology,* 69**,** 61-87.
ZHANG, Y. & FROHMAN, M. A. 1998. Cloning cDNA Ends Using RACE. *Methods in Molecular Medicine,* 13**,** 81-105.
ZHANG, Y. K., SHEN, G. F. & RU, B. G. 2006. Survival of human metallothionein-2 transplastomic *Chlamydomonas reinhardtii* to ultraviolet B exposure. *Acta Biochimica et Biophysica Sinica (Shanghai),* 38**,** 187-93.
ZHENG, Y., CHI, Z., LUCKER, B. & CHEN, S. 2012. Two-stage heterotrophic and phototrophic culture strategy for algal biomass and lipid production. *Bioresource Technology,* 103**,** 484-488.
ZITTELLI, G. C. 2004. Industrial production of microalgal cell mass and secondary products - species of high potential: mass cultivation of Nannochloropsis in closed systems. *Handbook of Microalgal Culture***,** 298-303.

APPENDICES

Appendix A

List of oligonucleotides used for amplification and sequencing of the *psbB-petB* insert

Name	Sequence
psbB.F	TTWCCAATRTARTCRTAAAAAGC
petB.R	AATTTCTTTTGGATCTWCWCC
RevM13 (standard)	AACAGCTATGACCATG
FwII.Cs.BPB	TGCCTAAACGTGTCATGAAAGG
FwIII.Cs.BPB	AGTATTCAACATGTAAATATGCC
FwIV.Cs.BPB	GAAAATACTCGCTCCTTTCTCC
FwM13 (standard)	TGACCGGCAGCAAAATG
RevII.Cs.BPB	TGGGTTATTTTTGGAGAAAGGAGC
RevIII.Cs.BPB	AAAAATTGAAAAGAATAAATCCC
RevIV.Cs.BPB	GAAAATTACTCGCTCCTTTCTCC

Appendix B

The sequence of the cloned region *psbB, clpP* and *petB* of *C. sorokiniana*

GenBank accession number FJ623757

TAATTTTTCTGGAATTTCTGCCCAAGCTTGTGAAGCTGATTTTCCTTCAGCAAGACTTGTTTGTACACGACGTTCAATTT
CTTGTTGGAAGAAACCTAAATCCCATTGGTAACGAGTTGGACCAAAAAGTTCAATTGGTGTTGCGGCAGAACCATACCAC
ATAGTTCCAGATACAACAAACGCTGCCCAGAAAACTGCTGCAATACTGCTAGAAAGAACAGTTTCAATATTACCCATACG
AAGCCCGTTATACAATCTTTGTGGTGGACGAACACAAAGGTGGAATAAACCAGCTAAAACTCCTAAGATACCTGCAGCAA
TGTGGTGTGCTGAAATGCCACCCGGGTTATAAGGATCAAATCCTGTAGCATCCCAAGAAGGAGCAACTGCTTGAACAGTT
CCTGTAATACCGTAAGGGTCTGAAACCCAAATACCAGGACCAAATAATCCAGTTACGTGGAAAGCTCCGAAACCAAAACA
AAGAAGACCTGATAAAAATAAGTGAATACCAAAAATTTTTGGAAGGTCTAATGCTGGATTTGATGTGCGTGGGTCACGGA
AAAGCTCAAGATCCCAATAAACCCAGTGCCAAATAGAAGCCGCAAATAGTAATCCAGATAAAATAATGTGAGCAGCAGCA
ACTCCTTCATAACTCCAAATCCCTGGATTTGATGCTGTTTCACCGCTAATAGTCCAACCACCCCAAGATTGAGTGATGCC
TAAACGTGTCATGAAAGGAAGAACAAACATCCCTTGACGCCACATTGGGTTTAAAACTGGATCAGAAGGATCAAAAACAG
CAAGTTCATAAAAAGCCATTGAACCAGCCCAACCAGAAACTAAAGAAGTATGCATTAAATGCACAGCAATTAAACGACCT
GGATCGTTTAAAACTACTGTATGAACTCTATACCATGGGAGACCCATAAAAAATTCTCTCTTTTGTAATTTTTTTTACTT
TCTTTCTTTTCGTATAAAAAAGCGTTTTATTTGCCCTTCCAAAAACAAACTTTAATACTATCCCGTCCATTCCCAAAAAC
ACTGTTTTTGGGATAAATTTTTCAAATTTTAAGATAAACGAAATTTTTGGTATTAAAACATTGTTTTTTAGAACAGGCTA
ATTTTATTTCTTTCTCATTCTGCGAAAAAAAATGTCGTTTTTATATGAAAATTGACGGTTTTAACTAAAAATTGAACTTT
GTGTGTCGCAAAAGCTTCCTTTTTAGGTTGAAACTAATTTTTAGCACTGACTTTCTAAATAATTTTAGAAGTTTTTTTGT
TTGTAAAATTTAACTGATGCCTTAAATAAAGAAAAGCTCATTAAAAATAAAATAAGCTTTAACAAAAAATTTGAAAAATT
TACCAATATCTAGAAGAGCATCAAAAACCGAGTCAAAAATGAGAAAAAAATAACTGCTTTTTTATACTACGTGTTCTTAA
TTTTAGTATATCACATTTTTTTTTACCTAGCATCTTTTCAAGATGCTACAAACTAAGATTTTTACTTATATAAAAAATTA
AAGCAGATAAAAAAGAACGATATTGTAATTTTAATGGGATTTATTCTTTTCAATTTTTTCTTTAATTTTTCGAACTTTTA
GAGTATACTAAATCTAGATCTTTAGTTTTAATAAAAGTCTTTCTTGTTATTTCTTATTATAATATTTTAAATAAAAAACT
ACTAAAGAAATAACAAGAATTTAGAAAGTATTCAACATGTAAATTATGCCAATCGGTGTCCCTAAAGTCCCTTTTCGTTT
ACCCGGTGAACCTGCAGCTCAATGGGTTGATTTATATAACCGTTTATATCGGGAACGCGTTTTATTTCTCTGTCAAGAAT
TAGATGATGAGTTGGCAAATCAACTCATTGGTATTATGCTTTATTTAAATGCAGAAGAACAAAATAAAGGTCTGTATATT

TATATCAATTCTCCTGGTGGATCTGTAACATGTGGTATTGCTGTATACGATGCAATGAACTATATTAAATCAGACGTGAC
TACAATTTGCGTAGGAACAGCAGCTTCAATGGCTTCTTTTATTTTAGCCGGTGGTGATCGTGGAAAACGAATTGCTCTTC
CACATTCTCGTATTATGATTCACCAACCTGAAGGCGGAAGTCAAGGGCAAGCTTCCGAAGTTCTTTCTGAATCTGAAGAA
GTAATGCGAATTAGACGACAAGTCGGTCGCATTTATGCCGAAAGAACAGGTCAACCATTAAGTCGAATTTCGAGGGATAT
GGATCGAGATCAGTTTTTATCCGCTCGTGAAGCAAAAGACTATGGTCTTGTGGATCAAGTAGCTGTCGATACAAAATGGT
CAACAAACTAATCATCCACAATTCGAAAATTACTCGCTCCTTTCTCCAAAAATAACCCAGTAAGTATGTTTTGGGTTAAA
AAAAATATTAATCGCTTTCCCTGTCCCAAAAACGAAGTTTTTAGGGCAAGGAAGCGCTTCCGTAATTTGTTTTTGAGAAT
GGAAGAAAAACGATAATTTTTTGCCTATTTTAGCCTAATTGTGGTCTAACTAGTTTTAATGCGTAGTTGCGTTTTTTTTA
GCCTAAATAGGGTTTATGTAAAAAACAAAGTACGCGTATTTACATCTTTTCAAAATTGTTTATGGGTAAAGTGTACGATT
GGTTTGAAGAACGTCTTGAAATTCAAGCCATTGCTGATGATATTTCAAGTAAATATGTTCCACCACATGTTAATATTTTT
TATTGTATTGGTGGAATCACTTTTACATGTTTTTTAGTTCAAGTAGCAACTGGTTTTGCTATGACTTTTTATTACCGTCC
TACAGTTGCAGAAGCTTTTGCTTCGGTACAGTACATCATGACTGAAGTGAACTTTGGTTGGTTAATTCGTTCAATTCACC
GTTGGTCAGCTAGTATGATGGTTCTTATGATGATCCTTCATGTTTGTCGAGTTTATTTAACTGGTGGATTCAAAAAGCCA
CGTGAACTTACTTGGGTTACTGGTGTTATCATGGCTGTATGTACTGTTTCTTTCGGTGTAACAGGCTATTCACTCCCATG
GGACCAAGTAGGCTACTGGGCGGTAAAAATTGTAACAGGTGTTCCTGACGCTATTCCTGGTGTTGGACCAGCTTTAGTAG
AATTATTAAGAGGTGGTGTT

Appendix C

pU.Cs3-*aadA* 7587bp

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT
TTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATA
ACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT
TTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGAT
GAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATAC
ACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA
TGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAAC
CGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACG
ACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT
TCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCT
CCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC
TCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATT
TAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGT
CAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGC
GCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAG
TTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGA
ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGC
CACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTT
TTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAA
GAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTC
CGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTT
GGCTGCAGGTCGACGGATCCCCTTAATTTTTCTGGAATTTCTGCCCAAGCTTGTGAAGCTGATTTTCCTTCAGCAAGACT
TGTTTGTACACGACGTTCAATTTCTTGTTGGAAGAAACCTAAATCCCATTGGTAACGAGTTGGACCAAAAAGTTCAATTG
GTGTTGCGGCAGAACCATACCACATAGTTCCAGATACAACAAACGCTGCCCAGAAAACTGCTGCAATACTGCTAGAAAGA
ACAGTTTCAATATTACCCATACGAAGCCCGTTATACAATCTTTGTGGTGGACGAACACAAAGGTGGAATAAACCAGCTAA
AACTCCTAAGATACCTGCAGCAATGTGGTGTGCTGAAATGCCACCCGGGTTATAAGGATCAAATCCTGTAGCATCCCAAG
AAGGAGCAACTGCTTGAACAGTTCCTGTAATACCGTAAGGGTCTGAAACCCAAATACCAGGACCAAATAATCCAGTTACG
TGGAAAGCTCCGAAACCAAAACAAAGAAGACCTGATAAAAATAAGTGAATACCAAAAATTTTTGGAAGGTCTAATGCTGG
ATTTGATGTGCGTGGGTCACGGAAAAGCTCAAGATCCCAATAAACCCAGTGCCAAATAGAAGCCGCAAATAGTAATCCAG
ATAAAATAATGTGAGCAGCAGCAACTCCTTCATAACTCCAAATCCCTGGATTTGATGCTGTTTCACCGCTAATAGTCCAA
CCACCCCAAGATTGAGTGATGCCTAAACGTGTCATGAAAGGAAGAACAAACATCCCTTGACGCCACATTGGGTTTAAAAC
TGGATCAGAAGGATCAAAAACAGCAAGTTCATAAAAAGCCATTGAACCAGCCCAACCAGAAACTAAAGAAGTATGCATTA
AATGCACAGCAATTAAACGACCTGGATCGTTTAAAACTACTGTATGAACTCTATACCATGGGAGACCCATAAAAAATTCT
CTCTTTTGTAATTTTTTTTACTTTCTTTCTTTTCGTATAAAAAAGCGTTTTATTTGCCCTTCCAAAAACAAACTTTAATA
CTATCCCGTCCATTCCCAAAAACACTGTTTTTGGGATAAATTTTTCAAATTTTAAGATAAACGAAATTTTTGGTATTAAA
ACATTGTTTTTTAGAACAGGCTAATTTTATTTCTTTCTCATTCTGCGAAAAAAAATGTCGTTTTTATATGAAAATTGACG
GTTTTAACTAAAAATTGAACTTTGTGTGTCGCAAAAGCTTCCTTTTTAGGTTGAAACTAATTTTTAGCACTGACTTTCTA
AATAATTTTAGAAGTTTTTTTGTTTGTAAAATTTAACTGATGCCTTAAATAAAGAAAAGCTCATTAAAAATAAAATAAGC

TTTAACAAAAAATTTGAAAAATTTACCAATATCTAGAAGAGCATCAAAAACCGAGTCAAAAATGAGAAAAAAATAACTGC
TTTTTTATACTACGTGTTCTTAATTTTAGGATGGGTACCGAGCTCCACCGCGGTGGCGGCCGCTCTAGCTAGAACTAGTG
GATCGCACTCTACCGATTGAGTTACATCCGCTTTAGTATGTTACTATTTCTTTTATTATAACTTATAAAATATAATACAT
AAAGATAAATTCTATAATAAAAAGCTAAGATTTTATTTTTCTGGCACATCGTAATTTATAAAGACAGGCAAATTTAAACA
AAAGATAACTTTAGAACTTAATTTTAAAAATGTAAAATGATGTTTAGGTATTTAACCTAAACACCATAAAAATAAAAACG
ATGTTTATGCTATTCACATAAACATCATGAAAAATAAAAATTAAAGTTTGTCAATAGTATCAAATTCGAATTTAATTTCT
TTCCAAACTTCACATGCAGCAGCAAGTTCTGGAGACCATTTACAAGCTGAACGAATTACGTCGCCACCTTCACGAGCAAG
GTCACGACCTTCGTTACGAGCTTGAGTACAAGCTTGCATGCCTGCAGTTATTTGCCAACTACCTTAGTGATCTCGCCTTT
CACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAG
CTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGGCGCGATTTTG
CCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGGCGAGTT
CCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGAC
CTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCTCGAAGATA
CCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTC
GTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGA
TCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCG
CCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGA
TAGTTGAGTTGATACTTCGGCGATAACCGCTTCACGAGCCATGGACATTTTCACTTCTGGAGTGTATTGTTCAATTAAAT
CTTTAATAAGATTACTAAGTTCTTCTGGAGTACGCATTGCCATAAAAAAGAAAAAATAAATAAAAGATTAAAAAAGTTTA
TTTTTAAAATCTTTCTCGAGAATTTTAAATAAGTTTAAAATTCAACAAAAATAGTGAGTGGTAAGATCACTTGTTAACAA
AAGTAATGGTTCACCCTTGTCATATTTAAATACTAAAATTCATTTGCCCGAAGAGGACAAATTTATTTATTGCATTAAAA
TCCCTAAGTTTACTTGCCCGTAAGGGGAAGGGGGGGACGTCCACAGGCGTCGTAAGCAACTAAAGTTTATGACGCCGATT
GCTTTGTTAGGAAAATATAAATATCCCATAAGAAAAGGTCCTTTAAAGGTTTTATGGACTAAATAAAAAAGATAGCATAA
GCATTAAAATCATGCAAATTAAAAAAAAGGTAAATGTATTTATAAAAAGGTAAATGTATTTATATAGTATTTATATTATA
GCATAATAATAAATATATTTATAAATTGATTGTTCTTAGAGCTAAAAGAGAAGAACAATGGGTTTATAGGTATTTTGAGA
TCCAGTTATAAAAATGACTTTTGACGTTTATGGTATATAAACACTGCCTCTAATAAAGTCATCGATAAGCTTGATATCAC
ATTTTTTTTTTACCTAGCATCTTTTCAAGATGCTACAAACTAAGATTTTTACTTATATAAAAAATTAAAGCAGATAAAAAA
GAACGATATTGTAATTTTAATGGGATTTATTCTTTTCAATTTTTTCTTTAATTTTTCGAACTTTTAGAGTATACTAAATC
TAGATCTTTAGTTTTAATAAAAGTCTTTCTTGTTATTTCTTATTATAATATTTTAAATAAAAAACTACTAAAGAAATAAC
AAGAATTTAGAAAGTATTCAACATGTAAATTATGCCAATCGGTGTCCCTAAAGTCCCTTTTCGTTTACCCGGTGAACCTG
CAGCTCAATGGGTTGATTTATATAACCGTTTATATCGGGAACGCGTTTTATTTCTCTGTCAAGAATTAGATGATGAGTTG
GCAAATCAACTCATTGGTATTATGCTTTATTTAAATGCAGAAGAACAAAATAAAGGTCTGTATATTTATATCAATTCTCC
TGGTGGATCTGTAACATGTGGTATTGCTGTATACGATGCAATGAACTATATTAAATCAGACGTGACTACAATTTGCGTAG
GAACAGCAGCTTCAATGGCTTCTTTTATTTTAGCCGGTGGTGATCGTGGAAAACGAATTGCTCTTCCACATTCTCGTATT
ATGATTCACCAACCTGAAGGCGGAAGTCAAGGGCAAGCTTCCGAAGTTCTTTCTGAATCTGAAGAAGTAATGCGAATTAG
ACGACAAGTCGGTCGCATTTATGCCGAAAGAACAGGTCAACCATTAAGTCGAATTTCGAGGGATATGGATCGAGATCAGT
TTTTATCCGCTCGTGAAGCAAAAGACTATGGTCTTGTGGATCAAGTAGCTGTCGATACAAAATGGTCAACAAACTAATCA
TCCACAATTCGAAAATTACTCGCTCCTTTCTCCAAAAATAACCCAGTAAGTATGTTTTGGGTTAAAAAAAATATTAATCG
CTTTCCCTGTCCCAAAAACGAAGTTTTTAGGGCAAGGAAGCGCTTCCGTAATTTGTTTTTGAGAATGGAAGAAAAACGAT
AATTTTTTGCCTATTTTAGCCTAATTGTGGTCTAACTAGTTTTAATGCGTAGTTGCGTTTTTTTTAGCCTAAATAGGGTT
TATGTAAAAAACAAAGTACGCGTATTTACATCTTTTCAAAATTGTTTATGGGTAAAGTGTACGATTGGTTTGAAGAACGT
CTTGAAATTCAAGCCATTGCTGATGATATTTCAAGTAAATATGTTCCACCACATGTTAATATTTTTTATTGTATTGGTGG
AATCACTTTTACATGTTTTTTAGTTCAAGTAGCAACTGGTTTTGCTATGACTTTTTATTACCGTCCTACAGTTGCAGAAG
CTTTTGCTTCGGTACAGTACATCATGACTGAAGTGAACTTTGGTTGGTTAATTCGTTCAATTCACCGTTGGTCAGCTAGT
ATGATGGTTCTTATGATGATCCTTCATGTTTGTCGAGTTTATTTAACTGGTGGATTCAAAAAGCCACAATTCACTGGCCG
TCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC
TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCG
GTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT
AGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGA

Appendix D

The *psbA* promoter region of *C. sorokiniana* H-1983

CCAACCCATGGATAAACGGTTTTCAGTGCTAGTAACCCACTCACAGAAACGAGCCCAAAGGCTAGTGCTTTCACGTCTTT
CTAAAATTGCTGTCATAAAAATTTAAAAGTTTTTTTCGTTTTTCCTCAATAAGGAAGTTTTCCCAAAATTAAAAATTTAA
TTTGTTACTTTTACTATAACCAAAATAAAAAACACATTCAAGTAATTAATTTATTATAAATTAAAATTAGTTTTTAATAA
AAGAGGGTCAAAGAGGCCAGTGTGCAACACTTTGGTTCTGAAACAAGCCAGAAAACTCTTAGTCTACGTCTCTCCGAATC
TTTCTTCAGTTACCTTGTGTTTCGTTTATGTTCCTGAAGATTCTCTTCGATCACCCCTCTCTAAACTTAGATATCAAGC

Appendix E

pUC.Cs.7.3-psbX-*aadA* 7297bp

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT
TTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATA
ACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT
TTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCA
CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGAT
GAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATAC
ACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA
TGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAAC
CGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACG
ACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT
TCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCT
CCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC
TCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATT
TAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGT
CAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGC
GCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAG
TTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGA
ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGC
CACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTT
TTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAA
GAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTC
CGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTT
GGCTGCAGGTCGACGGATCCCCTTAATTTTTCTGGAATTTCTGCCCAAGCTTGTGAAGCTGATTTTCCTTCAGCAAGACT
TGTTTGTACACGACGTTCAATTTCTTGTTGGAAGAAACCTAAATCCCATTGGTAACGAGTTGGACCAAAAAGTTCAATTG
GTGTTGCGGCAGAACCATACCACATAGTTCCAGATACAACAAACGCTGCCCAGAAAACTGCTGCAATACTGCTAGAAAGA
ACAGTTTCAATATTACCCATACGAAGCCCGTTATACAATCTTTGTGGTGGACGAACACAAAGGTGGAATAAACCAGCTAA
AACTCCTAAGATACCTGCAGCAATGTGGTGTGCTGAAATGCCACCCGGGTTATAAGGATCAAATCCTGTAGCATCCCAAG
AAGGAGCAACTGCTTGAACAGTTCCTGTAATACCGTAAGGGTCTGAAACCCAAATACCAGGACCAAATAATCCAGTTACG
TGGAAAGCTCCGAAACCAAAACAAAGAAGACCTGATAAAAATAAGTGAATACCAAAAATTTTTGGAAGGTCTAATGCTGG
ATTTGATGTGCGTGGGTCACGGAAAAGCTCAAGATCCCAATAAACCCAGTGCCAAATAGAAGCCGCAAATAGTAATCCAG
ATAAAATAATGTGAGCAGCAGCAACTCCTTCATAACTCCAAATCCCTGGATTTGATGCTGTTTCACCGCTAATAGTCCAA
CCACCCCAAGATTGAGTGATGCCTAAACGTGTCATGAAAGGAAGAACAAACATCCCTTGACGCCACATTGGGTTTAAAAC
TGGATCAGAAGGATCAAAAACAGCAAGTTCATAAAAAGCCATTGAACCAGCCCAACCAGAAACTAAAGAAGTATGCATTA
AATGCACAGCAATTAAACGACCTGGATCGTTTAAAACTACTGTATGAACTCTATACCATGGGAGACCCATAAAAAATTCT
CTCTTTTGTAATTTTTTTTACTTTCTTTCTTTTCGTATAAAAAAGCGTTTTATTTGCCCTTCCAAAAACAAACTTTAATA
CTATCCCGTCCATTCCCAAAAACACTGTTTTTGGGATAAATTTTTCAAATTTTAAGATAAACGAAATTTTTGGTATTAAA
ACATTGTTTTTTAGAACAGGCTAATTTTATTTCTTTCTCATTCTGCGAAAAAAAATGTCGTTTTTATATGAAAATTGACG
GTTTTAACTAAAAATTGAACTTTGTGTGTCGCAAAAGCTTCCTTTTTAGGTTGAAACTAATTTTTAGCACTGACTTTCTA
AATAATTTTAGAAGTTTTTTTGTTTGTAAAATTTAACTGATGCCTTAAATAAAGAAAAGCTCATTAAAAATAAAATAAGC
TTTAACAAAAAATTTGAAAAATTTACCAATATCTAGAAGAGCATCAAAAACCGAGTCAAAAATGAGAAAAAAATAACTGC
TTTTTTATACTACGTGTTCTTAATTTTAGGATGGGTACCGAGCTCCACCGCGGTGGCGGCCGCTCTAGCTAGAACTAGTG
GATCGCACTCTACCGATTGAGTTACATCCGCTTTAGTATGTTACTATTTCTTTTATTATAACTTATAAAATATAATACAT
AAAGATAAATTCTATAATAAAAAGCTAAGATTTTATTTTTCTGGCACATCGTAATTTATAAAGACAGGCAAATTTAAACA
AAAGATAACTTTAGAACTTAATTTTAAAAATGTAAAATGATGTTTAGGTATTTAACCTAAACACCATAAAAATAAAAACG
ATGTTTATGCTATTCACATAAACATCATGAAAAATAAAAATTAAAGTTTGTCAATAGTATCAAATTCGAATTTAATTTCT
TTCCAAACTTCACATGCAGCAGCAAGTTCTGGAGACCATTTACAAGCTGAACGAATTACGTCGCCACCTTCACGAGCAAG
GTCACGACCTTCGTTACGAGCTTGAGTACAAGCTTGCATGCCTGCAGTTATTTGCCAACTACCTTAGTGATCTCGCCTTT
CACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAG
CTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTG
CCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGGCGAGTT
CCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGAC
CTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCTCGAAGATA
CCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTC
GTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGA
TCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCG
CCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGA
TAGTTGAGTTGATACTTCGGCGATAACCGCTTCACGAGCCATGGATAAACGGTTTTCAGTGCTAGTAACCCACTCACAGA
AACGAGCCCAAAGGCTAGTGCTTTCACGTCTTTCTAAAATTGCTGTCATAAAAATTTAAAAGTTTTTTTCGTTTTTCCTC
AATAAGGAAGTTTTCCCAAAATTAAAAATTTAATTTGTTACTTTTACTATAACCAAAATAAAAAACACATTCAAGTAATT
AATTTATTATAAATTAAAATTAGTTTTTAATAAAAGAGGGTCAAAGAGGCCAGTGTGCAACACTTTGGTTCTGAAACAAG

CCAGAAAACTCTTAGTCTACGTCTCTCCGAATCTTTCTTCAGTTACCTTGTGTTTCGTTTATGTTCCTGAAGATTCTCTT
CGATCACCCCTCTCTAAACTTAGATATCACATTTTTTTTTACCTAGCATCTTTTCAAGATGCTACAAACTAAGATTTTTA
CTTATATAAAAAATTAAAGCAGATAAAAAAGAACGATATTGTAATTTTAATGGGATTTATTCTTTTCAATTTTTTCTTTA
ATTTTTCGAACTTTTAGAGTATACTAAATCTAGATCTTTAGTTTTAATAAAAGTCTTTCTTGTTATTTCTTATTATAATA
TTTTAAATAAAAAACTACTAAAGAAATAACAAGAATTTAGAAAGTATTCAACATGTAAATTATGCCAATCGGTGTCCCTA
AAGTCCCTTTTCGTTTACCCGGTGAACCTGCAGCTCAATGGGTTGATTTATATAACCGTTTATATCGGGAACGCGTTTTA
TTTCTCTGTCAAGAATTAGATGATGAGTTGGCAAATCAACTCATTGGTATTATGCTTTATTTAAATGCAGAAGAACAAAA
TAAAGGTCTGTATATTTATATCAATTCTCCTGGTGGATCTGTAACATGTGGTATTGCTGTATACGATGCAATGAACTATA
TTAAATCAGACGTGACTACAATTTGCGTAGGAACAGCAGCTTCAATGGCTTCTTTTATTTTAGCCGGTGGTGATCGTGGA
AAACGAATTGCTCTTCCACATTCTCGTATTATGATTCACCAACCTGAAGGCGGAAGTCAAGGGCAAGCTTCCGAAGTTCT
TTCTGAATCTGAAGAAGTAATGCGAATTAGACGACAAGTCGGTCGCATTTATGCCGAAAGAACAGGTCAACCATTAAGTC
GAATTTCGAGGGATATGGATCGAGATCAGTTTTTATCCGCTCGTGAAGCAAAAGACTATGGTCTTGTGGATCAAGTAGCT
GTCGATACAAAATGGTCAACAAACTAATCATCCACAATTCGAAAATTACTCGCTCCTTTCTCCAAAAATAACCCAGTAAG
TATGTTTTGGGTTAAAAAAAATATTAATCGCTTTCCCTGTCCCAAAAACGAAGTTTTTAGGGCAAGGAAGCGCTTCCGTA
ATTTGTTTTTGAGAATGGAAGAAAAACGATAATTTTTTGCCTATTTTAGCCTAATTGTGGTCTAACTAGTTTTAATGCGT
AGTTGCGTTTTTTTTAGCCTAAATAGGGTTTATGTAAAAAACAAAGTACGCGTATTTACATCTTTTCAAAATTGTTTATG
GGTAAAGTGTACGATTGGTTTGAAGAACGTCTTGAAATTCAAGCCATTGCTGATGATATTTCAAGTAAATATGTTCCACC
ACATGTTAATATTTTTTATTGTATTGGTGGAATCACTTTTACATGTTTTTTAGTTCAAGTAGCAACTGGTTTTGCTATGA
CTTTTTATTACCGTCCTACAGTTGCAGAAGCTTTTGCTTCGGTACAGTACATCATGACTGAAGTGAACTTTGGTTGGTTA
ATTCGTTCAATTCACCGTTGGTCAGCTAGTATGATGGTTCTTATGATGATCCTTCATGTTTGTCGAGTTTATTTAACTGG
TGGATTCAAAAAGCCACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAAT
CGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCG
CAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCA
CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGG
GCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTC
ATCACCGAAACGCGCGA

Appendix F

pBa3-AX-psbX-*aadA* 9005bp

GCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCC
CTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCG
CATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG
AATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAG
CTAACCGCTTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC
AAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCT
GGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG
GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTT
TAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTG
AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAA
AAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGC
AGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGAC
GATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC
ACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCG
GCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAA
CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAG
CGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCC
GACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTAT
GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCA
AGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGGTATATACATTCA
CCCTTTAAGGCTACCCGGCAGTTAGTTACGGCTTACGTTCCATAAAATATTGGCATATTTTATAAATTATTTTATAGATC
ATATATTTTGTAAATATATAATATTACGCATAACACATTTATTTAAAAACATATAATATTACGCATAACACATTTATTTA
AAAACAGCAAAAACTTGCGTCAAATCCCTATAGGATATTTTTTATTTATGGCACGAGCAGGTTTACACGCTCCGTCAGGA
CGCCGGCACGTAGTTGGAAAGTATGTCCCCCTTGCCCGGAAGGGGAAAGGAGGAGACAAATTTATTTATTGTATATAAAT
AGCATGACTTTCCAAGCGAGTTAACATAAACAAACTGCGCGAGTTAACATAAACAAACTTCCTCTCTGGGGAGGCATGGC
AGCAAATGGCACCCTAAATACATGCGCTTTTAACACAGATAATTATAAACAAGCATAGCGTTAAATCTGCTACCTTGGAT
TAAATCATTAGAAAGAATTTGAGCCGTGTGCAGTGAAAATTGCATGCACGGCTCTTAAGGTTTAAATAAATTTTTAAAGA

AGAAAATTTAACTCCTAACTATGTACTTCCATGGTGCACGTTTTTCAAACTATGAAGCTTGGTTAAGTGACCCTACTCAC
ATTAAACCAAGTGCTCAAGTAGTATGGCCTATTGTAGGTCAAGAAATTTTAAACGGTGATGTAGGTGGTGGTTTCCAAGG
TATTCAAATTACTTCTGGTTTCTTCCAATTATGGCGTGCTAGTGGTATTACTAGTGAATTACAACTTTATACTACAGCAA
TTGGTGGTTTAGTAATGGCTGCTGCAATGTTCTTTGCTGGTTGGTTCCACTACCACAAAGCTGCTCCAAAACTAGAATGG
TTCCAAAACGTTGAATCAATGTTAAACCACCACTTAGGTGGTCTTCTTGGTTTAGGTAGTTTAGCTTGGGCTGGTCACCA
AATTCACGTTTCTTTACCAGTAAACAAATTATTAGATGCTGGTGTAGATCCAAAAGAAATTCCACTTCCTCATGATTTAT
TATTAAATCGTGCTATTATGGCTGACTTATACCCAAGTTTTGCTAAAGGTATTGCTCCTTTCTTTACTTTAAACTGGAGT
GAATACAGTGATTTCTTAACATTTAAAGGTGGTTTAAACCCTGTTACTGGTGGTCTTTGGTTAAGTGATACTGCTCACCA
CCACGTAGCTATTGCTGTATTATTCTTAGTAGCTGGTCACATGTATCGTACTAACTGGGGTATTGGTCACAGTATGAAAG
AAATTTTAGAAGCTCACCGTGGTCCATTTACAGGTGAAGGTCACGTTGGTTTATATGAAATTTTAACAACTTCTTGGCAT
GCACAATTAGCTATTAACTTAGCTTTATTTGGTTCGTTATCAATTATTGTAGCTCACCACATGTACGCAATGCCTCCATA
CCCTTATTTAGCTACTGATTACGGTACACAATTATCATTATTTACACACCACACATGGATTGGTGGTTTCTGTATTGTTG
GTGCTGGTGCTCACGCAGCTATTTTCATGGTTCGTGACTACGATCCTACTAATAACTACAACAACTTATTAGACCGTGTA
ATTCGTCACCGTGATGCTATTATTTCTCACTTAAACTGGGTTTGTATTTTCTTAGGTTTCCACAGCTTTGGTTTATACAT
CCACAACGATACAATGAGTGCTTTAGGTCGTCCTCAAGACATGTTCTCAGATACTGCTATCCAACTTCAACCAGTATTTG
CTCAATGGATTCAAAATACACACTTCTTAGCTCCACAATTAACAGCACCAAATGCTTTAGCTGCTACAAGTTTAACTTGG
GGTGGTGAGCTAGGTGCTCATGGCGGTAAAGTAGCTATGATGCCTATTTCTTTAGGTACTTCTGACTTTATGGTTCACCA
CATTCACGCTTTCACAATTCACGTAACTGTGTTAATTCTTCTGAAAGGTGTTTTATTTGCTCGTAGCTCTCGTCTTATCC
CAGATAAAGCTAACTTAGGTTTCCGTTTCCCTTGTGACGGTCCTGGTCGTGGCGGTACTTGTCAGGTTTCTGCTTGGGAC
CACGTATTCTTAGGTCTTTTCTGGATGTACAACAGCTTATCAATTGTAATTTTCCACTTCAGCTGGAAGATGCAATCTGA
TGTTTGGGGTACGGTTACAGCTTCTGGTGTTTCTCACATTACTGGTGGTAACTTTGCACAAAGCGCTAACACAATCAACG
GTTGGTTACGTGACTTCTTATGGGCACAATCATCACAAGTAATCCAATCATACGGTTCAGCTCTATCTGCTTATGGTTTA
ATTTTCTTAGGTGCTCACTTCGTATGGGCATTCTCGTTAATGTTCTTATTCTCTGGTCGTGGTTACTGGCAAGAACTTAT
CGAATCAATTGTATGGGCTCACAACAAACTTAAAGTTGCACCTGCAATTCAACCACGTGCTTTAAGTATTACTCAAGGTC
GTGCTGTTGGTGTAGCTCACTACCTTTTAGGTGGTATTGCTACTACATGGTCGTTCTTCTTAGCACGTATCATTTCTGTA
GGTTAACATTTAATACTTTTTAATACATATATGCCTAAGTTTATCTTTAAAGATAAACTTAGCCATATGTGTTAAGTTAT
CTAACAAGGTTACCTTTTTATTTCTCTTTAGATATATAAACATTAAAAACTACCGTGATCGTTACACTTTAGATAACTGG
AAGGGGGAAAAATCATGTATTCGCTGGAAGGCGCACCTCCTACTGCCTACTGCGCAGCATTAAAATGCTGTAGATATTGG
TATCTTACAAAGGACAGTAGTACACAATTAAACGCGTTGGAATATTTATATTTTCATTAGAGAATTCTAGGATGGGTACC
GAGCTCCACCGCGGTGGCGGCCGCTCTAGCTAGAACTAGTGGATCGCACTCTACCGATTGAGTTACATCCGCTTTAGTAT
GTTACTATTTCTTTTATTATAACTTATAAAATATAATACATAAAGATAAATTCTATAATAAAAAGCTAAGATTTTATTTT
TCTGGCACATCGTAATTTATAAAGACAGGCAAATTTAAACAAAAGATAACTTTAGAACTTAATTTTAAAAATGTAAAATG
ATGTTTAGGTATTTAACCTAAACACCATAAAAATAAAAACGATGTTTATGCTATTCACATAAACATCATGAAAAATAAAA
ATTAAAGTTTGTCAATAGTATCAAATTCGAATTTAATTTCTTTCCAAACTTCACATGCAGCAGCAAGTTCTGGAGACCAT
TTACAAGCTGAACGAATTACGTCGCCACCTTCACGAGCAAGGTCACGACCTTCGTTACGAGCTTGAGTACAAGCTTGCAT
GCCTGCAGTTATTTGCCAACTACCTTAGTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGA
GGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCT
CCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGC
ACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATC
CTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCA
AGATAGCCAGATCAATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGC
AGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTC
GCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCA
CCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAAC
GTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTTGATACTTCGGCGATAACCGCTTCACGAGC
CATGGATAAACGGTTTTCAGTGCTAGTAACCCACTCACAGAAACGAGCCCAAAGGCTAGTGCTTTCACGTCTTTCTAAAA
TTGCTGTCATAAAAATTTAAAAGTTTTTTTCGTTTTTCCTCAATAAGGAAGTTTTCCCAAAATTAAAAATTTAATTTGTT
ACTTTTACTATAACCAAAATAAAAAACACATTCAAGTAATTAATTTATTATAAATTAAAATTAGTTTTTAATAAAAGAGG
GTCAAAGAGGCCAGTGTGCAACACTTTGGTTCTGAAACAAGCCAGAAAACTCTTAGTCTACGTCTCTCCGAATCTTTCTT
CAGTTACCTTGTGTTTCGTTTATGTTCCTGAAGATTCTCTTCGATCACCCCTCTCTAAACTTAGAATTCCCATATAGCCA
ATGGCTTAAGGAGTGTCATAGGAATAACTAGTCATGCACATTTTCCTAAAAATCTAAAATGTTATTAGAAGCTATACAAA
AATTAAAAATTATGTTATTATATATTTATTAAAATTTAGACTTATCACGGGGATATGGCGGAATGGTAGACGCTACGGAC
TTAAAATCCGTTCTTGTGCGAACAAGGTGAGGGTTCAAGTCCCTCTTTCCCCATTTATAAATAAAAATCAAGTCAATATT
TTAACATTGTAAACTAAAAAATAAGTGTAAAATAACTAGGGTAAATAGTTAAGAATTGGTAGTTTTTGAACTAAACAGTG
ACATAAAACTGGTTAAAAGACATCCCTGTAAGAGAAATGCATATGGTGAATTACACAATAAATTTAAAAATAAAGCTCTG
ACAAGCCTCTTCCCCTTTGCGATATACATGCTCTGTTAGATATAATTCTGCCTAAGTTCCTAACAAAGGGTTTACTTTTT
CTCGGGGAATGGTTTTGCTACTTTGTTTTAAATCTTTTTCAAAGAATGTCAATCTCCTATAAAGGGAGGAACATTACTCA
TGGTCTCACAATTTTAATAAACTAGTTTCTTATAATCTATTTTATAAAAATAAATCTAAATATATTTCTTATGAATTTTG
TTAATTTAGAACAAATTGAAAATTCTTTACGTAATGCTACTTTTTGCATGCTTTTTTTAACAACATTTTTATATTGGTTT
TATACTGCATTTTACAGTACAAATCCTCAACAAATTATAAATCCATTGTCATTAACAAACATTAAAACTAATTTTACTTA
TCCTGACGGGTTTCAAGCAAACAATATTAATGTTTCTACAACTTATTTACCGATTAACCCTGTATTAAATACCGAAAGAG
AAGAGCAACCCGAAGCTAACGGAACGAATGGGCTGTTGGGTGTTTCTTCATTAGTAGTAAATTTAAAATCAATAGCAATC
CCACGCATTATGATGGGTGTTTCTAATTTATTATTAGTATTATTATTACTTGTTCGTTGGGAAAAATCAGGTCATTTCCC
ATTAAGTAATTTATATGAATCTTTAATGTTTTTAGCGTGGTGTTGTACATTTTTATATTTATTATATTGTACAAGTTTTA
CTTTATTGGTTGAAAAAATGTTAGGTTCATTAATTGCACCTTGTAGTTTATTAATGAATGCATTTGCTACTTTTAGTTTA
CCAAAAGAAATGCAACAAGCATCACCATTAGTACCAGCTTTACAATCAAATTGGTTAATGATGCATGTTACTGTAATGAT
TATTAGTTATGCCACATTAATTATTGGATCGTTATTGTCGATTTTATTTTTGATTTTATTTAAACACAAAAAAGGTACAC
CCAAAAAGTATGATAACTTTATTAACAATTTAGATGCATTAAGTTATCGCATTATTGGATTAGGTTTTCCTTTTTTAACT
ATTGGGATTTTATCTGGGGCTGTGTGGGCTAATGAAGCATGGGGATCATATTGGAGTTGGGATCCCCCATCAAGCTTATC

GATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTT
TACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAAT
ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTA
TAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACT
CCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG
TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGT
GGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCA
CCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGTG

pBa3-AX-AAD 9281bp

GCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCC
CTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCG
CATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG
AATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAG
CTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC
AAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCT
GGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG
GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTT
TAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTG
AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAA
AAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGC
AGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGAC
GATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC
ACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCG
GCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAA
CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAG
CGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCC
GACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTAT
GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCA
AGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGGTATATACATTCA
CCCTTTAAGGCTACCCGGCAGTTAGTTACGGCTTACGTTCCATAAAATATTGGCATATTTTATAAATTATTTTATAGATC
ATATATTTTGTAAATATATAATATTACGCATAACACATTTATTTAAAAACATATAATATTACGCATAACACATTTATTTA
AAAACAGCAAAAACTTGCGTCAAATCCCTATAGGATATTTTTTATTTATGGCACGAGCAGGTTTACACGCTCCGTCAGGA
CGCCGGCACGTAGTTGGAAAGTATGTCCCCCTTGCCCGGAAGGGGAAAGGAGGAGACAAATTTATTTATTGTATATAAAT
AGCATGACTTTCCAAGCGAGTTAACATAAACAAACTGCGCGAGTTAACATAAACAAACTTCCTCTCTGGGGAGGCATGGC
AGCAAATGGCACCCTAAATACATGCGCTTTTAACACAGATAATTATAAACAAGCATAGCGTTAAATCTGCTACCTTGGAT
TAAATCATTAGAAAGAATTTGAGCCGTGTGCAGTGAAAATTGCATGCACGGCTCTTAAGGTTTAAATAAATTTTTAAAGA
AGAAAATTTAACTCCTAACTATGTACTTCCATGGTGCACGTTTTTCAAACTATGAAGCTTGGTTAAGTGACCCTACTCAC
ATTAAACCAAGTGCTCAAGTAGTATGGCCTATTGTAGGTCAAGAAATTTTAAACGGTGATGTAGGTGGTGGTTTCCAAGG
TATTCAAATTACTTCTGGTTTCTTCCAATTATGGCGTGCTAGTGGTATTACTAGTGAATTACAACTTTATACTACAGCAA
TTGGTGGTTTAGTAATGGCTGCTGCAATGTTCTTTGCTGGTTGGTTCCACTACCACAAAGCTGCTCCAAAACTAGAATGG
TTCCAAAACGTTGAATCAATGTTAAACCACCACTTAGGTGGTCTTCTTGGTTTAGGTAGTTTAGCTTGGGCTGGTCACCA
AATTCACGTTTCTTTACCAGTAAACAAATTATTAGATGCTGGTGTAGATCCAAAAGAAATTCCACTTCCTCATGATTTAT
TATTAAATCGTGCTATTATGGCTGACTTATACCCAAGTTTTGCTAAAGGTATTGCTCCTTTCTTTACTTTAAACTGGAGT
GAATACAGTGATTTCTTAACATTTAAAGGTGGTTTAAACCCTGTTACTGGTGGTCTTTGGTTAAGTGATACTGCTCACCA
CCACGTAGCTATTGCTGTATTATTCTTAGTAGCTGGTCACATGTATCGTACTAACTGGGGTATTGGTCACAGTATGAAAG
AAATTTTAGAAGCTCACCGTGGTCCATTTACAGGTGAAGGTCACGTTGGTTTATATGAAATTTTAACAACTTCTTGGCAT
GCACAATTAGCTATTAACTTAGCTTTATTTGGTTCGTTATCAATTATTGTAGCTCACCACATGTACGCAATGCCTCCATA
CCCTTATTTAGCTACTGATTACGGTACACAATTATCATTATTTACACACCACACATGGATTGGTGGTTTCTGTATTGTTG
GTGCTGGTGCTCACGCAGCTATTTTCATGGTTCGTGACTACGATCCTACTAATAACTACAACAACTTATTAGACCGTGTA
ATTCGTCACCGTGATGCTATTATTTCTCACTTAAACTGGGTTTGTATTTTCTTAGGTTTCCACAGCTTTGGTTTATACAT
CCACAACGATACAATGAGTGCTTTAGGTCGTCCTCAAGACATGTTCTCAGATACTGCTATCCAACTTCAACCAGTATTTG
CTCAATGGATTCAAAATACACACTTCTTAGCTCCACAATTAACAGCACCAAATGCTTTAGCTGCTACAAGTTTAACTTGG
GGTGGTGAGCTAGGTGCTCATGGCGGTAAAGTAGCTATGATGCCTATTTCTTTAGGTACTTCTGACTTTATGGTTCACCA
CATTCACGCTTTCACAATTCACGTAACTGTGTTAATTCTTCTGAAAGGTGTTTTATTTGCTCGTAGCTCTCGTCTTATCC

CAGATAAAGCTAACTTAGGTTTCCGTTTCCCTTGTGACGGTCCTGGTCGTGGCGGTACTTGTCAGGTTTCTGCTTGGGAC
CACGTATTCTTAGGTCTTTTCTGGATGTACAACAGCTTATCAATTGTAATTTTCCACTTCAGCTGGAAGATGCAATCTGA
TGTTTGGGGTACGGTTACAGCTTCTGGTGTTTCTCACATTACTGGTGGTAACTTTGCACAAAGCGCTAACACAATCAACG
GTTGGTTACGTGACTTCTTATGGGCACAATCATCACAAGTAATCCAATCATACGGTTCAGCTCTATCTGCTTATGGTTTA
ATTTTCTTAGGTGCTCACTTCGTATGGGCATTCTCGTTAATGTTCTTATTCTCTGGTCGTGGTTACTGGCAAGAACTTAT
CGAATCAATTGTATGGGCTCACAACAAACTTAAAGTTGCACCTGCAATTCAACCACGTGCTTTAAGTATTACTCAAGGTC
GTGCTGTTGGTGTAGCTCACTACCTTTTAGGTGGTATTGCTACTACATGGTCGTTCTTCTTAGCACGTATCATTTCTGTA
GGTTAACATTTAATACTTTTTAATACATATATGCCTAAGTTTATCTTTAAAGATAAACTTAGCCATATGTGTTAAGTTAT
CTAACAAGGTTACCTTTTTATTTCTCTTTAGATATATAAACATTAAAAACTACCGTGATCGTTACACTTTAGATAACTGG
AAGGGGGAAAAATCATGTATTCGCTGGAAGGCGCACCTCCTACTGCCTACTGCGCAGCATTAAAATGCTGTAGATATTGG
TATCTTACAAAGGACAGTAGTACACAATTAAACGCGTTGGAATATTTATATTTTCATTAGAGATATCAAGCTTATCGATG
ACTTTATTAGAGGCAGTGTTTATATACCATAAACGTCAAAAGTCATTTTTATAACTGGATCTCAAAATACCTATAAACCC
ATTGTTCTTCTCTTTTAGCTCTAAGAACAATCAATTTATAAATATATTTATTATTATGCTATAATATAAATACTATATAA
ATACATTTACCTTTTTATAAATACATTTACCTTTTTTTTAATTTGCATGATTTTAATGCTTATGCTATCTTTTTTATTTA
GTCCATAAAACCTTTAAAGGACCTTTTCTTATGGGATATTTATATTTTCCTAACAAAGCAATCGGCGTCATAAACTTTAG
TTGCTTACGACGCCTGTGGACGTCCCCCCCTTCCCCTTACGGGCAAGTAAACTTAGGGATTTTAATGCAATAAATAAATT
TGTCCTCTTCGGGCAAATGAATTTTAGTATTTAAATATGACAAGGGTGAACCATTACTTTTGTTAACAAGTGATCTTACC
ACTCACTATTTTTGTTGAATTTTAAACTTATTTAAAATTCTCGAGAAAGATTTTAAAAATAAACTTTTTTAATCTTTTAT
TTATTTTTTCTTTTTTATGGCAATGCGTACTCCAGAAGAACTTAGTAATCTTATTAAAGATTTAATTGAACAATACACTC
CAGAAGTGAAAATGTCCATGGCTCGTGAAGCGGTTATCGCCGAAGTATCAACTCAACTATCAGAGGTAGTTGGCGTCATC
GAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGA
TATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTT
CGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCGT
TATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGAT
CGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCT
TTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCT
GGCGATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGT
CGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATCTTGGAC
AAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACTAAGGTA
GTTGGCAAATAACTGCAGGCATGCAAGCTTGTACTCAAGCTCGTAACGAAGGTCGTGACCTTGCTCGTGAAGGTGGCGAC
GTAATTCGTTCAGCTTGTAAATGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATTCGAATTTGA
TACTATTGACAAACTTTAATTTTTATTTTTCATGATGTTTATGTGAATAGCATAAACATCGTTTTTATTTTTATGGTGTT
TAGGTTAAATACCTAAACATCATTTTACATTTTTAAAATTAAGTTCTAAAGTTATCTTTTGTTTAAATTTGCCTGTCTTT
ATAAATTACGATGTGCCAGAAAAATAAAATCTTAGCTTTTTATTATAGAATTTATCTTTATGTATTATATTTTATAAGTT
ATAATAAAAGAAATAGTAACATACTAAAGCGGATGTAACTCAATCGGTAGAGTGCGATCCACTAGTTCTAGCTAGAGCGG
CCGCCACCGCGGTGGAGCTCGGTACCCATATAGCCAATGGCTTAAGGAGTGTCATAGGAATAACTAGTCATGCACATTTT
CCTAAAAATCTAAAATGTTATTAGAAGCTATACAAAAATTAAAAATTATGTTATTATATATTTATTAAAATTTAGACTTA
TCACGGGGATATGGCGGAATGGTAGACGCTACGGACTTAAAATCCGTTCTTGTGCGAACAAGGTGAGGGTTCAAGTCCCT
CTTTCCCCATTTATAAATAAAAATCAAGTCAATATTTTAACATTGTAAACTAAAAAATAAGTGTAAAATAACTAGGGTAA
ATAGTTAAGAATTGGTAGTTTTTGAACTAAACAGTGACATAAAACTGGTTAAAAGACATCCCTGTAAGAGAAATGCATAT
GGTGAATTACACAATAAATTTAAAAATAAAGCTCTGACAAGCCTCTTCCCCTTTGCGATATACATGCTCTGTTAGATATA
ATTCTGCCTAAGTTCCTAACAAAGGGTTTACTTTTTCTCGGGGAATGGTTTTGCTACTTTGTTTTAAATCTTTTTCAAAG
AATGTCAATCTCCTATAAAGGGAGGAACATTACTCATGGTCTCACAATTTTAATAAACTAGTTTCTTATAATCTATTTTA
TAAAAATAAATCTAAATATATTTCTTATGAATTTTGTTAATTTAGAACAAATTGAAAATTCTTTACGTAATGCTACTTTT
TGCATGCTTTTTTTTAACAACATTTTTATATTGGTTTTATACTGCATTTTACAGTACAAATCCTCAACAAATTATAAATCC
ATTGTCATTAACAAACATTAAAACTAATTTTACTTATCCTGACGGGTTTCAAGCAAACAATATTAATGTTTCTACAACTT
ATTTACCGATTAACCCTGTATTAAATACCGAAAGAGAAGAGCAACCCGAAGCTAACGGAACGAATGGGCTGTTGGGTGTT
TCTTCATTAGTAGTAAATTTAAAATCAATAGCAATCCCACGCATTATGATGGGTGTTTCTAATTTATTATTAGTATTATT
ATTACTTGTTCGTTGGGAAAAATCAGGTCATTTCCCATTAAGTAATTTATATGAATCTTTAATGTTTTTAGCGTGGTGTT
GTACATTTTTATATTTATTATATTGTACAAGTTTTACTTTATTGGTTGAAAAAATGTTAGGTTCATTAATTGCACCTTGT
AGTTTATTAATGAATGCATTTGCTACTTTTAGTTTACCAAAAGAAATGCAACAAGCATCACCATTAGTACCAGCTTTACA
ATCAAATTGGTTAATGATGCATGTTACTGTAATGATTATTAGTTATGCCACATTAATTATTGGATCGTTATTGTCGATTT
TATTTTTGATTTTATTTAAACACAAAAAAGGTACACCCAAAAAGTATGATAACTTTATTAACAATTTAGATGCATTAAGT
TATCGCATTATTGGATTAGGTTTTCCTTTTTTAACTATTGGGATTTTATCTGGGGCTGTGTGGGCTAATGAAGCATGGGG
ATCATATTGGAGTTGGGATCCCCCATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAAT
CGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCG
CAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATT
TTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAG
TTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCA
CTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGG
CGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGT
G

Appendix G

RPL41

TCGCTCTACGCGCAGGGTGCGCCGCCAGCCCGCTGCTCTTGCCTCTGCTGCTGCTGCTGCCAATGTGTGCAGCTGCTGAG
TGGAGGCAGCTGAGTGTGTGTAAGAGCCTGGAGTGCAGCCCCCGATCGCTGTTCAAACATGGCATCCGCCGCCGCCGCTA
CATCATGCTCGCTGCCGCTGCCGCCGCCCGCCCGCAGGCAAGCGGCGTTACGACCGCAAGCAGTCCGGCTATGGTGGCCA
GACCAAGCCCGTGTTCCACAAGAAGGTGCGCAGTGGCAGCAGGGCCGGGGGCATGCAGCACCAGGATGTACTTCAAGGCG
CGACAGCTGTAGCGGGAGCCAACGGCAAGCATGGAGCTCTGGCGCACAGCAGTTGGGCGAAGTGACATGGCGCAGCAGCA
GCAGGGCCCATGGACGACAACAGTCGCAGCTTGGCGCCGCTGGCACAGCATGTGCACCCAGAATAGCCAGTCAGCAGTGC
TGCATGCCAAAGTCCCGGCTGCAGCAGCAGCATTCAAAAAGCAGCTGCGAAAGCGGTGGCAGCAGGAGCAGCAGCAGCAG
CAGCAGCTTGCTCGGCAGCAATGTGCCATATGTTTACAGCACGCATGTCGCATGGCAGTATTGCGGATCCGTGCAGCAAG
CAGGCAGCCGTTGATCCCGCGGCACAGGCTTTGCAGCATGTAGGCGCTACGGGCAGCTGCTCAGCAATGGCTGGCAGCAG
GGGGCCGAGGCCTTGGAAGCAGCATGGCAGCAATGGAGCACGCCAGCATTGGCAGCAATGGCAGCATTCTGCCACCAATG
GTCCCAGCTGCGTTGGCATGCAGCTGCTTGCCGGAGCTTTGTTGTCGTCCCACTGTGCTGGATGCTCACTGACATGTTCC
ATCTCCATGTCTCTGCCCCCCTGCAGGCCAAGACCACCAAGAAGATTGTGCTGCGTCTGGCCTGCACCGTGTGCAAGGCC
CAGCACATGCACGCCATCAAGGTGCGGGGCGGGGCGGCGGAGTGGGGGTGATGGGATGATGACGATCAGCAGCAGCAGTC
ATGGCTGTGGCATCGTGCAGCAGCTGTGGCAGCAGCAGATGGGGCGCTGCGGGCGTCAGGCGGCGGGGTGAGCTGACCGT
GGTGCCGCCCTACGCCGCCCCTGCCCCTGCAGCGCTGCAAGCACTTCGAGATCGGCGGCGACAAGAAGGTACGCTGCATG
GTTGTGATTAAGCTGCATTTGCCTGGCATGCACAGCACCTTGGGGGTGCACTGTAGCGGGCAAGGCAGGCGGTGCGTGCC
CGTGGCGGTCGGAGGAGGGAGCGAGCGCCTGATGCATGTTGAAGGGGTGCTGACTGCTGTTCTCTGCTTCTTATTCCCCT
TTGATCTATTAGGTCAAGGGGCTATACTAGTAAAAAAAAGGAGAGCGGCAGGCGGAGAGCATGACCGAGAAATGGAGGCA
GCGGCGCCCGCTCACCCTGGGCGGCCTGCTCCCCGCCCGGCGCCTCGCTTTGGCCGCCTCTGGCCCCTCGTACCCCTCCT
TGTCCATGCCCTTGTGCTTCCTGAGCTGCCCAACCTGGCATGCAAGCAACCTGGCCGTCTGTAAAGCCATTGGCGTGTGA
GTCTTTGCACCGCAGCTGGCCTTGGGATCTCATGAGACTACAGCCTAGGACGAGGGGTGTGAGATGGGATACACTGCCCG
ATGCACCAGCATTACCCTCTTTCTTGGTGAAC

Appendix H

pASapI 6623bp

AGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA
TCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCG
AATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATC
TGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCC
GGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG
CGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGC
ACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT
TTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT
GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAAT
GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCA
TACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA
ACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGG
CTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGC
CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAG
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA
AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAG
AGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGC
CTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCA
ATTCGAATCCGCGTTTTCTCCGTGAAAGGGAGGTGTCCTAGGCCTCTAGACGATGGGGGCTTTTTGTTATATTTTACTAA
ATATATATTATAATTAAAAAAAATTGAATTGTCAATTTTTAATGTACACTTAGTTGAAAGTGCCCCTGTCCCCTTGGCCA
TATTTAACAGAAGTTATTTATAACGCAGCTGTTTTTTGGAGTCTATAAATTTATAACATCAGTTACTATGGATTTCCCTT
TAGTTTTATGGCCTAGGACGTCCCCTTCCCCTTCGATGCTGGAGGCATCCTTTTACGGGACAATAAATAAATTTGTTGCC
TCGCCTATCGGCTAACAAGTTCCTTCGGAGTATATAAATATAGGATGTTAATACTGCTATAAACTTTAGTTGCCCAATAT
TTATATTAGGACGCCAGTGGCAGTGGTACCGCCACTGCCTGCTTCGCAGTATATAAATATAGGCAGTTGGCAGGCAACTG
CCACTGACGTCCTATTTTAATACTCCCAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACAAATTTATTTATTGTCCACTA

AAATTTATTTGCCCGAAGGGGACGTCCACTAAAATTTATTTACCCGAAGGGGACGTCCTAATATAAATATGGGGATGTCA
ATGCTCCGTTAGGAAGTAACTAACGTTTTTCAAATAAATTTTATCCCGGAGGGAAGTAGGCAGTAGCCCGCCACTGTCAT
CCTTTAAGTGGATCTCTCGTCAGGCAATTTGCTTACACCTTTAAATTAAAAATTAAATTTAAAGAAAAGTGAGCTATTAA
CGCGTCTCCAATATAGTAGACTTTATTAGAGGCAGTGTTTATATACCATAAACGTCAAAAGTCATTTTTATAACTGGATC
TCAAAATACCTATAAACCCATTGTTCTTCTCTTTTAGCTCTAAGAACAATCAATTTATAAATATATTTATTATTATGCTA
TAATATAAATACTATATAAATACATTTACCTTTTTATAAATACATTTACCTTTTTTTTTAATTTGCATGATTTTAATGCTT
ATGCTATCTTTTTTTATTTAGTCCATAAAACCTTTAAAGGACCTTTTCTTATGGGATATTTATATTTTCCTAACAAAGCAA
TCGGCGTCATAAACTTTAGTTGCTTACGACGCCTGTGGACGTCCCCCCCTTCCCCTTACGGGCAAGTAAACTTAGGGATT
TTAATGCAATAAATAAATTTGTCCTCTTCGGGCAAATGAATTTTAGTATTTAAATATGACAAGGGTGAACCATTACTTTT
GTTAACAAGTGATCTTACCACTCACTATTTTTGTTGAATTTTAAACTTATTTAAAATTCTCGAGAAAGATTTTAAAAATA
AACTTTTTTAATCTTTTATTTATTTTTTCTTTTTTATGGGAAGAGCTACTCCATGGATCCTCTAGAGTCGACCTGCAGGC
ATGCAAGCTTGTACTCAAGCTCGTAACGAAGGTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAA
ATGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATTCGAATTTGATACTATTGACAAACTTTAAT
TTTTATTTTTCATGATGTTTATGTGAATAGCATAAACATCGTTTTTATTTTTATGGTGTTTAGGTTAAATACCTAAACAT
CATTTTACATTTTTAAAATTAAGTTCTAAAGTTATCTTTTGTTTAAATTTGCCTGTCTTTATAAATTACGATGTGCCAGA
AAAATAAAATCTTAGCTTTTTATTATAGAATTTATCTTTATGTATTATATTTTATAAGTTATAATAAAAGAAATAGTAAC
ATACTAAAGCGGATGTAGCGCGTTTATCTTAACGGAAGGCCAGTGGCACGTCCGCCTGCCACTGCCGAATATAAATATGG
TTGAGTTGCTTAGTTTACCTTAGCGAAAAGAAGACTTAGCAGCTAGCCTTAACAAACAGTTTTATATTTTATGTTTGTGT
TAAATAAAATTAAGAAACTTTAGCTAAAGTTTCCCAACTCATAGAAACGTCATCTAAAATTAAAGAACTGTTGTAAATTT
CTAAAATGATTAATAAGAATGCTGCAAATAAAAGGATAAATACAGCCATTAAAACAGTTGTACCCCAGCCTGGTAATACT
TTACCTGCTTCTGAGTTAAGTGGACGTAATAAAGTACCTAATGGTGTAACTAAACCAGGTTCTTGGAAGTCTGAATTTAC
TTTTGATGGTTTAGCTTTAGAAGTTCCTGTTGCCATAATTGATTAAATGAATTAAGCGTTATTAGCGCTATTTTATTTAC
TTTCTGTAAAAAATAAGGAAAATATTCTTCAGTGCATTCCCTCTCAGGATTATAAATACTCTGAGGATAACGTTCTCTCG
TCAAGGGGTTGCTTCTTGTGAGTATAGAAACCTACTAGCACAAGAAATAAATTGCATAAAAATGTATTTACCTAGGACCG
CAGTAGGCAGTCCCTTTTCCCCTTCAGAACTGCCTGCTTTAAAAGAATGAAAAAACTGCCTTGTCTGGTAAGTAAAACTC
TTTAATTACTCACTAAAGACGATCTTAGAAGTTCTTTGTTCATTTTTTATTTAATATAATATTTGTTATATAAAAATTAA
ATAATTTTTAATTAATGTTTAACTTTGTAAGGACAGTTTCAAAGTGACATGAATGGCTACTGCAAAAACGAAGTAAGTTA
TTCTTTCTCAGGGCAAAATTTTGAGTAGATTAATTTTGTTTAAAAATGTGGGACACAGTCGTCAAGTCTTTTGAACTATC
TAAGAGATATGTTGAAAAGAGAATAATTTTATTATTAAATGAGCTATGGAAAGTCCAGCTTTTTTCTTTACCTTTTTTTT
ATGGTTTCTTCTGTTAAGTGTAACTGGCTATTCAGTTTATGTTAGTTTTGGTCCACCTTCAAAAAAATTACGTGATCCTT
TTGAAGAACACGAAGATTAAACAAGTTAAAAAGTACTATTTTTACAAGTGACTTCGGTGCCTCTGAGAACCCTAGTTATA
GTGATATAAAATAACTAGCTAACTACTTTATATTTTTATGAAAGTCATTTTGTCGAGCATATAAACAAAAACAAAATTGC
TATACTAGGCAGTCACAGTGCAACTGTCTCCGTCTCCTTAACCGAGAAAGGGTAAACGTCTTCGGTAAAGTAACAAACTT
TAGTTATGTTAACTGCTTGCGAGTTAACCATTTTTTTTCCTCCGAAGGACAACAGTTGGCAGTTGCCAAACTTTAGTGGT
CTAATATTTATATTAGGCAGTTGGCAGGCAACTGCACTGACGTCCCGAAGGGGAAGGGGTTTACTTACCTCCTAACGGAG
TATATAAATAGAATAAAATTTATTTCCTGCGCTAGCAGATTTACATACTAGGATTTTAATACTCCGAAGGAGGCAGTGGC
GGTACCACTGCCACTGGCGTCCTCCTTCCCCTTCGGGCAAATGCATTTTAGTGCCACTTAAGTTTACTTGCCTAGGCAGT
TGGCAGGACGTCAGTGGCAGTGGTACCGCGACTGCCTATATTTATATACTCCTAAGTTTACTTGCCTAGGCAGTTGGCAG
GCAACTGCCACTGACGTCCTTCCCCTTCCCCTTCGGGACGTCCCCTTACGGGAATATAAATATTAGTGGATATTTATATA
CTGCGATGTTTACATACTCCGAAGGAGGAGAGCTAGCAGTTGCCTGCCAACTGCCTAATATAAATATTGGGCAAGTAAAC
TTAGAATGTTTACATACTCCGAAGGAGGACGTCCCTTACGGGAATATAAATATTAGTGGCAGTGGTACCGCCACTGCCTC
CTTCGGAGTATTAAAATCCTAGTATATAATATACCGTAAGGGACGTCCTCCGACGGTGGCAGTGGCGGTACCACTGCCAC
CGGCGTCCTAATATACATATTGAAGTATTTAAACCTGTTAGCGCACGCTCTAACGAGTCAGTAAACTTCCCTTTTGGGGC
TTCTAGGCAGCGCATAAATTTTCTAGGACGAACGTCCACTGGCGTCTCGTAAGGAGCAGTGACAGGCCACTAATGTCCCC
TTAATGGGTAAATAAATGGCTATCGTCTATCCATGAAGAGACCATATATTCCAGTAGCACCGTTATGATCCTCAAAGGGT
AACACCATTTGTATAGTATTATGGTGAAATGCATCCCTTTCAGGGTAGATTTATATCTTACAG

pPSapI 6273bp

AGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA
TCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCG
AATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATC
TGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCC
GGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG
CGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGC
ACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT
TTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT
GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAAT
GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCA
TACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA
ACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGG
CTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGC

CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAG
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA
AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAG
AGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGC
CTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCA
ATTCGAATCCGCGTTTTCTCCGTGAAAGGGAGGTGTCCTAGGCCTCTAGACGATGGGGGCTTTTTGTTATATTTTACTAA
ATATATATTATAATTAAAAAAAATTGAATTGTCAATTTTTAATGTACACTTAGTTGAAAGTGCCCCTGTCCCCTTGGCCA
TATTTAACAGAAGTTATTTATAACGCAGCTGTTTTTTGGAGTCTATAAATTTATAACATCAGTTACTATGGATTTCCCTT
TAGTTTTATGGCCTAGGACGTCCCCTTCCCCTTCGATGCTGGAGGCATCCTTTTACGGGACAATAAATAAATTTGTTGCC
TCGCCTATCGGCTAACAAGTTCCTTCGGAGTATATAAATATAGGATGTTAATACTGCTATAAACTTTAGTTGCCCAATAT
TTATATTAGGACGCCAGTGGCAGTGGTACCGCCACTGCCTGCTTCGCAGTATATAAATATAGGCAGTTGGCAGGCAACTG
CCACTGACGTCCTATTTTAATACTCCCAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACAAATTTATTTATTGTCCACTA
AAATTTATTTGCCCGAAGGGGACGTCCACTAAAATTTATTTACCCGAAGGGGACGTCCTAATATAAATATGGGGATGTCA
ATGCTCCGTTAGGAAGTAACTAACGTTTTTCAAATAAATTTTATCCCGGAGGGAAGTAGGCAGTAGCCCGCCACTGTCAT
CCTTTAAGTGGATCTCTCGTCAGGCAATTTGCTTACACCTTTAAATTAAAAATTAAATTTAAAGAAAAGTGAGCTATTAA
CGCGTAATACTCCGAAGGAGGCAGTTGGCAGGCAACTGCCACTGACGTCCCGTAAGGGTAAGGGGACGTCCACTGGCGTC
CCGTAAGGGGAAGGGGACGTAGGTACATAAATGTGCTAGGTAACTAACGTTTGATTTTTTGTGGTATAATATATGTACCA
TGCTTTTAATAGAAGCTTGAATTTATAAATTAAAATATTTTTACAATATTTTACGGAGAAATTAAAACTTTAAAAAAATT
AACATATGGGAAGAGCTACTCCATGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGTACTCAAGCTCGTAACGAA
GGTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAAATGGTCTCCAGAACTTGCTGCTGCATGTGA
AGTTTGGAAAGAAATTAAATTCGAATTTGATACTATTGACAAACTTTAATTTTTATTTTTCATGATGTTTATGTGAATAG
CATAAACATCGTTTTTATTTTTATGGTGTTTAGGTTAAATACCTAAACATCATTTTACATTTTTAAAATTAAGTTCTAAA
GTTATCTTTTGTTTAAATTTGCCTGTCTTTATAAATTACGATGTGCCAGAAAAATAAAATCTTAGCTTTTTATTATAGAA
TTTATCTTTATGTATTATATTTTATAAGTTATAATAAAAGAAATAGTAACATACTAAAGCGGATGTAGCGCGTTTATCTT
AACGGAAGGCCAGTGGCACGTCCGCCTGCCACTGCCGAATATAAATATGGTTGAGTTGCTTAGTTTACCTTAGCGAAAAG
AAGACTTAGCAGCTAGCCTTAACAAACAGTTTTATATTTTATGTTTGTGTTAAATAAAATTAAGAAACTTTAGCTAAAGT
TTCCCAACTCATAGAAACGTCATCTAAAATTAAAGAACTGTTGTAAATTTCTAAAATGATTAATAAGAATGCTGCAAATA
AAAGGATAAATACAGCCATTAAAACAGTTGTACCCCAGCCTGGTAATACTTTACCTGCTTCTGAGTTAAGTGGACGTAAT
AAAGTACCTAATGGTGTAACTAAACCAGGTTCTTGGAAGTCTGAATTTACTTTTGATGGTTTAGCTTTAGAAGTTCCTGT
TGCCATAATTGATTAAATGAATTAAGCGTTATTAGCGCTATTTTATTTACTTTCTGTAAAAAATAAGGAAAATATTCTTC
AGTGCATTCCCTCTCAGGATTATAAATACTCTGAGGATAACGTTCTCTCGTCAAGGGGTTGCTTCTTGTGAGTATAGAAA
CCTACTAGCACAAGAAATAAATTGCATAAAAATGTATTTACCTAGGACCGCAGTAGGCAGTCCCTTTTCCCCTTCAGAAC
TGCCTGCTTTAAAAGAATGAAAAAACTGCCTTGTCTGGTAAGTAAAACTCTTTAATTACTCACTAAAGACGATCTTAGAA
GTTCTTTGTTCATTTTTTATTTAATATAATATTTGTTATATAAAAATTAAATAATTTTTAATTAATGTTTAACTTTGTAA
GGACAGTTTCAAAGTGACATGAATGGCTACTGCAAAAACGAAGTAAGTTATTCTTTCTCAGGGCAAAATTTTGAGTAGAT
TAATTTTGTTTAAAAATGTGGGACACAGTCGTCAAGTCTTTTGAACTATCTAAGAGATATGTTGAAAAGAGAATAATTTT
ATTATTAAATGAGCTATGGAAAGTCCAGCTTTTTTCTTTACCTTTTTTTTTATGGTTTCTTCTGTTAAGTGTAACTGGCTA
TTCAGTTTATGTTAGTTTTGGTCCACCTTCAAAAAAATTACGTGATCCTTTTGAAGAACACGAAGATTAAACAAGTTAAA
AAGTACTATTTTTACAAGTGACTTCGGTGCCTCTGAGAACCCTAGTTATAGTGATATAAAATAACTAGCTAACTACTTTA
TATTTTTATGAAAGTCATTTTGTCGAGCATATAAACAAAAACAAAATTGCTATACTAGGCAGTCACAGTGCAACTGTCTC
CGTCTCCTTAACCGAGAAAGGGTAAACGTCTTCGGTAAAGTAACAAACTTTAGTTATGTTAACTGCTTGCGAGTTAACCA
TTTTTTTTCCTCCGAAGGACAACAGTTGGCAGTTGCCAAACTTTAGTGGTCTAATATTTATATTAGGCAGTTGGCAGGCA
ACTGCACTGACGTCCCGAAGGGGAAGGGGTTTACTTACCTCCTAACGGAGTATATAAATAGAATAAAATTTATTTCCTGC
GCTAGCAGATTTACATACTAGGATTTTAATACTCCGAAGGAGGCAGTGGCGGTACCACTGCCACTGGCGTCCTCCTTCCC
CTTCGGGCAAATGCATTTTAGTGCCACTTAAGTTTACTTGCCTAGGCAGTTGGCAGGACGTCAGTGGCAGTGGTACCGCG
ACTGCCTATATTTATATACTCCTAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACTGCCACTGACGTCCTTCCCCTTCCC
CTTCGGGACGTCCCCTTACGGGAATATAAATATTAGTGGATATTTATATACTGCGATGTTTACATACTCCGAAGGAGGAG
AGCTAGCAGTTGCCTGCCAACTGCCTAATATAAATATTGGGCAAGTAAACTTAGAATGTTTACATACTCCGAAGGAGGAC
GTCCCTTACGGGAATATAAATATTAGTGGCAGTGGTACCGCCACTGCCTCCTTCGGAGTATTAAAATCCTAGTATATAAT
ATACCGTAAGGGACGTCCTCCGACGGTGGCAGTGGCGGTACCACTGCCACCGGCGTCCTAATATACATATTGAAGTATTT
AAACCTGTTAGCGCACGCTCTAACGAGTCAGTAAACTTCCCTTTTGGGGCTTCTAGGCAGCGCATAAATTTTCTAGGACG
AACGTCCACTGGCGTCTCGTAAGGAGCAGTGACAGGCCACTAATGTCCCCTTAATGGGTAAATAAATGGCTATCGTCTAT
CCATGAAGAGACCATATATTCCAGTAGCACCGTTATGATCCTCAAAGGGTAACACCATTTGTATAGTATTATGGTGAAAT
GCATCCCTTTCAGGGTAGATTTATATCTTACAG

pCSap2 6537bp

AGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA

TCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCG
AATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATC
TGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCC
GGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG
CGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGC
ACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT
TTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT
GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAAT
GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCA
TACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA
ACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGG
CTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGC
CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAG
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA
AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAG
AGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGC
CTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCA
ATTCGAATCCGCGTTTTCTCCGTGAAAGGGAGGTGTCCTAGGCCTCTAGACGATGGGGGCTTTTTGTTATATTTTACTAA
ATATATATTATAATTAAAAAAAATTGAATTGTCAATTTTTAATGTACACTTAGTTGAAAGTGCCCCTGTCCCCTTGGCCA
TATTTAACAGAAGTTATTTATAACGCAGCTGTTTTTTGGAGTCTATAAATTTATAACATCAGTTACTATGGATTTCCCTT
TAGTTTTATGGCCTAGGACGTCCCCTTCCCCTTCGATGCTGGAGGCATCCTTTTACGGGACAATAAATAAATTTGTTGCC
TCGCCTATCGGCTAACAAGTTCCTTCGGAGTATATAAATATAGGATGTTAATACTGCTATAAACTTTAGTTGCCCAATAT
TTATATTAGGACGCCAGTGGCAGTGGTACCGCCACTGCCTGCTTCGCAGTATATAAATATAGGCAGTTGGCAGGCAACTG
CCACTGACGTCCTATTTTAATACTCCCAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACAAATTTATTTATTGTCCACTA
AAATTTATTTGCCCGAAGGGGACGTCCACTAAAATTTATTTACCCGAAGGGGACGTCCTAATATAAATATGGGGATGTCA
ATGCTCCGTTAGGAAGTAACTAACGTTTTTCAAATAAATTTTATCCCGGAGGGAAGTAGGCAGTAGCCCGCCACTGTCAT
CCTTTAAGTGGATCTCTCGTCAGGCAATTTGCTTACACCTTTAAATTAAAAATTAAATTTAAAGAAAAGTGAGCTATTAA
CGCGTGTCACCTCTTACCACCTCTTACCATATTTCTATACTCCAAAGTAACTTTTTACATAAATGTCCCCTCTGGGGCTG
CCTCCTTCCCCTTCCCCTTCGGTATATAAATATAGGGCAAGTAAACTTAGCATAAACTTTAGTTGCCCGAAGGGGTTTAC
ATACTCCGAAGGAGGACAAATTTATTTATTGTGGTACAATAAATAAATTGTATGTAAACCCCTTTCGGGTAACTAAAGTT
TATCACGGCAATAAGTTTCTGCTTACGCAGTATTATATCTGACGCAGTATTATATAAGAAGTTGGCAGGATAAAAATGTG
TAAGTATGGCAATCTTTTAAAATAGTGTTCAATTCATTTAAGGCAGATAAAAAGAAAAAAGTCCACAGGATTTAATTTTG
AATAGTTCTCTATCAAAAAAAGGTTTGCCGAACAATGTTTTTATTCCTGGAGTTTGATTTTATGAAATTAGCTGTTTACG
GAAAAGGTGGTATTGGAAAATCAACGACAATGGGAAGAGCTACTCCATGGATCCTCTAGAGTCGACCTGCAGGCATGCAA
GCTTGTACTCAAGCTCGTAACGAAGGTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTCGTTCAGCTTGTAAATGGTC
TCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATTCGAATTTGATACTATTGACAAACTTTAATTTTTAT
TTTTCATGATGTTTATGTGAATAGCATAAACATCGTTTTTATTTTTATGGTGTTTAGGTTAAATACCTAAACATCATTTT
ACATTTTTAAAATTAAGTTCTAAAGTTATCTTTTGTTTAAATTTGCCTGTCTTTATAAATTACGATGTGCCAGAAAAATA
AAATCTTAGCTTTTTATTATAGAATTTATCTTTATGTATTATATTTTATAAGTTATAATAAAAGAAATAGTAACATACTA
AAGCGGATGTAGCGCGTTTATCTTAACGGAAGGCCAGTGGCACGTCCGCCTGCCACTGCCGAATATAAATATGGTTGAGT
TGCTTAGTTTACCTTAGCGAAAAGAAGACTTAGCAGCTAGCCTTAACAAACAGTTTTATATTTTATGTTTGTGTTAAATA
AAATTAAGAAACTTTAGCTAAAGTTTCCCAACTCATAGAAACGTCATCTAAAATTAAAGAACTGTTGTAAATTTCTAAAA
TGATTAATAAGAATGCTGCAAATAAAAGGATAAATACAGCCATTAAAACAGTTGTACCCCAGCCTGGTAATACTTTACCT
GCTTCTGAGTTAAGTGGACGTAATAAAGTACCTAATGGTGTAACTAAACCAGGTTCTTGGAAGTCTGAATTTACTTTTGA
TGGTTTAGCTTTAGAAGTTCCTGTTGCCATAATTGATTAAATGAATTAAGCGTTATTAGCGCTATTTTATTTACTTTCTG
TAAAAAATAAGGAAAATATTCTTCAGTGCATTCCCTCTCAGGATTATAAATACTCTGAGGATAACGTTCTCTCGTCAAGG
GGTTGCTTCTTGTGAGTATAGAAACCTACTAGCACAAGAAATAAATTGCATAAAAATGTATTTACCTAGGACCGCAGTAG
GCAGTCCCTTTTCCCCTTCAGAACTGCCTGCTTTAAAAGAATGAAAAAACTGCCTTGTCTGGTAAGTAAAACTCTTTAAT
TACTCACTAAAGACGATCTTAGAAGTTCTTTGTTCATTTTTTATTTAATATAATATTTGTTATATAAAAATTAAATAATT
TTTAATTAATGTTTAACTTTGTAAGGACAGTTTCAAAGTGACATGAATGGCTACTGCAAAAACGAAGTAAGTTATTCTTT
CTCAGGGCAAAATTTTGAGTAGATTAATTTTGTTTAAAAATGTGGGACACAGTCGTCAAGTCTTTTGAACTATCTAAGAG
ATATGTTGAAAAGAGAATAATTTTATTATTAAATGAGCTATGGAAAGTCCAGCTTTTTTCTTTACCTTTTTTTTTATGGTT
TCTTCTGTTAAGTGTAACTGGCTATTCAGTTTATGTTAGTTTTGGTCCACCTTCAAAAAAATTACGTGATCCTTTTGAAG
AACACGAAGATTAAACAAGTTAAAAAGTACTATTTTTACAAGTGACTTCGGTGCCTCTGAGAACCCTAGTTATAGTGATA
TAAAATAACTAGCTAACTACTTTATATTTTTATGAAAGTCATTTTGTCGAGCATATAAACAAAAACAAAATTGCTATACT
AGGCAGTCACAGTGCAACTGTCTCCGTCTCCTTAACCGAGAAAGGGTAAACGTCTTCGGTAAAGTAACAAACTTTAGTTA
TGTTAACTGCTTGCGAGTTAACCATTTTTTTTCCTCCGAAGGACAACAGTTGGCAGTTGCCAAACTTTAGTGGTCTAATA

TTTATATTAGGCAGTTGGCAGGCAACTGCACTGACGTCCCGAAGGGGAAGGGGTTTACTTACCTCCTAACGGAGTATATA
AATAGAATAAAATTTATTTCCTGCGCTAGCAGATTTACATACTAGGATTTTAATACTCCGAAGGAGGCAGTGGCGGTACC
ACTGCCACTGGCGTCCTCCTTCCCCTTCGGGCAAATGCATTTTAGTGCCACTTAAGTTTACTTGCCTAGGCAGTTGGCAG
GACGTCAGTGGCAGTGGTACCGCGACTGCCTATATTTATATACTCCTAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACT
GCCACTGACGTCCTTCCCCTTCCCCTTCGGGACGTCCCCTTACGGGAATATAAATATTAGTGGATATTTATATACTGCGA
TGTTTACATACTCCGAAGGAGGAGAGCTAGCAGTTGCCTGCCAACTGCCTAATATAAATATTGGGCAAGTAAACTTAGAA
TGTTTACATACTCCGAAGGAGGACGTCCCTTACGGGAATATAAATATTAGTGGCAGTGGTACCGCCACTGCCTCCTTCGG
AGTATTAAAATCCTAGTATATAATATACCGTAAGGGACGTCCTCCGACGGTGGCAGTGGCGGTACCACTGCCACCGGCGT
CCTAATATACATATTGAAGTATTTAAACCTGTTAGCGCACGCTCTAACGAGTCAGTAAACTTCCCTTTTGGGGCTTCTAG
GCAGCGCATAAATTTTCTAGGACGAACGTCCACTGGCGTCTCGTAAGGAGCAGTGACAGGCCACTAATGTCCCCTTAATG
GGTAAATAAATGGCTATCGTCTATCCATGAAGAGACCATATATTCCAGTAGCACCGTTATGATCCTCAAAGGGTAACACC
ATTTGTATAGTATTATGGTGAAATGCATCCCTTTCAGGGTAGATTTATATCTTACAG

pSSapI 6316bp

AGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA
TCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCG
AATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATC
TGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCC
GGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG
CGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGC
ACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT
TTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT
GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAAT
GATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCA
TACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA
ACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTA
GCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGG
CTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGC
CCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT
GCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAG
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA
AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAG
AGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGC
CTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACC
GTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCA
ATTCGAATCCGCGTTTTCTCCGTGAAAGGGAGGTGTCCTAGGCCTCTAGACGATGGGGGCTTTTTGTTATATTTTACTAA
ATATATATTATAATTAAAAAAAATTGAATTGTCAATTTTTAATGTACACTTAGTTGAAAGTGCCCCTGTCCCCTTGGCCA
TATTTAACAGAAGTTATTTATAACGCAGCTGTTTTTTGGAGTCTATAAATTTATAACATCAGTTACTATGGATTTCCCTT
TAGTTTTATGGCCTAGGACGTCCCCTTCCCCTTCGATGCTGGAGGCATCCTTTTACGGGACAATAAATAAATTTGTTGCC
TCGCCTATCGGCTAACAAGTTCCTTCGGAGTATATAAATATAGGATGTTAATACTGCTATAAACTTTAGTTGCCCAATAT
TTATATTAGGACGCCAGTGGCAGTGGTACCGCCACTGCCTGCTTCGCAGTATATAAATATAGGCAGTTGGCAGGCAACTG
CCACTGACGTCCTATTTTAATACTCCCAAGTTTACTTGCCTAGGCAGTTGGCAGGCAACAAATTTATTTATTGTCCACTA
AAATTTATTTGCCCGAAGGGGACGTCCACTAAAATTTATTTACCCGAAGGGGACGTCCTAATATAAATATGGGGATGTCA
ATGCTCCGTTAGGAAGTAACTAACGTTTTTCAAATAAATTTTATCCCGGAGGGAAGTAGGCAGTAGCCCGCCACTGTCAT
CCTTTAAGTGGATCTCTCGTCAGGCAATTTGCTTACACCTTTAAATTAAAAATTAAATTTAAAGAAAAGTGAGCTATTAA
CGCGTAAGCTTTCTTAATTCAACATTTTTAAGTAAATACTGTTTAATGTTATACTTTTACGAATACACATATGGTAAAAA
ATAAAACAATATCTTTAAAATAAGTAAAAATAATTTGTAAACCAATAAAAAATATATTTATGGTATAATATAACATATGA
TGTAAAAAAAACTATTTGTCTAATTTAATAACCATGCATTTTTTATGAACACATAATAATTAAAAGCGTTGCTAATGGTG
TAAATAATGTATTTATTAAATTAAATAATTGTTATTATAAGGAGAAATATGGGAAGAGCTACTCCATGGATCCTCTAGAG
TCGACCTGCAGGCATGCAAGCTTGTACTCAAGCTCGTAACGAAGGTCGTGACCTTGCTCGTGAAGGTGGCGACGTAATTC
GTTCAGCTTGTAAATGGTCTCCAGAACTTGCTGCTGCATGTGAAGTTTGGAAAGAAATTAAATTCGAATTTGATACTATT
GACAAACTTTAATTTTTATTTTTCATGATGTTTATGTGAATAGCATAAACATCGTTTTTATTTTTATGGTGTTTAGGTTA
AATACCTAAACATCATTTTACATTTTTAAAATTAAGTTCTAAAGTTATCTTTTGTTTAAATTTGCCTGTCTTTATAAATT
ACGATGTGCCAGAAAAATAAAATCTTAGCTTTTTATTATAGAATTTATCTTTATGTATTATATTTTATAAGTTATAATAA
AAGAAATAGTAACATACTAAAGCGGATGTAGCGCGTTTATCTTAACGGAAGGCCAGTGGCACGTCCGCCTGCCACTGCCG
AATATAAATATGGTTGAGTTGCTTAGTTTACCTTAGCGAAAAGAAGACTTAGCAGCTAGCCTTAACAAACAGTTTTATAT

TTTATGTTTGTGTTAAATAAAATTAAGAAACTTTAGCTAAAGTTTCCCAACTCATAGAAACGTCATCTAAAATTAAAGAA
CTGTTGTAAATTTCTAAAATGATTAATAAGAATGCTGCAAATAAAAGGATAAATACAGCCATTAAAACAGTTGTACCCCA
GCCTGGTAATACTTTACCTGCTTCTGAGTTAAGTGGACGTAATAAAGTACCTAATGGTGTAACTAAACCAGGTTCTTGGA
AGTCTGAATTTACTTTTGATGGTTTAGCTTTAGAAGTTCCTGTTGCCATAATTGATTAAATGAATTAAGCGTTATTAGCG
CTATTTTATTTACTTTCTGTAAAAAATAAGGAAAATATTCTTCAGTGCATTCCCTCTCAGGATTATAAATACTCTGAGGA
TAACGTTCTCTCGTCAAGGGGTTGCTTCTTGTGAGTATAGAAACCTACTAGCACAAGAAATAAATTGCATAAAAATGTAT
TTACCTAGGACCGCAGTAGGCAGTCCCTTTTCCCCTTCAGAACTGCCTGCTTTAAAAGAATGAAAAAACTGCCTTGTCTG
GTAAGTAAAACTCTTTAATTACTCACTAAAGACGATCTTAGAAGTTCTTTGTTCATTTTTTATTTAATATAATATTTGTT
ATATAAAAATTAAATAATTTTTAATTAATGTTTAACTTTGTAAGGACAGTTTCAAAGTGACATGAATGGCTACTGCAAAA
ACGAAGTAAGTTATTCTTTCTCAGGGCAAAATTTTGAGTAGATTAATTTTGTTTAAAAATGTGGGACACAGTCGTCAAGT
CTTTTGAACTATCTAAGAGATATGTTGAAAAGAGAATAATTTTATTATTAAATGAGCTATGGAAAGTCCAGCTTTTTTCT
TTACCTTTTTTTTATGGTTTCTTCTGTTAAGTGTAACTGGCTATTCAGTTTATGTTAGTTTTGGTCCACCTTCAAAAAAA
TTACGTGATCCTTTTGAAGAACACGAAGATTAAACAAGTTAAAAAGTACTATTTTTACAAGTGACTTCGGTGCCTCTGAG
AACCCTAGTTATAGTGATATAAAATAACTAGCTAACTACTTTATATTTTTATGAAAGTCATTTTGTCGAGCATATAAACA
AAAACAAAATTGCTATACTAGGCAGTCACAGTGCAACTGTCTCCGTCTCCTTAACCGAGAAAGGGTAAACGTCTTCGGTA
AAGTAACAAACTTTAGTTATGTTAACTGCTTGCGAGTTAACCATTTTTTTTCCTCCGAAGGACAACAGTTGGCAGTTGCC
AAACTTTAGTGGTCTAATATTTATATTAGGCAGTTGGCAGGCAACTGCACTGACGTCCCGAAGGGGAAGGGGTTTACTTA
CCTCCTAACGGAGTATATAAATAGAATAAAATTTATTTCCTGCGCTAGCAGATTTACATACTAGGATTTTAATACTCCGA
AGGAGGCAGTGGCGGTACCACTGCCACTGGCGTCCTCCTTCCCCTTCGGGCAAATGCATTTTAGTGCCACTTAAGTTTAC
TTGCCTAGGCAGTTGGCAGGACGTCAGTGGCAGTGGTACCGCGACTGCCTATATTTATATACTCCTAAGTTTACTTGCCT
AGGCAGTTGGCAGGCAACTGCCACTGACGTCCTTCCCCTTCCCCTTCGGGACGTCCCCTTACGGGAATATAAATATTAGT
GGATATTTATATACTGCGATGTTTACATACTCCGAAGGAGGAGAGCTAGCAGTTGCCTGCCAACTGCCTAATATAAATAT
TGGGCAAGTAAACTTAGAATGTTTACATACTCCGAAGGAGGACGTCCCTTACGGGAATATAAATATTAGTGGCAGTGGTA
CCGCCACTGCCTCCTTCGGAGTATTAAAATCCTAGTATATAATATACCGTAAGGGACGTCCTCCGACGGTGGCAGTGGCG
GTACCACTGCCACCGGCGTCCTAATATACATATTGAAGTATTTAAACCTGTTAGCGCACGCTCTAACGAGTCAGTAAACT
TCCCTTTTGGGGCTTCTAGGCAGCGCATAAATTTTCTAGGACGAACGTCCACTGGCGTCTCGTAAGGAGCAGTGACAGGC
CACTAATGTCCCCTTAATGGGTAAATAAATGGCTATCGTCTATCCATGAAGAGACCATATATTCCAGTAGCACCGTTATG
ATCCTCAAAGGGTAACACCATTTGTATAGTATTATGGTGAAATGCATCCCTTTCAGGGTAGATTTATATCTTACAG

Appendix I

FatB1

ATGGCTTCAGCTCCACCAAAAATTAATGGTTCACCAGTTGGTTTAAAATCAGGTGGTTTAAAAACACAAGAAGATGCTCA
TTCTGCTCCACCACCACGTACATTTATTAATCAATTACCAGATTGGTCAATGTTATTAGCTGCTATTACAACAGTTTTTT
TAGCTGCTGAAAAACAATGGATGATGTTAGATTGGAAACCAAAACGTCCAGATATGTTAGTTGATCCATTTGGTTTAGGT
TCAATTGTTCAAGATGGTTTAGTTTTTCGTCAAAATTTTTCAATTCGTTCATATGAAATTGGTGCTGATCGTACAGCTTC
AATTGAAACAGTTATGAATCACTTACAAGAAACAGCTTTAAATCACGTTAAAATTGCTGGTTTATCAAATGATGGTTTTG
GTCGTACACCAGAAATGTATAAACGTGATTTAATTTGGGTTGTTGCTAAAATGCAAGTTATGGTTAATCGTTATCCAACA
TGGGGTGATACAGTTGAAGTTAATACATGGGTTGCTAAATCAGGTAAAAATGGTATGCGTCGTGATTGGTTAATTTCAGA
TTGTAATACAGGTGAAATTTTAACACGTGCTTCATCAGTTTGGGTTATGATGAATCAAAAAACACGTCGTTTATCAAAAA
TTCCAGATGAAGTTCGTAATGAAATTGAACCACACTTTGTTGATTCACCACCAGTTATTGAAGATGATGATCGTAAATTA
CCAAAATTAGATGAAAAAACAGCTGATTCAATTCGTAAAGGTTTAACACCACGTTGGAATGATTTAGATGTTAATCAACA
CGTTAATAATGTTAAATATATTGGTTGGATTTTAGAATCAACACCACCAGAAGTTTTAGAAACTCAAGAATTATGTTCAT
TAACATTAGAATATCGTCGTGAATGTGGTCGTGAATCAGTTTTAGAATCATTAACAGCTATGGATCCATCAGGTGGTGGT
TATGGTTCACAATTTCAACACTTATTACGTTTAGAAGATGGTGGTGAAATTGTTAAAGGTCGTACAGAATGGCGTCCAAA
AAATGGTGTTATTAATGGTGTTGTTCCAACAGGTGAATCATCTCCAGGTGATTATTCATATCCATATGATGTTCCAGATT
ATGCTTAATAA

Appendix J

>Synechocystis (Chr) 2510000-2512300 reverse

ACCGGGGGTTAATTTCTTAACCTTTCTTTACTTTGATTCTCTGGTCTTCATCGACCTTTTAACCGTCCATCCACAGGAGT
CCAACGCCCCTATGCCCGAGCTTGCTGTCCGCACCGAATTTGACTATTCCAGCGAAATTTACAAAGACGCCTATAGCCGC
ATCAACGCCATTGTCATTGAAGGCGAACAGGAAGCCTACAGCAACTACCTCCAGATGGCGGAACTCTTGCCGGAAGACAA
AGAAGAGTTGACCCGCTTGGCCAAAATGGAAAACCGCCATAAAAAAGGTTTCCAAGCCTGTGGCAACAACCTCCAAGTGA
ACCCTGATATGCCCTATGCCCAGGAATTTTTCGCCGGTCTCCATGGCAATTTCCAGCACGCTTTTAGCGAAGGGAAAGTT
GTTACCTGTTTATTGATCCAGGCTTTGATTATCGAAGCTTTTGCGATCGCCGCCTATAACATATATATCCCTGTGGCGGA
CGACTTTGCTCGGAAAATCACTGAGGGCGTAGTCAAGGACGAATACACCCACCTCAACTACGGGGAAGAATGGCTAAAGG
CCAACTTTGCCACCGCTAAGGAAGAACTGGAGCAGGCCAACAAAGAAAACCTACCCTTAGTGTGGAAAATGCTCAACCAA
GTGCAGGGGGACGCCAAGGTATTGGGCATGGAAAAAGAAGCCCTAGTGGAAGATTTTATGATCAGCTACGGCGAAGCCCT

CAGTAACATCGGCTTCAGCACCAGGGAAATTATGCGTATGTCTTCCTACGGTTTGGCCGGAGTCTAGTCCCGGCTAGAAA
TATCTTCAACTATTTCATGAGCCCACAAATCCCAAGCTTCTTAGGATCTGTGGGTTTTTCCCTATTCTGGAGTGCTTTTA
GGAAGCTGGCCCATGCCAACGGGAAGTTCGTCAAATTGTGCCATACTGTGTAAAGCTTTATGGCGATCGGAGTCGTTAAC
TGAGCGGAAAGCGATACCTTCCCCTATTTATTCTGACTATTATCAATGTTTGGTCTTATTGGTCATCTCACGAGTTTAGA
ACACGCCCAAGCGGTTGCTGAAGATTTAGGCTATCCTGAGTACGCCAACCAAGGCCTGGATTTTTGGTGTTCGGCTCCTC
CCCAAGTGGTTGATAATTTTCAGGTGAAAAGTGTGACGGGGCAGGTGATTGAAGGCAAATATGTGGAGTCTTGCTTTTTG
CCGGAAATGTTAACCCAACGGCGGATCAAAGCGGCCATTCGTAAAATCCTCAATGCTATGGCCCTGGCCCAAAAGGTGGG
CTTGGATATTACGGCCCTGGGAGGCTTTTCTTCAATCGTATTTGAAGAATTTAACCTCAAGCAAAATAATCAAGTCCGCA
ATGTGGAACTAGATTTTCAGCGGTTCACCACTGGTAATACCCACACCGCTTATGTGATCTGCCGTCAGGTCGAGTCTGGA
GCTAAACAGTTGGGTATTGATCTAAGTCAGGCAACGGTAGCGGTTTGTGGCGCCACGGGAGATATTGGTAGCGCCGTATG
TCGTTGGTTAGATAGCAAACATCAAGTTAAGGAATTATTGCTAATTGCCCGTAACCGCCAAAGATTGGAAAATCTCCAAG
AGGAATTGGGTCGGGGCAAAATTATGGATTTGGAAACAGCCCTGCCCCAGGCAGATATTATTGTTTGGGTGGCTAGTATG
CCCAAGGGGGTAGAAATTGCGGGGGAAATGCTGAAAAAGCCCTGTTTGATTGTGGATGGGGGCTATCCCAAGAATTTAGA
CACCAGGGTGAAAGCGGATGGGGTGCATATTCTCAAGGGGGGGATTGTAGAACATTCCCTTGATATTACCTGGGAAATTA
TGAAGATTGTGGAGATGGATATTCCCTCCCGGCAAATGTTCGCCTGTTTTGCGGAGGCCATTTTGCTAGAGTTTGAGGGC
TGGCGCACTAATTTTTCCTGGGGCCGCAACCAAATTTCCGTTAATAAAATGGAGGCGATTGGTGAAGCTTCTGTCAAGCA
TGGCTTTTGCCCTTTAGTAGCTCTTTAGGCCGACAGGATAGGGCGTGTGTGGAGCTGGTGAACAAGGTCTGAAACAGTGG
CCAGTAGTCAATATTTCCCCGTGCTTTCCATAGAAGATATCTGCTTAATCTAAAATATTCCTGCCTTCCTAGAGACTTTA
TAAATGGGTGAATTCAGAGGATTCGTGCCTAAAACTGCTGGGCTAGCTCAAAGAGGGACTGTCTTAGGTAAACTTTTGAA
TCAAAGACTAATTTGACAACCAATTATCAAAGCTCAATGCTGCTGGTCAAAGAGTCTGGCT

Appendix K

HPV16-*L1*

ATGCAAGTTACATTCATTTACATTTTAGTTATTACATGTTACGAAAACGACGTTAACGTTTACCACATTTTCTTCCAAAT
GTCTTTATGGTTACCATCAGAAGCTACAGTTTACTTACCACCAGTTCCAGTTTCAAAAGTTGTTTCAACAGACGAATACG
TTGCTCGTACAAACATTTACTACCACGCTGGTACATCACGTTTATTAGCTGTTGGTCACCCATACTTCCCAATTAAAAAA
CCAAACAACAACAAAATTTTAGTTCCAAAAGTTTCAGGTTTACAATACCGTGTTTTCCGTATTCATTTACCAGACCCAAA
CAAATTCGGTTTCCCAGACACATCATTCTACAACCCAGACACACAACGTTTAGTTTGGGCTTGTGTAGGTGTTGAAGTTG
GACGTGGTCAACCATTAGGTGTTGGTATTTCAGGTCACCCATTATTAAACAAATTAGACGACACAGAAAACGCTTCAGCT
TATGCTGCTAACGCTGGTGTTGATAACCGTGAATGTATTTCAATGGACTACAAACAAACTCAATTATGTTTAATTGGTTG
TAAACCACCAATTGGTGAACACTGGGGTAAAGGTTCACCATGTACAAACGTTGCTGTTAACCCAGGTGACTGTCCACCAT
TAGAATTAATTAACACAGTTATTCAAGACGGTGACATGGTTCACACAGGTTTTGGTGCTATGGACTTCACAACATTACAA
GCTAACAAATCAGAAGTTCCATTAGACATTTGTACATCAATTTGTAAATACCCAGACTACATTAAAATGGTTTCAGAACC
ATACGGTGACTCATTATTCTTCTACTTACGTCGTGAACAAATGTTCGTTCGTCACTTATTCAACCGTGCTGGTACAGTTG
GTGAAAACGTTCCAGACGACTTATACATTAAAGGTTCAGGTTCAACTGCTAATTTAGCTTCATCAAACTACTTCCCAACA
CCAAGTGGTTCAATGGTTACATCAGACGCTCAAATTTTCAACAAACCATACTGGTTACAACGTGCTCAAGGTCACAACAA
CGGTATTTGTTGGGGTAACCAATTATTCGTTACAGTTGTTGACACAACACGTTCAACAAACATGTCATTATGTGCTGCTA
TTTCAACATCAGAAACAACATACAAAAACACAAACTTCAAAGAATACTTACGTCACGGTGAAGAATACGACTTACAATTC
ATTTTCCAATTATGCAAAATCACATTAACAGCTGACGTTATGACATACATTCACTCAATGAACTCAACAATTTTAGAAGA
TTGGAACTTCGGTTTACAACCACCTCCAGGTGGTACATTAGAAGATACATACCGTTTCGTTACACAAGCAATCGCTTGTC
AAAAACACACACCTCCAGCTCCAAAAGAAGATGACCCATTAAAAAAATACACATTCTGGGAAGTTAACTTAAAAGAAAAA
TTCTCAGCTGACTTAGACCAATTCCCATTAGGTCGTAAATTCTTATTACAAGCAGGTTTAAAAGCTAAACCAAAATTCAC
ATTAGGTAAACGTAAAGCAACACCAACAACATCATCTACATCAACAACAGCAAAACGTAAAAAACGTAAATTATAATAAG
CATGC