DR. ANGELA SUN (Orcid ID : 0000-0002-6009-0066)

Article type : Regular Article

Comparative assessment of the *Euglena gracilis* var. *saccharophila* variant strain as a producer of the β-1,3-glucan paramylon under varying light conditions

Angela Sun, Mafruha Tasnin Hasan,
Department of Molecular Sciences, Macquarie University, Sydney, Australia
Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia

Graham Hobba,
Agritechnology Pty Ltd, Borenore, Australia

Helena Nevalainen,
Department of Molecular Sciences, Macquarie University, Sydney, Australia
Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia

Junior Te'o
Department of Molecular Sciences, Macquarie University, Sydney, Australia
Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia

School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, Australia

Corresponding author:
Assoc. Prof. Valentino Setoa Junior Te'o
Queensland University of Technology
Brisbane, QLD, 4001, Australia
Phone: +61 7 3138 1687
Email: junior.teo@qut.edu.au

Running title: *Euglena gracilis*-paramylon and proteomic studies

Editorial Responsibility: L. Graham (Associate Editor)

Abstract

Euglena gracilis Z and a 'sugar loving' variant strain *Euglena gracilis* var. *saccharophila*, were investigated as producers of paramylon, a β-1,3-glucan polysaccharide with potential medicinal and industrial applications. The strains were grown under diurnal or dark growth conditions on a glucose-yeast extract medium supporting high-level paramylon production. Both strains produced the highest paramylon yields (7.4 to 8 $g \cdot L^{-1}$, respectively) while grown in the dark, but the maximum yield was achieved faster by *E. gracilis* var. s*accharophila* (48 h vs 72 h). The glucose-to-paramylon yield coefficient $Y_{par/glu}$ = 0.46±0.03 in the *Euglena gracilis* var. *saccharophila* cultivation, obtained in this study, is the highest reported to date.

Proteomic analysis of the metabolic pathways provided molecular clues for the strain behavior observed during cultivation. For example, overexpression of enzymes in the gluconeogenesis/glycolysis pathways including fructokinase-1 and chloroplastic fructose-1,6-bisphosphatase may have contributed to the faster rate of paramylon accumulation in *E. gracilis* var. *saccharophila*. Differentially expressed proteins in the early steps of chloroplastogenesis pathway including plastid uroporphyrinogen decarboxylases, photoreceptors, and a highly abundant (68 fold increase) plastid transketolase may have

provided the *E. gracilis* var. *saccharophila* strain an advantage in paramylon production during diurnal cultivations.

In conclusion, the variant strain *E. gracilis* var. *saccharophila* seems to be well suited for producing large amounts of paramylon. This work has also resulted in the identification of molecular targets for future improvement of paramylon production in *E. gracilis*, including the chloroplastic fructose-1,6-bisphosphatase (FBP) and phosophofructokinase 1 (PFK-1), the latter being a key regulator of glycolysis.

Keywords

Paramylon, β-1,3-glucan, *Euglena gracilis*, *E. gracilis* var. *saccharophila*, proteomic analysis, metabolic pathways

Abbreviations

NSAF, Normalised spectral abundance factor

PFK-1, phosphofructokinase 1

FBP, chloroplastic fructose-1,6-bisphosphatase

PEPCKs, phosphoenolpyruvate carboxykinase isoforms

UROD, plastid uroporphyrinogen decarboxylase

GSA, glutamate 1-semialdehyde 2,1-aminotransferases

TK, plastid transketolase

Introduction

The flagellated fresh water protist *Euglena gracilis* produces a wide range of compounds of medical and industrial interest. These include lipids for potential use as biofuel, dietary supplements such as vitamin E and free *amino acids*, and the bioactive carbohydrate paramylon (Price 1990, Mahapatra et al. 2013). The linear β-1,3-glucan polysaccharide paramylon is the principal energy-storage compound in *E. gracilis*. Paramylon has attracted special interest in the context of medical and veterinary applications (Rodriguez-Zavala et al. 2010). It also finds uses as a dietary supplement for lowering cholesterol levels (Wang et al. 1997) and moderating postprandial glycaemic responses (Wood 1994).

Euglena gracilis has been cultivated heterotrophically- and photoautotrophically on various nutrient compositions, ranging from simple and chemically defined formulae to complex media with industrial byproducts (e.g., molasses, corn steep solids and yeast extract) to accumulate 50-75% of the cell biomass as paramylon (Barsanti et al. 2001, Šantek et al. 2009, Ivušić and Šantek 2015, Ogawa et al. 2015b). However, the yield is negatively impacted by photo-induced degradation of paramylon and inhibition of growth and carbon assimilation (Rodriguez-Zavala et al. 2010). In order to avoid photo-inhibitory effects, most cultivations of *E. gracilis* have been performed under heterotrophic conditions, either in the dark using the Z strain, or using permanently bleached variants which are no longer capable of photosynthetic growth (Barsanti et al. 2001, Šantek et al. 2009, Rodriguez-Zavala et al. 2010, Šantek et al. 2012).

Euglena gracilis var. *saccharophila*, a variant of *E. gracilis*, was first isolated in Germany in the 1950s. It was described as more efficient at assimilating glucose than the Z strain and capable of utilizing a wider range of carbohydrates including glutamate, malate, pyruvate, succinate and ethanol (Rodriguez-Zavala et al. 2006). Apart from this information, the properties of *Euglena gracilis* var. *saccharophila* have remained largely uncharacterized.

Recent transcriptomic studies on various metabolic pathways in *Euglena gracilis* (O'Neill et al. 2015, Yoshida et al. 2016) have provided further insights into the molecular nature and gene function of the organism (Ogawa et al. 2015a, Tanaka et al. 2017). There are also recent developments in the studies into the genome sequence and organization but no reports in high throughput proteomic studies with *Euglena* (Ebenezer et al. 2017).

In this study, we have compared paramylon production by *Euglena gracilis Z* and var *saccharophila* strains during a diurnal cultivation cycle and in a completely dark environment on a medium optimized for paramylon production. The experiments were performed with strains trained on the optimized medium containing glucose and yeast extract to maximize the paramylon yield. Proteome profiles of the two strains cultivated in the diurnal and dark conditions were compared for differentially expressed proteins, focussing on pathways and enzymes that may contribute to paramylon metabolism or regulation of paramylon synthesis in varying light conditions, with a view of finding targets for strain improvement.

Materials and Methods

Strains and cultivation media

The *Euglena gracilis* Z strain was obtained from the Southern Biological Pty Ltd (Australia) and the *Euglena gracilis* var. *saccharophila* strain was from the University of Texas Culture Collection (https://utex.org/). Cultures were maintained in the *Euglena* liquid medium (EM) which consists of (per liter): 1 g sodium acetate, 1 g peptone, 2 g tryptone, 2 g yeast extract and 0.01 g $CaCl_2 \cdot 2H_2O$.

The medium optimized for paramylon production (1G1YE) contained (per litre): 17.7 g glucose, 10 g yeast extract, 0.2 g $CaCO_3$, 0.5 g $MgSO_4$, 0.4 g $(NH_4)_2HPO_4$, 0.2 g KH_2PO_4, 1.81 g NH_4Cl, 2 mL of trace mineral stock A and 1 mL of trace mineral stock B. Trace mineral stock A was made by dissolving 2.2 g $ZnSO_4 \cdot 7H_2O$, 2 g $MnSO_4 \cdot 4H_2O$, 0.5 g $Na_2MoO_4 \cdot 2H_2O$, 0.04 g $CoCl_2 \cdot 6H_2O$ in 3 mL of concentrated HCl and the solution was diluted with ddH_2O to 50 mL. Trace mineral stock B was made following the recipe by Rodriguez-Zavala et al. (2010). All chemicals were purchased from Sigma Aldrich (Australia) unless stated otherwise.

Both strains used in the study have been passaged in 1G1YE for over 100 generations (culture diluted 1:5 every seven d) over the course of two y.

Culture conditions

A seed culture was produced by inoculating 10 mL of a stock culture maintained in the EM medium into 50 mL of 1G1YE in a 250 mL Erlenmeyer flask and cultivating until the cell density reached $1x10^7$ cells $\cdot$ mL^{-1}. The seed culture was then used to start the experimental culture in a fresh 1G1YE medium with an initial concentration of 10^7 cells $\cdot$ 50 mL^{-1} in a 250

mL Erlenmeyer flask. The cultures were incubated at 22°C on an orbital shaker at 120 rpm. This regime was used for all cultivations. All experimental cultures were performed with three biological replicates.

Cultures under a diurnal cycle were exposed to incandescent light of approximately 1500 lumens for 12 h and then kept in the dark for 12 h with the shaker lid covered by aluminium foil. The diurnal cycle was repeated throughout the entire nine days of cultivation. Cultures grown totally in the dark were kept in 250 mL Erlenmeyer flasks covered with aluminium foil for the entire duration of the cultivation except for sample taking.

In order to deplete the residual paramylon from the cultures for the paramylon accumulation study, the *E. gracilis* Z and the *E. gracilis var. saccharophila* strains were seeded into the 1G1YE medium without glucose and allowed to grow for three days before starting the actual cultivation cycle by adding the appropriate amount of glucose ($17.7\ g \cdot L^{-1}$).

Measurement of biomass and cell concentration

Triplicates of 10 mL aliquots of each culture were centrifuged at 2,000*g* and the cell pellets washed with warm ddH_2O twice before resuspending in 10 mL ddH_2O. The cell pellets were subsequently dried in pre-weighed aluminium boats at 70°C for 48 h prior to weighing. Cell concentration of each aliquot was measured on a hemocytometer (BioRad, Australia).

Determination of the paramylon content

Methods for the extraction and determination of the amount of paramylon produced were adapted from Rodriguez-Zavala et al. (2010). Cells (10 mL) were washed by centrifugation at 2,000*g* and resuspended in 9 mL ddH_2O before addition of 1 mL of 30% (v/v) ice cold perchloric acid (PCA) for cell lysis. The mixture was vortexed vigorously before centrifuged at 1,100*g* for 2 min. Supernatant was removed, the pellet resuspended in 10 mL of 1% (w/v) SDS and placed in boiling water for 15 min. The suspension was then centrifuged at 1,100*g* for 15 min and washed twice in ddH_2O before resuspending in 10 mL of 1N NaOH. To determine the paramylon concentration in the samples, acid hydrolysis assay was carried out as per Rodriguez-Zavala et al. (2010). Paramylon concentration was determined by plotting the absorbance of the sample at 492 nm to a standard curve prepared with purified paramylon purchased from Sigma-Aldrich. All measurements were made in technical triplicates.

Paramylon amount in the cultures is expressed in the following ways: (i) volumetric yield (the amount of paramylon produced per litre of culture as $g \cdot L^{-1}$); (ii) mass fraction (the amount of paramylon per gram of biomass as $g \cdot g^{-1}$), and (iii) glucose-to-paramylon coefficient $Y_{par/glu}$ (the ratio of the initial glucose in the culture converted to paramylon).

Statistical analysis

All experimental data were collected from at least three independent cell cultures. Student's t-test was performed on unpaired samples to determine whether the sample values differed significantly.

Protein extraction, fractionation and trypsin in-gel digestion

The cells were harvested on day 4 of cultivation in the dark, 24 h after the addition of glucose (day 4; Fig. 1). Three biological replicates were cultivated for each condition. Cell samples of 100 mg wet weight were collected from both *Euglena gracilis* Z and *Euglena gracilis* var. *saccharophila* cultures and sonicated for five cycles of 10 s sonication, 30 s rest at 25% of the maximum amplitude (Digital Sonifier 450, Branson-Emerson, US) in a lysis buffer containing 50 mM Tris-HCl (pH 7.4), 3 mM EDTA, 250 mM sucrose, 0.04% β-mercaptoethanol, and protease inhibitor (cOmplete™ Protease Inhibitor Cocktail, Roche US). Supernatants of the sonicated samples were collected after centrifugation at 1,500*g* for 5 min followed by precipitation, extraction and fractionation as described previously (Hasan et al. 2017). Protein concentration was determined by the BCA assay according to manufacturer's recommendations (ThermoFisher, Australia). Fractionation of the protein samples was carried out using the XCell SureLock Mini-Cell Electrophoresis System (ThermoFisher, AU). All gels and reagents used for SDS-PAGE were purchased from ThermoFisher (AU) and the procedures were carried out as per manufacturer's instructions unless mentioned otherwise. The in-gel trypsin digestion of the proteins and the subsequent peptide extraction and purification process was carried out as previously described (Hasan et al. 2017).

Nanoflow LC-MS/MS and data analysis The peptides were analysed on a reverse-phase nanoLC-MS/MS QExactive mass spectrometer (Thermo Scientific, US) following the method outlined in Hasan et al. (2017).

Protein identification

Raw nanoLC-MS/MS files were converted into mzXML format and analyzed using the global proteome machine (GPM) Fury version 3.0 with the X!Tandem algorithm (http://www.thegpm.org; Craig and Beavis 2003, 2004). The MS/MS files were searched against custom databases consisting of protein sequences retrieved from the NCBI, Uniprot and TrEMBL databases in April 2016 (NCBI: http://www.ncbi.nlm.nih.gov/, Uniprot and TrEMBL: http://www.uniprot.org/), and against common human and trypsin peptide contaminants. The data resultant GPM hits were filtered according to criteria previous described (Hasan et al. 2017). The sixteen fractions of each replicate were processed separately on GPM with output files for each fraction and merged into a non-redundant output file for peptide and protein identifications with a log (e) value cut-off of -1.5. The peptide FDR was calculated as 2 x (total number of peptides representing reversed hits/total number of peptides representing all hits) x 100, and the protein FDR was calculated as (total number of reversed protein hits/total number of all proteins) x 100.

Quantitative analysis

Quantitative analysis was carried out by calculating the normalized abundance factor (NSAF) for each protein (Mirzaei et al. 2012, Neilson et al. 2014). Log-transformed NSAF data was then used for the two-sample unpaired t tests and the proteins with a t test *p*-value of less than 0.05 and t-value higher than 2 or lower than -2 were considered to be differentially expressed. This gave the relative abundance value of a protein between the Z and variant cultures, which is expressed as fold change (FC). The enzyme class (EC), expected values (E-value) and accession numbers were extracted from the BLAST search results. Protein hits are arranged based on the pathways involved, with a separate category “Glycolysis/gluconeogenesis, reversible” designated for enzymes that catalyse reversible reactions, thus involved in both pathways.

Chlorophyll extraction

Chlorophyll was extracted and the amount calculated according to the methods described in Jeffrey and Humphrey (1975). Briefly, 1 mL of cells was harvested daily during both diurnal and dark cultivations and resuspended in 90% (v/v) acetone, followed by incubation on ice for 30 min. The cells were then centrifuged at 2,000*g* to sediment solid matter and the absorbance of the supernatant was measured at 664 nm and 647 nm. Chlorophyll concentration was calculated using the absorbance readings according to the equation for higher plans and green algae containing chlorophylls *a* and *b*: Chlorophyll *a* = 11.93 A_{664} –

1.93 A_{647} and Chlorophyll *b* = 20.36 A_{647} – 5.50 $A_{664,}$ described in Jeffrey and Humphrey (1975). Total chlorophyll amount was calculated as the sum of chlorophyll *a* and *b*.

Results and Discussion

Accumulation of the biomass and paramylon under different light conditions

The *Euglena gracilis* var. *saccharophila* and the *E. gracilis Z* strains were cultured under a diurnal cycle (12:12 h light:dark), and in constant darkness for nine d in order to compare accumulation of paramylon as a function of time. The strains were first grown with no carbon for 72 h to deplete the paramylon reserve before addition of glucose on day 3. The day 3+ samples were collected on day 3 immediately after the addition of glucose; this served as the initial reference point for the residual glucose analysis (Fig. 1, A and B).

Following the addition of glucose, the *Euglena gracilis* var. *saccharophila* cultures reached the peak concentration of paramylon at approximately 24 h (Fig. 1B; day 4) when cultivated in the dark, while the *E. gracilis Z* cultures required a longer cultivation period of approximately another 24 h (Fig. 1A; day 5) to reach the maximum paramylon concentration. This could be due to the noticeably higher rate of glucose assimilation by *E. gracilis* var. *saccharophila* (residual glucose depleted in 48 h, day 5) compared to the *E. gracilis Z* (residual glucose depleted in 72 h, day 6).

The highest paramylon levels produced by *Euglena gracilis* var. *saccharophila* (8.1 ± 0.3 g · L^{-1}) and *E. gracilis Z* strain (7.5 ± 0.4 g · L^{-1}) were similar when the strains were cultivated in the dark, whereas under diurnal condition the *E. gracilis* var. *saccharophila* cultures produced significantly more paramylon (4.9 ± 0.4 g · L^{-1}) than the Z strain (3.6 ± 0.1 g · L^{-1}) (Fig. 2, A and B). Although there appeared to be no significant differences in the total amount of paramylon accumulated by the two strains in the dark, the faster metabolism of sugar and accumulation of paramylon shown by the variant strain may have facilitated a shorter turnaround time thus potentially reducing the production costs. Glucose depletion also happened faster in the variant strain under both conditions (Figure 1B).

The highest amount of paramylon produced per gram of biomass (mass fraction) achieved in this study was about 80% for the variant strain *Euglena gracilis var. saccharophila*, when cultivated in the dark. This is higher than the most commonly reported 50-60% (Šantek et al.

2009, Ivušić and Šantek 2015, Ogawa et al. 2015b) and on par with the 75% achieved with repeated-batch cultivation of the Z strain with an over-supply of glucose (Šantek et al. 2012).

We also examined the efficiency of conversion of glucose to paramylon. The highest glucose to paramylon coefficient value obtained in this study was $Y_{par/glu} = 0.46 \pm 0.03$ (Fig. 1A; dark cultivation of the *Euglena gracilis Z* strain), which is comparable to, if not exceeding the numbers published so far (Barsanti et al. 2001, Rodriguez-Zavala et al. 2010, Grimm et al. 2015, Ivušić and Šantek 2015). Other approaches to increase volumetric production of paramylon using very high amounts of glucose and other complex carbohydrate sources have returned much lower $Y_{par/glu}$ values (Šantek et al. 2009, Šantek et al. 2012).

During the diurnal cycle, the yield of paramylon in the *Euglena gracilis* var. *saccharophila* cultures was significantly higher to that of the Z strain cultures as a direct consequence of having a higher paramylon mass fraction (Fig. 2B). This indicates that the photo-inducible degradation of paramylon was not as pronounced in the variant strain as it was in the Z strain.

In summary, both strains were found to produce more paramylon when cultivated in the dark, producing approximately 8 $g \cdot L^{-1}$ of paramylon compared to 4-5 $g \cdot L^{-1}$ under diurnal cultivation. Although both strains accumulated less paramylon under the diurnal condition, the variant strain was still somewhat more productive than the Z strain. Since the paramylon catabolism has been demonstrated to be photo-inducible, it is possible that the differences of paramylon levels between the strains arise from their apparent differing sensitivity, or the extent of the photoinduced reaction to light (Schwartzbach et al. 1975, Barsanti et al. 2001).

Accumulation of chlorophylls in the Euglena gracilis *Z and E.* gracilis var. saccharophila *strains*

Chlorophylls play important roles in the photosynthesis process and chlorophyll concentration is often used as an indicator of photosynthesis activity level in plants and other photosynthetic organisms (Nagaraj et al. 2002, Papageorgiou 2007, Zarco-Tejada et al. 2013). Here the chlorophyll content was evaluated biochemically to support the proteomic analysis of the photosynthesis and chloroplast biogenesis pathways in section 3.3.2. The total chlorophyll content (chl *a* and *b*) in the variant *Euglena gracilis* var. *saccharophila* strain was lower than that of the Z strain under both diurnal and dark cultivation, with more pronounced differences under the diurnal condition (Fig. 3).

The amount of chlorophyll remained at 20 to 30 $\mu g \cdot g^{-1}$ biomass during the first three d of cultivation, when both *Euglena* strains were grown in the absence of glucose thus relying on photosynthesis for survival. Chlorophyll levels declined sharply to less than 5 $\mu g \cdot g^{-1}$ after the addition of glucose (day 3+ to day 4) as both strains entered heterotrophic growth, but steadily recovered after day 5 when glucose was exhausted and the strains returned to photoautotrophic growth (Figs. 1 and 3).

Proteome changes between the dark and diurnal cultivation conditions

A total of 1647 non-redundant proteins could be reproducibly identified and quantified from the *Euglena gracilis* Z and *E. gracilis* var. *saccharophila* strains on day 4 of the cultivation period. From these, the relative abundances of a subset of proteins involved in glycolysis, gluconeogenesis, paramylon synthesis and degradation, and chloroplast and plastid biosynthesis were analyzed from the strains grown under dark and diurnal growth conditions, based on their NSAF coefficient.

The NSAF coefficient is calculated based on a protein's spectral count and length, which allows comparison of the relative abundance of a protein against the entire pool of proteins at a given time point. This method does not require a reference (0 hour time point) for comparison of differentially expressed proteins and has been widely applied to various organisms as a reliable indicator of relative protein quantity (Voelckel et al. 2010, Neilson et al. 2011, Mirzaei et al. 2012, Monavarfeshani et al. 2013, George et al. 2015, Kaufman et al. 2015). To add to the stringency, fold change ratio (as calculated from NSAF coefficient values) of a protein was considered valid only when the protein of interest can be detected in all three replicates and that the NSAF distribution in the triplicates satisfy the t-test (*p*-value <0.05, t-value >2 or <-2).

Differentially expressed enzymes in the sugar metabolic pathways

The majority of enzymes involved in the sugar metabolism pathways - glycolysis, gluconeogenesis, paramylon synthesis and degradation – were identified from both strains under both cultivation conditions (dark and diurnal). Exceptions to this were glucan endo-β-D-1,3-glucosidase and phosphofructokinase-1 enzymes that were not found in the diurnal

cultures, and isoform II of β-D-1,3-glucan synthase and isoform VI of glucan endo-β-D-1,3-glucosidase that were not found in dark cultivations (Table 1).

Glycolysis and gluconeogenesis

The analysis of proteins involved in glycolysis and gluconeogenesis (Table 1) showed that phosphofructokinase 1 (PFK-1), a key regulator of glycolysis, was more than 7-fold higher in *Euglena gracilis* var. *saccharophila* than in the Z strain in the diurnal cultivation. PFK-1 catalyses the first reaction that commits glucose to the glycolytic pathway (Mor et al. 2011) and an overexpression of PFK-1 may help contributing to the faster metabolism of glucose in the *E. gracilis* var. *saccharophila* strain observed in the cultivation studies in both cultivation modes (Fig. 1B).

Most enzymes in the gluconeogenesis pathway were more abundant in the *Euglena gracilis* var. *saccharophila* strain, including chloroplastic fructose-1,6-bisphosphatase (FBP) with a two-fold increase during diurnal cultivation. The increased abundance of chloroplastic FBP was as expected as it plays a role in photosynthetic carbon metabolism (Ogawa et al. 2015a). The higher abundance of FBP may also facilitate paramylon accumulation in the *E. gracilis* var. *saccharophila* strain, as the product of FBP, fructose-6-phosphate, serves as a precursor to the paramylon synthesis pathway. Other enzymes that were more abundant in the *E. gracilis* var. *saccharophila* strain were phosphoenolpyruvate carboxykinase isoforms (PEPCKs) in the range of 1.6 to 2-fold increase (Table 1) during diurnal cultivation. The levels of cytosolic fructose-1,6-bisphosphatase and pyrophosphate-fructose 6-phosphate 1-phosphotransferase were not affected by the cultivation mode.

Paramylon synthesis and degradation

Enzymes (e.g., glucan endo-β-D-1,3-glucan synthases and glucan endo-β-D-1/3-glucosidases) involved in paramylon synthesis and degradation in the *Euglena gracilis* var. *saccharophila* strain were almost uniformly lower with a few remaining unchanged compared to the Z strain in both diurnal and dark cultivations (Table 1). For the *E. gracilis* var. *saccharophila* strain, this may be indicative of a transitional phase in its paramylon metabolism, as it coincides with the highest point of paramylon production at day 4, before a decline (Fig. 1B; diurnal and dark cultivations).

The overall higher expression level of the gluconeogenesis pathway, coupled with the lower expression level of paramylon synthesis/degradation in the *Euglena gracilis* var. *saccharophila* strain in the diurnal cultivation also signifies a shift from the heterotrophic growth (medium glucose-dependent) to photoautotrophic growth (light and CO_2-dependent). This metabolic shift was corroborated by the elevated levels of photoreceptor proteins and chlorophyll biosynthesis enzymes detected in the same culture (Table 2).

During dark cultivation, the levels of most enzymes were similar between the two strains. Some enzymes involved in the paramylon synthesis and gluconeogenesis pathways became more abundant (1.4 to 30-fold increase) in the Z strain (Table 1), which may have contributed to the accumulation of paramylon during dark cultivation (Fig. 1A; days 4-5). The enzymes that were overexpressed included the chloroplastic fructose-1,6-bisphosphatase, phosphoenolpyruvate carboxykinases, phosphoglucomutase, UDP–glucose pyrophosphorylase and a bi-functional enzyme capable of catalysing both the phosphoglucomutase and UDP–glucose pyrophosphorylase reactions. However, fluctuation of the enzyme levels in paramylon and glucose metabolic pathways did not result in a significant difference in maximal paramylon yields between the Z and *Euglena gracilis* var. *saccharophila* strains in the dark even though *E. gracilis* var. *saccharophila* accumulated paramylon faster (Fig. 2, A and B; days 4-5). This finding further reinforces the view that while differential protein expression leads to a different rate of paramylon accumulation between the strains studied, their maximal growth and paramylon yield are still regulated by similar physiological bottlenecks or rate limiting enzymes discussed below.

Differentially expressed enzymes in photosynthesis and chloroplast biogenesis pathways

Since photoinduced degradation is an important factor affecting the yields, the relationship between chlorophyll levels, chloroplast/photosynthesis-associated proteins and paramylon accumulation were also explored. A decrease in paramylon production was expected in the diurnal cultivations due to the fact that light may enhance the activity of enzymes responsible for paramylon degradation (Price 1990).

The overall abundance of chloroplast and plastid proteins identified from the diurnal cultivations in *Euglena gracilis* var. *saccharophila* and Z strains was similar and approximately two-fold higher than in their corresponding dark cultures. Some chlorophyll

biosynthesis-related proteins (e.g., plastid uroporphyrinogen synthase, chlorophyll synthase andtetrapyrrole-binding protein) were absent or in relatively low abundance in the dark comparing to their diurnal counterparts (Table 2). While the change in expression levels of the individual chlorophyll biosynthesis proteins is not proportionate to the drastic crash of chlorophyll levels in the dark cultivations (Fig. 3), it is possible that the low abundance, or absence, of multiple proteins in this pathway collectively formed a bottleneck for the biosynthesis and accumulation of chlorophylls.

Photoreceptor proteins in the *Euglena gracilis* var. *saccharophila* strain were generally more abundant than those in strain Z, especially during the diurnal cultivations (Table 2). For example, during diurnal growth, the phototropins and an adagio 2-like protein, both of which are light-oxygen-voltage (LOV) blue light receptors, were found to be up to 5-fold more abundant in the *E. gracilis* var. *saccharophila* strain (Table 2). Phototropins are responsible for mediating phototropism, light-induced stomatal opening, and chloroplast movement in response to changes in light intensity (Christie 2007). The adagio 2 proteins are involved in regulation of the chloroplast development and degradation in relation to circadian rhythm (Mittag et al. 2005, Zoltowski and Imaizumi 2014). Differently to earlier reports (Schwartzbach et al. 1975) where paramylon degradation were found to be regulated by photoreceptors, relatively higher levels of the above photoreceptors in the *E. gracilis* var. *saccharophila* strain did not readily result in lower paramylon accumulation. This suggests that more factors connecting the chloroplast development to paramylon degradation were involved.

The development of chloroplasts in *Euglena gracilis* is catabolite sensitive, and will be supressed if a utilizable carbon source is present in the medium (Monroy and Schwartzbach 1984). This mechanism was reflected by a crash of chlorophyll levels within 24 h of addition of glucose in both strains in both dark and diurnal cultivations and staying low until day 9 (Fig. 3). However, at the proteome level, some of the enzymes involved in the chlorophyll biosynthesis pathways, such as plastid uroporphyrinogen decarboxylase (UROD, 4 to 44 fold increase) and glutamate 1-semialdehyde 2,1-aminotransferases (GSA, 1.5 to 1.9 fold increase), were found to be higher in the *E. gracilis* var. *saccharophila* strain (Table 2). The abundance of URODs and GSAs did not seem to directly translate to significant increase in the chlorophyll levels, possibly gated by the downstream enzymes. For example, no

454 significant changes were observed in the expression of enzymes (e.g., Mg-protoporphyrin IX
455 chelatase, protochlorophyllide reductase, divinyl chlorophyllide *a* 8-vinyl-reductase and
456 chlorophyll synthase) involved in the chloroplast biosynthesis sub-pathway that converts
457 protoporphyrin IX into Mg-Protoporphyrin, protochlorophyllide, and ultimately chlorophyll *a*
458 and *b* (Terry and Smith 2013). Moreover, the level of tetrapyrrole-binding protein which
459 regulates chlorophyll synthesis and plastid-to-nucleus signal transduction (Brzezowski et al.
460 2014) was 50% of that in the *Z* strain. Taken together, although many of the enzymes in the
461 early steps of the chloroplastogenesis pathway were more abundant, the generation of
462 chloroplasts in *Euglena gracilis* may be controlled by the later steps that convert
463 protoporphyrin into chlorophylls.
464
465 Comparing to the diurnal cultivation, many chloroplast and plastid proteins were not found or
466 had fallen below the detection limit in the dark cultivations (grey-out rows; Table 2, both
467 strains), reflecting the corresponding low chlorophyll levels of <10 $\mu g \cdot g^{-1}$ biomass (Fig. 3).
468 Of the more abundant proteins in the *Euglena gracilis* var. *saccharophila* strain during dark
469 cultivation, plastid transketolase (TK) was found to be 68 times more abundant than in the Z
470 strain. The transketolase plays a strategic role in carbohydrate metabolism, catalyzing
471 reversible conversions of glyceraldehyde-3-phosphate and fructose-6-phosphate to xylulose-
472 5-phosphate and erythrose-4-phosphate, both of which are substrates for enzymes in the
473 glycolysis/gluconeogenesis pathways (Gontero et al. 2007). Moreover, TK was found to limit
474 the maximum rate of photosynthesis and growth through regulation of carbon allocation
475 (Henkes et al. 2001, Gontero et al. 2007, Khozaei et al. 2015). A look into the TK abundance
476 in the diurnal samples revealed that the enzyme levels were high in both conditions in *E.*
477 *gracilis* var. *saccharophila* (NSAF coefficient 1.711 in diurnal and 1.602 in dark).
478
479 The elevated levels of TK in *Euglena gracilis* var. *saccharophila* may thus be a contributing
480 factor to its better performance in paramylon production over the Z strain during diurnal
481 cultivation (Fig. 3) as the enzyme modulates photosynthesis and growth. However, further
482 investigation would be required to establish this relationship.
483
484 **Concluding remarks**
485 The *Euglena gracilis* var. *saccharophila* variant strain was established in this study as a
486 desirable producer of paramylon due to its ability to metabolize glucose and convert to
487 paramylon at a faster rate than the frequently used Z strain. Paramylon yield, up to 8.1 $g \cdot L^{-1}$

was achieved by *E. gracilis* var. *saccharophila* in the dark cultivations and the glucose-to-paramylon conversion rate of $Y_{par/glu}$ = 0.46±0.03 is the highest reported in the literature to date.

A total of 1647 non-redundant proteins were reproducibly identified and quantified from the *Euglena gracilis* var. *saccharophila* and the Z strains grown under dark and diurnal growth conditions and the relative abundances of a subset of proteins involved in glycolysis, gluconeogenesis, paramylon synthesis and degradation, and chloroplast and plastid biosynthesis were analyzed. A higher abundance of PFK-1, a key regulator of glycolysis; FBP, chloroplastic fructose-1,6-bisphosphatase; UROD, plastid uroporphyrinogen decarboxylase; GSA, glutamate 1-semialdehyde 2,1-aminotransferases and TK, plastid transketolase and photoreceptors may have contributed to the improved ability of *E. gracilis* var. *saccharophila* to produce and retain paramylon. The overall abundance of chloroplast and plastid proteins was associated with chlorophyll levels in both strains. It was also notable that many enzymes in the early chloroplastogenesis pathways were more abundant in *E. gracilis* var. *saccharophila* strain.

The findings in this study have pointed out defined targets for the improvement of paramylon production in *Euglena gracilis.* For example, overexpression of the chloroplastic FBP may enhance paramylon accumulation as the protein serves as a precursor for the paramylon synthesis pathway, and overepxression of PFK-1 may enhance sugar metabolism, which could result in a quicker turnaround time in cell growth and paramylon production.

Acknowledgements

This research and the authors are supported by ARC-ITTC funding, Project IC130100009.

Authors contributions and competing interests

AS, GH, HN and JT were involved in the conception and design of the study. Data acquisition, analysis and interpretations were carried out by AS and MH. Manuscript was drafted by AS and revised by HN, GH and JT to its final shape. The manuscript was read and approved by all authors.

Conflict of Interest Disclosure

The authors declare no conflict of interest directly relevant or indirectly related to this work.

References

Barsanti, L., Vismara, R., Passarelli, V. & Gualtieri, P. 2001. Paramylon (β-1,3-glucan) content in wild type and WZSL mutant of *Euglena gracilis*. Effects of growth conditions. *J. Appl. Phycol.* 13:59-65.

Brzezowski, P., Schlicke, H., Richter, A., Dent, R. M., Niyogi, K. K. & Grimm, B. 2014. The GUN4 protein plays a regulatory role in tetrapyrrole biosynthesis and chloroplast-to-nucleus signalling in *Chlamydomonas reinhardtii*. *Plant J.* 79:285-98.

Christie, J. M. 2007. Phototropin blue-light receptors. *Annu. Rev. Plant Biol.* 58:21-45.

Craig, R. & Beavis, R. C. 2003. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun. Mass Spectrom.* 17:2310-6.

Craig, R. & Beavis, R. C. 2004. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20:1466-7.

Ebenezer, T. E., Carrington, M., Lebert, M., Kelly, S. & Field, M. C. 2017. *Euglena gracilis* genome and transcriptome: organelles, nuclear genome assembly strategies and initial features. *In* Schwartzbach, S. D. & Shigeoka, S. [Eds.] *Euglena: Biochemistry, Cell and Molecular Biology*. Springer International Publishing, Cham, Switzerland, pp. 125-40.

George, I. S., Pascovici, D., Mirzaei, M. & Haynes, P. A. 2015. Quantitative proteomic analysis of cabernet sauvignon grape cells exposed to thermal stresses reveals alterations in sugar and phenylpropanoid metabolism. *Proteomics* 15:3048-60.

548

549 Gontero, B., Avilan, L. & Lebreton, S. 2007. Control of carbon fixation in chloroplasts. *In*
550 Plaxton, W.C. & McManus, M.T. [Eds.] *Annual Plant Reviews Volume 22: Control of*
551 *Primary Metabolism in Plants.* Blackwell Publishing Ltd, Hoboken, NJ, USA, pp. 187-218.

552

553 Grimm, P., Risse, J. M., Cholewa, D., Müller, J. M., Beshay, U., Friehs, K. & Flaschel, E.
554 2015. Applicability of *Euglena gracilis* for biorefineries demonstrated by the production of α-
555 tocopherol and paramylon followed by anaerobic digestion. *J. Biotechnol.* 215:72-79.

556

557 Hasan, M., Sun, A., Mirzaei, M., Te'o, J., Hobba, G., Sunna, A. & Nevalainen, H. 2017.
558 A comprehensive assessment of the biosynthetic pathways of ascorbate, α-tocopherol and
559 free amino acids in *Euglena gracilis var. saccharophila. Algal Res.* 27:140-151.

560

561 Henkes, S., Sonnewald, U., Badur, R., Flachmann, R. & Stitt, M. 2001. A small decrease of
562 plastid transketolase activity in antisense tobacco transformants has dramatic effects on
563 photosynthesis and phenylpropanoid metabolism. *Plant Cell* 13:535-51.

564

565 Ivušić, F. & Šantek, B. 2015. Optimization of complex medium composition for
566 heterotrophic cultivation of *Euglena gracilis* and paramylon production. *Bioproc. Biosyst.*
567 *Eng.* 38:1103-12.

568

569 Jeffrey, S. & Humphrey, G. 1975. New spectrophotometric equations for determining
570 chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. *Biochem.*
571 *Physiol. Pflanzen.* 167:191-4.

572

573 Kaufman, K. L., Jenkins, Y., Alomari, M., Mirzaei, M., Best, O. G., Pascovici, D., Mactier,
574 S., Mulligan, S. P., Haynes, P. A. & Christopherson, R. I. 2015. The Hsp90 inhibitor SNX-
575 7081 is synergistic with fludarabine nucleoside via DNA damage and repair mechanisms in
576 human, p53-negative chronic lymphocytic leukemia. *Oncotarget* 6:40981-97.

577

578 Khozaei, M., Fisk, S., Lawson, T., Gibon, Y., Sulpice, R., Stitt, M., Lefebvre, S. C. & Raines,
579 C. A. 2015. Overexpression of plastid transketolase in tobacco results in a thiamine
580 auxotrophic phenotype. *Plant Cell* 27:432-47.

581

582 Mahapatra, D., Chanakya, H. N. & Ramachandra, T. V. 2013. *Euglena sp.* as a suitable
583 source of lipids for potential use as biofuel and sustainable wastewater treatment. *J Appl.*
584 *Phycol.* 25:855-65.

585

586 Mirzaei, M., Soltani, N., Sarhadi, E., George, I. S., Neilson, K. A., Pascovici, D., Shahbazian,
587 S., Haynes, P. A., Atwell, B. J. & Salekdeh, G. H. 2014. Manipulating root water supply
588 elicits major shifts in the shoot proteome. *J. Proteome Res.* 13:517-26.

589

590 Mirzaei, M., Soltani, N., Sarhadi, E., Pascovici, D., Keighley, T., Salekdeh, G. H., Haynes, P.
591 A. & Atwell, B. J. 2012. Shotgun proteomic analysis of long-distance drought signaling in
592 rice roots. *J. Proteome Res.* 11:348-58.

593

594 Mittag, M., Kiaulehn, S. & Johnson, C. H. 2005. The circadian clock in *Chlamydomonas*
595 *reinhardtii*. What is it for? What is it similar to? *Plant Physiol.* 137:399-409.

596

597 Monavarfeshani, A., Mirzaei, M., Sarhadi, E., Amirkhani, A., Khayam Nekouei, M., Haynes,
598 P. A., Mardi, M. & Salekdeh, G. H. 2013. Shotgun proteomic analysis of the Mexican Lime
599 tree infected with "Candidatus *Phytoplasma aurantifolia*". *J. Proteome Res.* 12:785-95.

600

601 Monroy, A. F. & Schwartzbach, S. D. 1984. Catabolite repression of chloroplast development
602 in *Euglena*. *Proc. Natl. Acad. Sci. USA* 81:2786-90.

603

604 Mor, I., Cheung, E. C. & Vousden, K. H. 2011. Control of glycolysis through regulation of
605 PFK1: old friends and recent additions. *Cold Spring Harb. Symp. Quant. Biol.* 76:211-6.

606

607 Nagaraj, N., Reese, J. C., Kirkham, M. B., Kofoid, K., Campbell, L. R. & Loughin, T. M.
608 2002. Relationship between chlorophyll loss and photosynthetic rate in greenbug (Homoptera,
609 Aphididae) damaged sorghum. *J. Kans. Entomol. Soc.* 75:101-09.

610

611 Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee,
612 A., van Sluyter, S. C. & Haynes, P. A. 2011. Less label, more free: Approaches in label-free
613 quantitative mass spectrometry. *Proteomics* 11:535-53.

614

615 O'Neill, E. C., Trick, M., Hill, L., Rejzek, M., Dusi, R. G., Hamilton, C. J., Zimba, P. V.,
616 Henrissat, B. & Field, R. A. 2015. The transcriptome of *Euglena gracilis* reveals unexpected
617 metabolic capabilities for carbohydrate and natural product biochemistry. *Mol. Biosyst.*
618 11:2808-20.

619

620 Ogawa, T., Kimura, A., Sakuyama, H., Tamoi, M., Ishikawa, T. & Shigeoka, S. 2015a.
621 Characterization and physiological role of two types of chloroplastic fructose-1,6-
622 bisphosphatases in *Euglena gracilis*. *Arch. Biochem. Biophys.* 575:61-8.

623

624 Ogawa, T., Tamoi, M., Kimura, A., Mine, A., Sakuyama, H., Yoshida, E., Maruta, T., Suzuki,
625 K., Ishikawa, T. & Shigeoka, S. 2015b. Enhancement of photosynthetic capacity in *Euglena*
626 *gracilis* by expression of cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase leads
627 to increases in biomass and wax ester production. *Biotechnol. Biofuels* 8:80.

628

629 Papageorgiou, G. C. & Govindjee. 2007. *Chlorophyll a Fluorescence: A Signature of*
630 *Photosynthesis*. Springer Dordrecht, Netherlands, 820 pp.

631

632 Price, C. A. 1990. The Biology of *Euglena*. *In* Buetow, D.E. [Ed.] *Vol. 4, Subcellular*
633 *Biochemistry and Molecular Biology*. Academic Press, San Diego, CA, 528 pp.

634

635 Rodriguez-Zavala, J. S., Ortiz-Cruz, M. A., Mendoza-Hernandez, G. & Moreno-Sanchez, R.
636 2010. Increased synthesis of alpha-tocopherol, paramylon and tyrosine by *Euglena gracilis*
637 under conditions of high biomass production. *J. Appl. Microbiol.* 109:2160-72.

Rodriguez-Zavala, J. S., Ortiz-Cruz, M. A. & Moreno-Sanchez, R. 2006. Characterization of an aldehyde dehydrogenase from *Euglena gracilis*. *J. Eukaryot. Microbiol.* 53:36-42.

Šantek, B. i., Felski, M., Friehs, K., Lotz, M. & Flaschel, E. 2009. Production of paramylon, a β-1,3-glucan, by heterotrophic cultivation of *Euglena gracilis* on a synthetic medium. *Eng. Life. Sci*. 9:23-28.

Šantek, B. i., Friehs, K., Lotz, M. & Flaschel, E. 2012. Production of paramylon, a β-1,3-glucan, by heterotrophic growth of *Euglena gracilis* on potato liquor in fed-batch and repeated-batch mode of cultivation. *Eng. Life. Sci.* 12:89-94.

Schwartzbach, S. D., Schiff, J. A. & Goldstein, N. H. 1975. Events surrounding the early development of *Euglena* chloroplasts: v. control of paramylum degradation. *Plant Physiol.* 56:313-7.

Tanaka, Y., Ogawa, T., Maruta, T., Yoshida, Y., Arakawa, K. & Ishikawa, T. 2017. Glucan synthase-like 2 is indispensable for paramylon synthesis in *Euglena gracilis*. *FEBS Lett.* 591:1360-70.

Terry, M. & Smith, A. 2013. A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. *Front. Plant Sci.* 4:14.

Voelckel, C., Mirzaei, M., Reichelt, M., Luo, Z., Pascovici, D., Heenan, P. B., Schmidt, S., Janssen, B., Haynes, P. A. & Lockhart, P. J. 2010. Transcript and protein profiling identify candidate gene sets of potential adaptive significance in New Zealand Pachycladon. *BMC Evol. Biol.* 10:151.

Wang, L., Behr, S. R., Newman, R. K. & Newman, C. W. 1997. Comparative cholesterol-lowering effects of barley β-glucan and barley oil in Golden Syrian hamsters. *Nutr. Res.* 17:77-88.

669

670 Wood, P. J. 1994. Evaluation of oat bran as a soluble fibre source. Characterization of oat β-
671 glucan and its effects on glycaemic response. *Carbohydr. Polymers* 25:331-36.

672

673 Yoshida, Y., Tomiyama, T., Maruta, T., Tomita, M., Ishikawa, T. & Arakawa, K. 2016. De
674 novo assembly and comparative transcriptome analysis of *Euglena gracilis* in response to
675 anaerobic conditions. *BMC Genomics* 17:182.

676

677 Zarco-Tejada, P. J., Catalina, A., González, M. R. & Martín, P. 2013. Relationships between
678 net photosynthesis and steady-state chlorophyll fluorescence retrieved from airborne
679 hyperspectral imagery. *Remote Sens. Environ.* 136:247-58.

680

681 Zoltowski, B. D. & Imaizumi, T. 2014. Structure and function of the ZTL/FKF1/LKP2 group
682 proteins in *Arabidopsis*. *Enzymes* 35:213-39.

683

684

685 **Tables and Figures**

686 **Figure 1.** Changes in the biomass and paramylon production by *Euglena gracilis* Z and
687 *Euglena gracilis* var. *saccharophila* strains during the 9-day cultivation. The day 3 samples
688 were collected before the addition of glucose and day 3+ samples were collected after the
689 addition of glucose. Days 4-9 represent time points after glucose addition. Yeast extract (1YE)
690 was provided as a nitrogen source at the start of culturing. A. *Euglena gracilis* Z cultures. B.
691 *Euglena gracilis* var. *saccharophila* cultures. ○: paramylon under diurnal cycle, ●: paramylon
692 under dark cultivation, □: biomass under diurnal cycle, ■: biomass under dark cultivation.
693 Dotted grey lines represent residual glucose in the medium, △: residual glucose under
694 diurnal cycle, ▲: residual glucose under dark cultivation.

695

696 **Figure 2.** The maximum volumetric yield (A) and paramylon mass fraction of paramylon (B)
697 obtained in the *Euglena gracilis* Z and the *Euglena gracilis* var. *saccharophila* cultures.
698 Statistical significance indicated by * $p<0.05$, t = 3.165, 95% confidence interval 0.1622 to
699 2.48.

700

701 **Figure 3.** Chlorophyll content (μg chlorophyll · g biomass^{-1}) in the *Euglena gracilis* Z and
702 the *Euglena gracilis* var. *saccharophila* cells cultivated under diurnal and dark conditions. ○:
703 *Euglena gracilis* Z under diurnal cycle, ●: *Euglena gracilis* Z under dark cycle, □: *Euglena*
704 *gracilis* var. *saccharophila* under diurnal cycle, ■: *Euglena gracilis* var. *saccharophila* under
705 the dark cycle.

706

Table 1. Level changes of the predicted proteins that are involved in the glycolysis and gluconeogenesis pathways from the *Euglena gracilis* var. *saccharophila* and *Euglena gracilis Z* strains during the mid-log phase of growth (day 4, 24 h post glucose addition). Greyscale heat map: proteins were color-coded if the level differences between the two strains were statistically significant, dark-grey indicates higher abundance while light-grey indicates lower abundance. Uncolored (blank) cells: proteins were present in these cultivation conditions, but the level differences between strains were insignificant. Greyed-out cells: proteins were absent or below level of detection in these cultivation conditions. NSAF coefficient values in each cell represents relative abundance of the protein across different cultivation conditions. Fold-changes represent the relative protein levels of the var. *saccharophila* culture compared to that in the *Z* strain under the same cultivation condition. Fold change ratios were not displayed for proteins that were absent or with level differences that were statistically insignificant. Up- and Down-ward arrows indicate proteins that were differentially expressed in the var. *saccharophila* strain. Values under the "Fold change – Diurnal to Dark" columns indicate the change in protein abundance in the dark cultures as compared to the Diurnal cultures. Statistical significance here defined as (*p*-value <0.05, t >2 or <-2).

Diurnal cultivation			Dark cultivation			Fold-change Diurnal to Dark				
E. gracilis var. *saccharophila*	Fold-change	*E. gracilis Z*	*E. gracilis* var. *saccharophila*	Fold-change	*E. gracilis Z*	*E. gracilis* var. *saccharophila*	*E. gracilis Z*	EC Number	Description	Accession number
Glycolysis										
0.612	↓ 2.7	1.647	1.062	-	0.678	-	-	2.7.1.2	Glucokinase; GK	T0RE29
0.981	-	0.812	1.570	↑ 2.4	0.637	-	-	5.3.1.9	Glucose-6-phosphate isomerase; GPI	R0GDQ1
1.232	↑ 7.4	0.166	1.568	-	1.034	-	↑ 6.2	2.7.1.11	Phosphofructokinase 1; PFK-1	B4CY52
-	-	-	0.709	-	2.372	-	-	2.7.1.11	ATP-dependent 6-phosphofructokinase	C9ZKZ6
1.043	↓ 1.2	1.230	0.830	-	0.897	-	-	4.1.2.13	Fructose-bisphosphate aldolase	Q42728
1.232	↑ 1.8	0.668	1.411	↑ 2.0	0.689	-	-	2.7.1.40	Pyruvate Kinase	B3RNS0
0.801	-	0.934	1.044	-	1.222	-	-	1.2.4.1	Pyruvate dehydrogenase complex, component E1	Q6KCM1
1.085	↑ 1.4	0.753	1.126	-	1.036	-	-	2.3.1.12	Pyruvate dehydrogenase complex, component E2	Q6KCM0
0.950	↓ 1.6	1.526	1.042	↑ 2.1	0.481	-	↓ 3.2	1.8.1.4	Pyruvate dehydrogenase complex, component E3	Q6KCL9
Glycolysis / Gluconeogenesis, reversible										
0.253	↓ 4.9	1.238	0.315	↓ 7.1	2.195	-	-	5.3.1.1	Triosephosphate isomerase	Q6VEG5
0.449	↓ 2.9	1.301	0.771	-	1.479	-	-	-	Triosephosphate isomerase	A0A0S3IUI5
0.911	↑ 1.9	0.492	1.256	-	1.340	-	↑ 2.7	1.2.1.12	Glyceraldehyde-3-phosphate dehydrogenase; GAPDH	Q43311
1.072	↑ 1.6	0.655	1.166	-	1.107	-	-	2.7.2.3	Phosphoglycerate kinase; PGK	Q66PT2
1.382	↓ 1.2	1.603	0.569	-	0.446	↓ 2.4	↓ 3.6	2.7.2.3	Phosphoglycerate kinase; PGK	Q66PT3
1.147	↑ 1.3	0.856	1.578	-	0.420	-	-	1.2.1.9	Glyceraldehyde-3-phosphate dehydrogenase (NADP+), GAPN	R1BKM5
1.011	-	0.760	1.238	-	0.991	-	-	5.4.2.11	Phosphoglycerate mutase	X6P0R7
1.125	-	0.961	1.064	-	0.849	-	-	4.2.1.11	Enolase	Q9LEK7
0.963	-	0.757	0.878	↓ 1.6	1.402	-	-	4.2.1.11	Enolase	Q9LEK7
Gluconeogenesis										
1.120	-	0.892	1.131	-	0.856	-	-	3.1.3.11	Fructose-1,6-bisphosphatase , cytosolic, EgFBPaseIII	A0A0U4MTX7
2.115	↑ 2.0	1.050	0.160	↓ 4.2	0.675	↓ 13.2	-	3.1.3.11	Fructose-1,6-bisphosphatase, chloroplastic, FBP	A3QSS7
0.861	-	0.932	1.027	-	1.180	-	-	2.7.1.90	Pyrophosphate-fructose 6-phosphate 1-phosphotransferase	D7FV23
1.374	↑ 1.6	0.858	0.849	-	0.918	-	-	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform I	L8GQL5
0.931	↑ 2.1	0.441	1.041	-	1.588	-	↑ 3.6	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform II	L8GQL5
0.716	↑ 1.4	0.495	0.860	↓ 2.3	1.928	-	↑ 3.9	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform III	L8GQL5
0.750	↑ 2.0	0.373	1.202	↓ 1.4	1.675	-	↑ 4.5	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform IV	L8GQL5
1.056	↑ 2.2	0.476	1.247	-	1.222	-	↑ 2.6	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform V	L8GQL5
0.909	↑ 1.6	0.612	1.221	-	1.258	-	-	4.1.1.32	Phosphoenolpyruvate carboxykinase; PEPCK; Isoform VI	L8GQL5
Paramylon synthesis										
0.403	↓ 3.1	1.239	0.531	↓ 3.5	1.827	-	-	5.4.2.2	Phosphoglucomutase	T0QKN9
0.169	↓ 7.0	1.180	0.801	↓ 2.3	1.850	↑ 4.7	-	5.4.2.2/ 2.7.7.9	Phosphoglucomutase /UDP–glucose pyrophosphorylase bifunctional	W7T5H8 A0A0N7L4L5
1.074	-	0.904	1.093	-	0.929	-	-	2.7.7.9	UDP–glucose pyrophosphorylase; Isoform I	B9XHR8
0.421	-	0.725	0.091	↓ 30.8	2.764	↓ 4.6	↑ 3.8	2.7.7.9	UDP–glucose pyrophosphorylase; Isoform II	I0YVS3
1.192	-	1.066	1.038	-	0.705	-	-	2.4.1.34	β-D-1,3-glucan synthase (GT48 family); Isoform I	A0A0S4KKB3
0.919	-	2.517	-	-	-	-	-	2.4.1.34	β-D-1,3-glucan synthase (GT48 family); Isoform II	A0A0S4IWM0
Parmaylon degradation										
-	-	-	2.704	↑ 7.5	0.354	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH64 family)	A0A0W0VYE4
0.463	-	1.082	1.160	-	1.295	↑ 2.5	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform I	A0A0E4B7Q1
0.687	↓ 2.1	1.429	1.142	-	0.742	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform II	A0A0E4B7Q1
0.564	-	1.220	1.126	-	1.090	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform III	A9ER91
0.540	↓ 2.9	1.575	0.966	-	0.919	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform IV	A0A150RBA9
0.802	-	1.153	1.219	-	0.826	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform V	A0A0E4B9H2
0.521	↓ 3.4	1.769	-	-	-	-	-	3.2.1.58/39	Glucan endo-β-D-1,3-glucosidase (GH81 family); Isoform VI	A0A150GTC4
0.341	-	0.786	0.940	-	1.932	-	-	3.2.1.21	Exo-β-D-1,3-glucanase (GH2 family); Isoform I	A4BD51
0.418	-	1.231	0.643	-	1.708	-	-	3.2.1.21	Exo-β-D-1,3-glucanase (GH5 family)	A0A0S8KM75

719

720 **Table 2.** Level changes of the predicted proteins associated with photoreceptors and proteins found in the chloroplasts and plastids. Data shown
721 are from the *Euglena gracilis* var. *saccharophila* and *Euglena gracilis Z* cultures grown in either diurnal or dark conditions during the mid-log
722 phase of growth (day 4, 24 h post glucose addition). Greyscale heat map: proteins were color-coded if the level differences between the two
723 strains were significant. Numerical values in the color-coded cells is the NSAF coefficient that represents relative abundance of the protein
724 comparing to other cultivation conditions, dark-grey indicates more abundant while light-grey indicates less abundant proteins. Uncolored (blank)
725 cells: proteins were present in these cultivation conditions, but the level differences between strains were insignificant. Greyed-out cells: proteins
726 were absent or could not be identified in these cultivation conditions. Fold-changes represent the protein levels of the *E. gracilis* var.
727 *saccharophila* cultures compared to that in the *Z* strain cultures under the same cultivation condition. The values under the "Fold change –
728 Diurnal to Dark" columns indicate the change in protein abundance in the dark cultures as compared to the Diurnal cultures. Up- and down-
729 arrows indicated proteins that were differentially expressed, respectively, in the *E. gracilis* var. *saccharophila* strain. Fold change ratios were not
730 calculated for proteins that were absent or with level differences that were statistically insignificant. Statistical significance here defined as (*p*-
731 value <0.05, t >2 or <-2).

Diurnal cultivation			Dark cultivation			Fold-change Diurnal to Dark				
E. gracilis var. *saccharophila*	Fold-change	*E. gracilis Z*	*E. gracilis* var. *saccharophila*	Fold-change	*E. gracilis Z*	*E. gracilis* var. *saccharophila*	*E. gracilis Z*	E-value	Description	Accession number
Photoreceptors										
1.608	↑ 4.4	0.363	-	-	-	-	-	3.0E-16	Adagio protein 2	Q8W420
1.271	↑ 4.0	0.319	1.251	→ 1.1	1.159	-	-	4.4E-06	Neochrome	Q401Q6
1.263	↑ 1.9	0.672	1.282	↑ 1.6	0.783	-	-	2.4E-07	Neochrome (fragment)	A0A059UB02
1.531	↑ 5.4	0.283	-	-	-	-	-	8.5E-12	Phototropin	A0A059UB53
1.219	↑ 1.7	0.736	1.076	→ 1.1	0.969	-	-	3.1E-18	Phototropin	A0A059UB72
0.777	↓ 2.0	1.533	1.103	↑ 1.9	0.587	-	-	5.9E-25	UVB-receptor	I0YXW7
Chlorophyll biosnhthesis										
1.020	↓ 1.6	1.615	0.452	↓ 2.0	0.912	↓ 2.3	↓ 1.8	0.0E+00	Plastid glutamyl tRNA synthetase	A0A1L3A2Y0
1.108	↑ 1.5	0.737	1.074	-	1.081	-		1.2E-26	Plastid glutamate 1-semialdehyde 2,1-aminotransferase	A0A1L3A3C6
1.098	↑ 1.9	0.581	1.320	↑ 1.3	1.001	-	-	8.5E-18	Plastid glutamate 1-semialdehyde 2,1-aminotransferase	A0A1L3A3C7
0.784	-	1.871	-	-	-	-	-	2.2E-06	Plastid glutamate 1-semialdehyde 2,1-aminotransferase	A0A1L3A3C9
1.377	-	2.037	-	-	-	-	-	0.0E+00	Plastid uroporphyrinogen synthase	A0A1L3A3L8
2.610	↑ 4.8	0.538	-	-	-	-	-	0.0E+00	Plastid uroporphyrinogen decarboxylase isoform 1	A0A1L3A3Y5
2.171	↑ 44.0	0.049	0.289	↓ 5.2	1.490	↓ 7.5	↑ 30.2	1.9E-17	Plastid uroporphyrinogen decarboxylase isoform 2	A0A1L3A403
1.513	-	1.149	0.609	-	0.729	↓ 2.5	↓ 1.6	0.0E+00	Plastid uroporphyrinogen decarboxylase isoform 3	A0A1L3A425
1.710	→ 1.1	1.526	-	-	-	-	-	3.5E-118	Plastid coproporphyrinogen oxidase isoform 3	A0A1L3A3U9
1.501	-	1.287	0.459	↓ 1.7	0.753	↓ 3.3	↓ 1.7	0.0E+00	Plastid protoporphyrinogen oxidase	A0A1L3A3P7
1.166	-	1.325	0.403	↓ 2.8	1.106	↓ 2.9	-	1.7E-60	Mg-protoporphyrin IX chelatase	A0A0G3FB04
1.482	-	0.678	-	-	-	-	-	1.4E-25	Protochlorophyllide reductase	O48741
1.013	-	0.625	1.197	-	1.165	-	-	4.6E-28	Protochlorophyllide reductase	Q8W3D9
0.985	-	0.372	2.314	↑ 7.0	0.329	-	-	3.5E-20	Light-dependent protochlorophyllide reductase	O66148
1.724	-	0.951	1.122	↑ 5.5	0.203	-	-	5.3E-146	Divinyl chlorophyllide *a* 8-vinyl-reductase, chloroplastic	J9RYI6
2.432	-	0.642	-	-	-	-	-	4.2E-141	Chlorophyll Synthase	Q5W6H5
1.062	↓ 1.9	2.013	-	-	-	-	-	3.8E-35	Tetrapyrrole-binding protein, chloroplastic	Q9LX31
Other chloroplast and plastid proteins										
1.819	↑ 1.6	1.172	0.259	↓ 2.9	0.751	↓ 7.0	↓ 1.6	1.7E-12	Elongation factor Tu, plastid	P14634
1.236	↑ 5.2	0.236	0.157	↓ 15.3	2.371	↓ 7.9	↑ 10.0	9.8E-38	DNA-directed RNA polymerase subunit beta	P23580
1.271	↑ 2.2	0.576	1.348	↑ 1.7	0.805	-	-	2.8E-06	6-phosphogluconate dehydrogenase, decarboxylating	B2NIV9
1.295	↑ 2.2	0.601	1.078	→ 1.0	1.025	-	↑ 1.7	2.2E-04	30S ribosomal protein S8, chloroplastic	P21508
1.234	↑ 1.6	0.772	1.382	↑ 2.2	0.612	-	-	5.8E-13	Elongation factor Tu, chloroplastic	M1EV34
1.021	↑ 1.5	0.683	1.480	↑ 1.8	0.816	-	-	1.5E-164	Elongation factor Tu, chloroplastic	A0A0G3FBA4
0.702	↓ 1.2	0.875	1.104	-	1.320	↑ 1.6	-	4.5E-32	ATP synthase subunit alpha, chloroplastic	A0A1B0UKY3
0.721	↓ 1.4	1.035	0.892	↓ 1.5	1.352	-	-	2.0E-26	ATP synthase subunit beta, chloroplastic	A0A1B0UKY9
0.874	↓ 2.7	2.367	0.111	↓ 5.9	0.648	↓ 7.9	↓ 3.7	3.7E-15	Oxygen-evolving enhancer protein 2	P83687
0.313	↓ 9.6	3.005	0.042	↓ 15.5	0.640	↓ 7.5	↓ 4.7	0.0E+00	ADP,ATP carrier protein	E6Y2N7
1.711	-	0.664	1.602	↑ 68.0	0.023	-	↓ 28.5	1.7E-161	Plastid transketolase	A6YAZ5
0.904	-	1.015	1.419	↑ 2.1	0.662	-	-	8.0E-37	Translation initiation factor IF-2, chloroplastic	Q9XEK9
1.195	-	1.095	0.403	↓ 3.3	1.308	↓ 3.0	-	1.3E-21	Chloroplast light-harvesting complex I protein Lhca2	A4QPI4
1.335	-	2.073	0.107	↓ 4.6	0.485	↓ 12.5	↓ 4.3	8.0E-71	Chloroplast photosystem II protein M	D8VEQ8
0.390	-	1.018	0.382	↓ 5.9	2.210	-	↑ 2.2	4.2E-19	DNA-directed RNA polymerase subunit	M1EVD2
2.430	↑ 9.8	0.247	-	-	-	-	-	4.5E-14	Ferredoxin	P22341
2.224	↑ 3.8	0.592	-	-	-	-	-	8.0E-13	50S ribosomal protein L12, chloroplastic	M1EWE9
1.357	↓ 1.9	2.611	-	-	-	-	-	0.0E+00	Chloroplast light-harvesting complex II protein	A8HPF9
1.280	↓ 1.9	2.488	-	-	-	-	-	0.0E+00	Cytochrome f, chloroplastic	Q8GZR2
0.197	↓ 19.2	3.784	-	-	-	-	-	0.0E+00	Chloroplast light-harvesting complex I protein	A8HPC6
-	-	-	0.330	↓ 9.2	2.992	-	-	3.2E-06	Chloroplast enolase	Q9LEK6
-	-	-	0.296	↓ 8.6	2.525	-	-	7.3E-04	Photosystem II reaction center protein Z (PSII-Z)	M1FT16
1.048	-	0.747	1.873	-	0.333	-	-	0.0E+00	Porin-like protein	Q9FPM7
0.594	-	1.152	1.441	-	0.813	-	-	2.0E-45	50S ribosomal protein L2, chloroplastic	P19165
0.996	-	0.880	1.288	-	0.835	-	-	2.0E-07	30S ribosomal protein S9, plastid	P58135
0.812	-	0.576	1.380	-	1.232	-	-	7.9E-26	DNA-directed RNA polymerase subunit beta	A0A0G3VQU0
2.149	-	0.949	-	-	-	-	-	1.4E-04	Chloroplast light-harvesting complex I protein	A8HPD3
0.555	-	3.011	-	-	-	-	-	5.1E-151	Chloroplast light-harvesting complex II protein Lhcbm4	A4QPI0
1.744	-	2.175	-	-	-	-	-	0.0E+00	Light harvesting chlorophyll a /b binding protein of PSII	Q39725
1.887	-	0.768	-	-	-	-	-	8.9E-166	Chloroplast Phosphoglycerate kinase	Q66PT3
1.250	-	0.907	-	-	-	-	-	2.3E-177	Plastid ribose-5-phosphate isomerase	A0A0S3IU33
1.447	-	0.987	-	-	-	-	-	1.4E-25	Plastid fructose 1,6-bisphosphatase isoform 1	A0A0S3IUC8