UPPSALA
UNIVERSITET

The Direct Photobiological Conversion of CO_2 into C(1) Compounds

Phosphoenolpyruvate carboxylase

Malate dehydrogenase

Claudia Durall de la Fuente

Degree project in Applied Biotechnology, Master of Science (2 years), 2013
Examensarbete i 45 hp till masterexamen, 2013 Biology Education Centre and Department of Photochemistry and Molecular Science, Ångstrom, Uppsala University.
Supervisor: Peter Lindblad and Elias Englund

Contents

Abstract

Using bioinformatics as a tool, Bar-Even *et al.* (2010) created two putative non-native synthetic carbon fixation pathways called MOG pathways. After a careful analysis of the MOG pathways, it was observed that the C4 Glyoxylate Cycle/Lactate option is the pathway and enzymes with most available information and therefore selected for further work. In this study, the two first genes (MDH and PEPc) of the selected MOG pathway were introduced and overexpressed in *Synechocystis* PCC 6803. Both genes are present as a single copy gene in the *Synechocystis* genome. To begin with, an extra copy of the each native gene was inserted under the control of a strong promoter. Two transformants were created which each contain a native gene under the control of the native promoter and an additional copy of the native gene under the control of a strong promoter. The overexpression of the genes were examined at a transcriptional (MDH and PEPc) and translational level (PEPc) using Semiquantitative RT-PCR technique and Western Blot, respectively. The amount of transcript of the genes encoding PEPc and MDH gene were higher in the transformants compared to the wild type. Also, in the case of the PEPc transformant the cells contained significantly higher level of the protein. Thus, the additional copies that were inserted into *Syenchocystis* genome were transcribed and in the case of the PEPc transformant, it led to more protein. Further experiments are required in order to examine if the overexpressed proteins are active and if the catalysis of the reactions are more efficient.

Aim of the project:

The aim of this study is the overexpression of two of the eleven enzymes involved in the MOG pathway (Lactate option) into the cyanobacterium *Synechocystis* PCC 6803, Phosphoenolpyruvate carboxylase (PEPc) and Malate dehydrogenase (MDH). PEPc is the primary carbon incorporative enzyme in the pathway and MDH is the enzyme converting the product of PEPc into malate. The overexpression of the introduced genes is examined at transcript and protein level using Semiquantitative RT-PCR and Western Blot, respectively.

Introduction

It has been proven that petroleum reserves will be exhausted in a few years. This is the reason why researchers are trying to find alternatives in order to supply the high demand of energy needed (Sakata and Kawai 1981 and Adachi *et al*. 1994). Also, the greenhouse effect is one of the main environmental troubles that man is trying to solve. The level of greenhouse gases has increased dramatically during the last two decades (Solomon *et al.* 2007). This will increase the global temperature leading to environmental consequences.

Researchers have tried to develop on alternative energies like wind, hydraulic or solar power. However, the efficiency of these alternative methods is not good enough to supply all the energy demanded (Dresselhaus and Thomas 2001). Other researches have used genetically modified plants to produce biofuels. However, several negative impacts have been identified and the use of those is under controversy (Rosgaard *et al.* 2012).

It seems that a unique solution to replace the current fuel is not viable. A combination of different technologies may supply all the energy required. Interestingly, several groups of researchers have worked with photosynthetic microorganisms, such as blue-green algae or cyanobacteria with the focus to use them as photosynthetic factories for the direct production of solar fuels (Lindblad *et al.* 2012).

It is assumed that cyanobacteria are the oldest organisms on earth (Dvornyk *et al.* 2003). These organisms are photosynthetic and it is believed they were responsible for raising the O_2 levels in the atmosphere when the earth was anoxic, 2.3 billion years ago (Kasting *et al.* 2002). These photosynthetic bacteria are broadly distributed in different habitats such as: sea, oceans, freshwater and extreme environments (Tomitani *et al.* 2006). Also, several strains possess nitrogenase enzyme, thus they are able to fix N_2. This is an important feature because they are the main organisms fixing N_2 on the earth (Soltani *et al.* 2007).

Since cyanobacteria perform oxygenic photosynthesis, they grow under light, with inorganic nutrients, water and CO_2 (Coleman and Colman 1980 and Deng and Coleman 1999). Supplying these components and using molecular biology as a tool, cyanobacteria can produce synthetic compounds which can be commercialized (Figure 1) (Parmar *et al.* 2011 and Rosgaard *et al.* 2012). However, it has not been possible to obtain high concentration of these products (Deng and Coleman 1999 and Peralta-Yahya *et al.* 2012). It is known that carbon fixation is one of the limiting steps to increase the performance of these products, thus the process of carbon fixation and possibilities to make it more efficient has attracted significant attention in the last few years.

Cyanobacteria contribute to almost a quarter of the global carbon fixation (Field *et al.* 1998). If carbon fixation reaction could be increased, it may improve the efficiency to obtain more commercial substances and to help to reduce the amount of CO_2 in the atmosphere contributing to decrease the greenhouse effect.

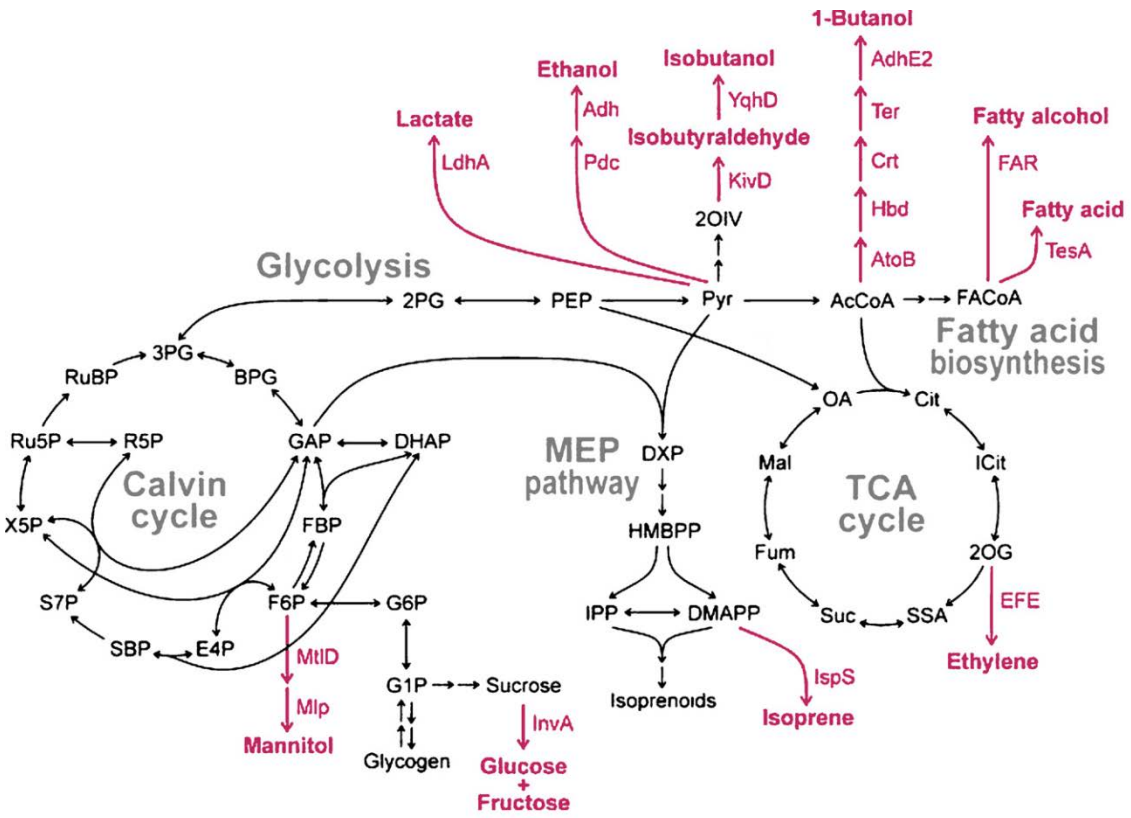

Figure 1. Main pathways present in cyanobacteria. The black color represents the native pathways. The violet color represents the different enzymes or pathways successfully introduced in cyanobacteria in order to produce different compounds of human interest. Abbreviations: 2OG, 2-oxoglutarate; 2OIV, 2-oxoisovalerate; 2PG, 2-phosphoglycerate; 3PG, 3-phosphoglycerate; AcCoA, acetyl-CoA; BPG, 1,3-bisphosphoglycerate; Cit, citrate; DHAP, dihydroxyacetone-phosphate; DMAPP, dimethylallyl-pyrophosphate; DXP, 1-deoxyxylulose-5-phosphate; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; FACoA, fatty acyl-CoA; FBP, fructose-1,6-bisphosphate; Fum, fumarate; G1P, glucose-1-phosphate; G6P, glucose-6-phosphate; GAP, glyceraldehyde-3-phosphate; HMBPP, 1-hydroxy-2-methyl-2-butenyl-4-pyrophosphate; ICit, isocitrate; IPP,isopentenyl-pyrophosphate; Mal, malate; OA, oxaloacetate; PEP, phosphoenolpyruvate; Pyr, pyruvate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; RuBP, ribulose-1,5-bisphosphate; S7P, sedoheptulose-7-phosphate; SBP, sedoheptulose-1,7-bisphosphate; SSA, succinic semialdehyde; Suc, succinate; X5P, xylulose-5-phosphate (Figure modified from Rosgaard *et al.* 2012).

In photoautotroph organisms such as cyanobacteria, algae and C_3 plants, Calvin-Benson-Bassham cycle (C_3) is the main cycle to assimilate carbon. The specific enzyme which fix carbon in this cycle is ribulose-1,5-biphosphate carboxylase/oxygenase (RuBisCO). In cyanobacteria RuBisCO is located in carboxysomes (Iwaki *et al.* 2006).

RuBisCO presents four different isoforms I,II,III and IV. Isoform I is the most abundant and it is found in plants, algae and cyanobacteria. When RuBisCO performs the carboxylation reaction, it uses CO_2 and ribulose-1,5-biphosphate (RuBP) to produce two molecules of 3-phosphoglycerate (3PGA). Then, 3PGA is used to regenerate RuBP or to increase the biomass and grow. It is in this point where SBPase proteins are imporant. In cyanobacteria there are two SBPase proteins (SBPase and SBPase/FBPase) (Tamoi *et al.* 1996). The SBPase/FBPase protein is present in different cyanobacteria including *Synechocystis* PCC 6803 and it is the responsible for directing the carbon into regeneration of RuBP or for leaving Calvin cycle to go into carbon metabolism (Tamoi *et al.* 1996).

Besides, RuBisCO's carboxylase activity, this enzyme can act as an oxygenase during photorespiration. In this case, RuBisCO forms 3-phosphoglycerate and 2-phosphoglycolate. The 2-phosphoglycolate is used to form glycolate. The latter can be secreted or used to synthesize amino acids or other compounds. Even though photorespiration leads to loses of CO_2, it is essential for the survival of photosynthetic organisms (Eisenhut *et al.* 2008).

Many studies have tried to overexpress or enhance the RuBisCO activity (Rosgaard *et al.* 2012). This led to a biomass increase in the cells, but RuBisCO was not incorporated in carboxysomes. Atsumi *et al*. 2009 produced isobutyraldehyde engineering cyanobacteria. They showed that overexpressing RuBisCO allows to increase CO_2 fixation leading to the augmentation of isobutyraldehyde production. Therefore, this study demonstrated that the CO_2 fixation is a bottleneck in the production of biofuels using photosynthetic organisms.

However, cyanobacteria, algae and C_3 plants contain another enzyme which fix carbon, Phosphoenolpyruvate carboxylase (PEPc). PEPc is the main enzyme responsible for the carbon fixation during photosynthesis in C_4 and CAM plants (Chollet *et al.* 1996 and Chen *et al.* 2002). In cyanobacteria, this enzyme is responsible for fixating 20% of the total carbon. It is essential as it plays an important anaplerotic role (Luinenburg and Coleman 1990). PEPc fix carbon to produce oxaloacetate which is an intermediate in the TCA cycle. Thus, cyanobacteria mainly fixate carbon into the C_3 cycle but they also contain the C_4 pathway (Coleman and Colman 1980 and Luinenburg and Coleman 1990).

The main problem in aquatic systems, where most of cyanobacteria live, is the inorganic carbon (Ci) availability. The CO_2 diffusion in water is much slower than in air. Furthermore, the equilibrium between CO_2 and HCO_3^- is slow in pH between 7 and 8.5. Nonetheless, cyanobacteria have developed different transporters in order to uptake inorganic carbon (Ci) efficiently when high or low levels are available (Shibata *et al.* 2002a, Benschop *et al.* 2003 and Price 2011). In addition, it seems that light is a pre-requisite for the expression of CO_2 response genes (Price 2011). The mechanism which transports inorganic carbon and concentrates the CO_2 around RuBisCO enzyme is called CO_2-concentrating mechanism (CCM).

Cyanobacteria have five different transporters to acquire inorganic carbon into the cell (Figure 2). Three of them transport HCO_3^- while the other two CO_2. BCT1 is a HCO_3^- transporter which is inducible under low levels of HCO_3^-. It has a high affinity to HCO_3^- and it is encoded by *cmpABCD* operon (Omata *et al*.1999 and Price 2011). SbtA is an inducible HCO_3^- transporter as well. This transporter is Na^+-dependent and it presents high affinity to HCO_3^- (Shibata *et al.* 2002b and Price 2011). There is another HCO_3^- transporter (BIcA) which is also Na^+ dependent. However, this transporter presents low affinity to HCO_3^- and the genes which encode these proteins are mainly constitutive expressed (Price *et al.* 2004 and Price 2011). The other two CCM are CO_2 transporters. NDH-I_4 is a constitutive CO_2 transporter while NDH-I_3 is an inducible one (Shibata *et al.* 2001). All of these transporters are present in the cyanobacterium *Synechocystis* PCC 6803 (Price 2011).

When CO_2 is up taken into the cytoplasm of the cell, it is converted to HCO_3^- by one type of Carbonic anhydrase (CA-like reaction). The CA-like reaction is associated to the plasma membrane and to some of the above discussed carbon transporters (active CO_2 transport and Na^+-independent HCO_3^- transport) (So *et al.*1998 and So *et al.* 2002). The HCO_3^- is poorly

permeable through the membranes. At this point, it is transported into the carboxysome where RuBisCO and another type of Carbonic anhydrase (CA) are present. In this case, CA is the responsible for converting HCO_3^- into CO_2. The action of this enzyme allows to increase the concentration of CO_2 around RuBisCO in order to lead to the carbon fixation reaction (Figge *et al.* 2001). The product of carbon fixation by RuBisCO is 3-PGA. This product is able to pass through the carboxysome shell and go to the cytoplasm, where it will be converted into other substances (Figure 2) (Price 2011).

Nevertheless, when CA in carboxysome converts HCO_3^- into CO_2, some CO_2 is lost. This lost is due to the fact that CO_2 can diffuse through the membranes even though the cabroxysome shell helps to prevent the CO_2 leakage. The major part of CO_2 that diffuses from the carboxysome is recycled by the CO_2 transporters. Hence, the CO_2 transporters can accept CO_2 from outer or inner sources. Mutants which lack CO_2 transporters have shown more CO_2 leakage than wild type (WT) cells (Maeda *et al.* 2002) which means that the CO_2 transporters act helping to avoid CO_2 leakage.

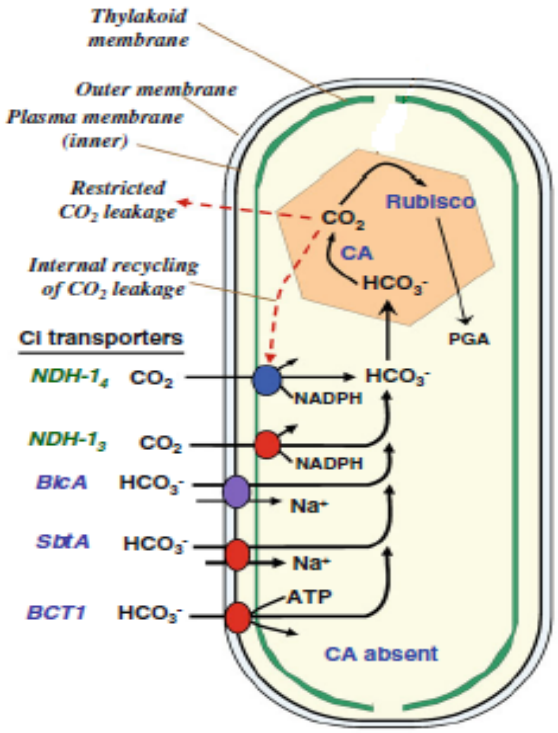

Figure 2. The five kinds of Ci transporters in β-cyanobacteria. The red transporters are induced when the Ci conditions are limited. The blue ones are expressed constitutively. The purple transporter can be expressed constitutively or under induction. Three of the transporters are located on the plasma membrane while the other two are present on the thylakoid membrane. The orange hexamer corresponds to carboxysome, where Carbonic anhydrase (CA) and RuBisCO are present. The substrates and products of both enzymes are annoted (Figure modified from Price 2011).

Although most research tried to study existing cycles in order to increase carbon fixation (Atsumi *et al*. 2009 and Rosgaard *et al.* 2012), Bar-Even *et al*. (2010) created novel non-existing synthetic pathways based on existing carbon fixing cycles, enzymes, and available information in protein data-bases. These synthetic pathways can be used to potentially increase carbon fixation. The pathways were designed using other enzymes rather than RuBisCO to increase

the carbon fixation because it is assumed that RuBisCO has been optimized naturally over millions of years.

The different synthetic pathways designed were based on 5000 enzymes and on five existing carbon fixation cycles (reductive tricarbxoylic cycle, oxygen-sensitive reductive acetyl-CoA pathway, 3-hydroxypropionate cycle, 3-hydroxypropionate/4-hydroxybutyrate cycle and dicarboxylate/ 4- hydroxybutyrate cycle). In addition, four criteria were established in order to find the best pathway that accomplishes the criteria. The four criteria were: (*i*) Specific affinity of the enzymes, measured as the maximum rate to generate 1 mg of product, (*ii*) Energetic cost, which refers to the efficiency to use the resources regenerated by light reaction, measured in terms of ATP and NADPH, (*iii*) Thermodynamically favorable, which means that the total free Gibbs energy required in the reactions is negative, (*iv*) The topology, that involves how the synthetic pathway would affect the native pathways in the host organism. This latter criterion takes in account the number of enzymes used in the pathway and the compatibility of the synthetic pathway in the guest organism.

After using these criteria, several carbon fixation pathways were created. However, all of them involved one or more carboxylation enzymes. In addition, the final product was a compound of at least two carbons. Small cycles with only four enzymes were found and they could be suitable because of the low number of enzymes but the favorable thermodynamics was not accomplished.

Also, different carbon fixation enzymes were compared (Phosphoenolpyruvate carboxylase, pyruvate carboxylase, acetyl-CoA and propionyl-CoA carboxylases and isocitrate dehydrogenase). PEPc was the most efficient candidate. Based on the C_4 plants cycle, different malonyl-CoA-oxaloacetate-glyoxylate (MOG) pathways were found. All MOG pathways produce glyoxylate, hydrolyze ATP and present a negative free Gibbs energy. Moreover, different ion strength and a broad pH range can be used. Glyoxylate is an intermediate compound in four different pathways in the cyanobacterium *Synechocystis* PCC 6803; Glycerate pathway, decarboxylation pathway, C_2 cycle during photorespiration and the unusual TCA cycle found to operate in cyanobacteria (Eisenhut *et al.* 2008 and Steinhauser *et al.* 2012). Two of the four pathways could improve the generation of products of interest using cyanobacteria (Glycerate pathway and the unusual TCA cycle). The glycerate pathway is where glyoxylate is converted to 3-phosphoglycerate (Eisenhut *et al.* 2008) and therefore could enhance the production of lactate, ethanol, isobutanol, 1-butanol and fatty acids (Figure 1). During several years, it has been thought that cyanobacteria did not have the complete TCA cycle because of the lacking of α-ketoglutarate dehydrogenase and NADH oxidase (Smith *et al*. 1967). Interestingly, a recent study has demonstrated that cyanobacteria have an unusual TCA cycle (Steinhauser *et al.* 2012). This unusual TCA cycle uses glyoxylate as an intermediate between isocitrate and malate. Since these two compounds are present in the TCA cycle, the ethylene production could be increased as well (Figure 1).

Despite of all putative MOG pathways, two the C4 Glyoxylate/ Alanine option and the C4 Glyoxylate/Lactate option were proved to be the optimal (Figure 3) since they were the shortest and the most thermodynamically favorable. These two MOG pathways present the same six first enzymes, but they differ in the following ones. A comparison between these two

MOG pathways was performed in order to examine which one could fit better in the cyanobacterium *Syenchocystis* PCC 6803, the organism in which we want to increase carbon fixation. The comparison led to the conclusion that the Lactate option is more appropriate compared to Alanine option. Even though the Lactate option requires more enzymes; the enzymes are better characterized under the use of Lactate, less number of organisms are needed to obtain all the eleven enzymes and the compatibility of the enzymes seems to be proper in *Syenochocystis* PCC 6803. However, a modification of the Lactate pathway was required. The enzyme number 6 (Methylmalonyl-CoA carboxytransferase) was replaced by Acetyl-CoA carboxylase since the original enzyme is not present in all the 4 microorganisms that we have selected to obtain the eleven enzymes. Thus, Lactate option has been chosen to follow during this project and it will be introduced into the cyanobacterium *Synechocystis* PCC 6803. The reason why this microorganism has been chosen is because it has been characterized properly and it is naturally transformed (Ikeuchi and Tabata 2001).

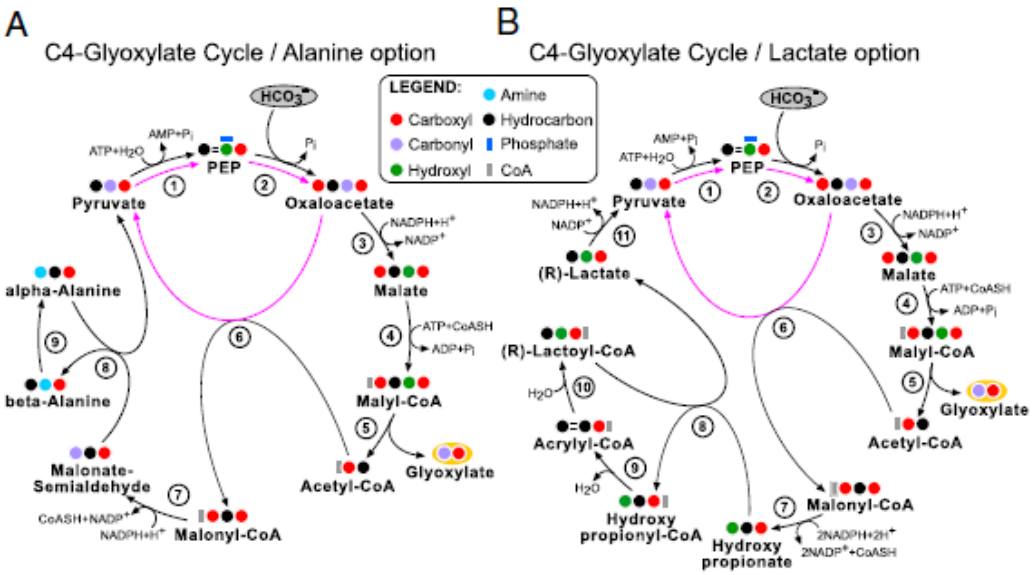

Figure 3. The two MOG pathways considered to study. A) C4-Glyoxylate Cycle/Alanine option which has 9 enzymes. The numbers correspond to the enzymes: 1, pyruvate water/phosphate dikinase; 2, PEP carboxylase; 3, malate dehydrogenase; 4, malyl-CoA synthetase; 5, malyl-CoA lyase; 6, methylmalonyl-CoA carboxytransferase; 7, malonate-semialdehyde dehydrogenase; 8, β-alanine-pyruvate transaminase; and 9, alanine 2,3-aminomutase. B) C4-Glyoxylate cycle / Lactate option. In this option there are 11 enzymes which correspond to: 1, pyruvate water/phosphate dikinase; 2, PEP carboxylase; 3, malate dehydrogenase; 4, malyl-CoA synthetase; 5, malyl-CoA lyase; 6, methylmalonyl-CoA carboxytransferase; 7, malonyl-CoA reductase; 8, propionate CoA transferase; 9, enoyl-CoA hydratase; 10, lactoyl-CoA dehydratase; and 11, lactate dehydrogenase.(Figure modified from Bar-Even *et al.* 2012).

Phosphoenolpyruvate carboxylase (PEPc):

Three different Phosphoenolpyruvate enzymes (Phosphoenolpyruvate carboxylase, Phosphoenolpyruvate carboxykinase and Phosphoenolpyruvate transphosphorylase) have been found in nature. All of them perform the β-carboxylation of PEP in order to lead to oxaloacetate. The main difference among them is the inorganic phosphate acceptor (Owttrim and Colman 1986).

PEP carboxylase catalyzes the irreversible reaction by using HCO_3^- and PEP in order to form oxaloacetate and inorganic phosphor in presence of Mg^{2+} (Figure 4) (O'Leary 1982 and Kai *et al.* 1999). Firstly Mg^{2+} binds the enzyme, followed by PEP and finally HCO_3^- is bound (Chollet *et al*. 1996).

This enzyme is present in bacteria, alga and vascular plants (Owttrim and Colman 1986 and Chen *et al*. 2002). Nevertheless, this enzyme has not been found in yeast, fungi and animal tissue due to it was replaced by pyruvate carboxylase (Owttrim and Colman 1986 and Lepiniec *et al.* 1994).

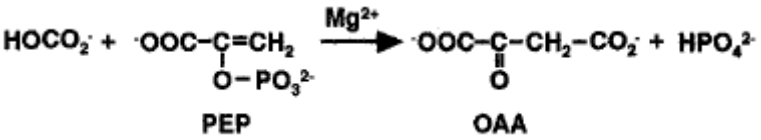

Figure 4. Reaction that Phosphoenolpyruvate carboxylase catalyzes. The substrates are bicarbonate ($HOCO_2$) and phosphoenolpyruvate (PEP). The cofactor is Mg^{2+} while the products are oxaloacetate (OAA) and inorganic phosphor (HPO_4^{2-})(Kai *et al.* 1999).

The main function of PEPc is to contribute to anaplerotic role into TCA cycle. It confers carbon skeletons for nitrogen assimilation and synthesis of amino acids (Luinenburg and Coleman 1990 and Lepiniec *et al.* 1994). In addition, it seems that in cyanobacteria, green algae and plant cell cultures, carbon fixation by PEPc can be enhanced during darkness when nitrogen assimilation occurs (Lepiniec *et al.* 1994).

Different PEPc isoforms have been found in different organisms. In higher plants, there are at least two different isoforms which are related to photosynthesis. CAM plants possess a PEPc which is active during darkness while C_4 plants contain one that is reversible active by a dependent light mechanism. Also, there are several PEPc isoforms which are no related to photosynthesis and they are involved in processes such as C/N partitioning in C_3 leaves, seed formation, fruit ripening, formation of C_4 acids in N_2 fixing legume root and in some plants seems to help to guard carbon metabolism when stoma are open (Chollet *et al.* 1996). Also, in *Chlamydomonas reinhardtii* and *Selenastrum minutum*, both unicellular green algae, two isoforms and four isoforms were found, respectively (Rivoal *et al.* 1996 and Rivoal *et al.* 1998). However, it has been not demonstrated that cyanobacteria have more than one PEPc isoform (Owttrim and Coleman 1986) but it has been concluded that PEPc is essential for cyanobacteria growth (Luinenburg and Coleman 1990).

The gene that expresses PEPc contains around 3000 nucleotides regardless of the organism it comes from. Thus, the amino acid sequence is formed by approximately 1000 amino acids (Lepiniec *et al.* 1994, Katagiri *et al.* 1985, Chen *et al*. 2002 and Smith *et al.* 2008). The molecular weight is around 100 KDa in all PEPc examined so far (Kodaki *et al.* 1985, O'Leary 1982, Luinenbug and Coleman 1992, Lepiniec *et al.* 1994, Rivoal *et al.* 1998, Chen *et al.* 2002, Smith *et al.* 2008). At least in various species of plants and in *Escherichia coli*, the active form of PEPc protein is a homotetramer (Ting *et al.* 1973, Uedan *et al.* 1976, Mares *et al.* 1979, Mares *et al.* 1980 and Kai *et al*. 1999).

When PEPc amino acid sequence is used to compare among organisms, PEPc from cyanobacteria is not grouped within the bacteria Kingdom in a phylogenetic tree (Rivoal *et al.* 1998 and Chen *et al.* 2002). This is in agreement with the low similarity (29.7%) that PEPc from *E.coli* and *Synechocystis* PCC 6803 showed when the amino acid sequences were compared using ClustalW2 (data not shown). The sequences were different but the essential domains and residues which are involved in activity were present in both sequences. The domains and the residues which carry out the catalysis are located in the C-terminal (Figure 6) of the protein and therefore, these two organisms show higher similarity in C-terminal than N-terminal (Smith *et al.* 2008). Besides, the comparison of PEPc amino acid sequence of different cyanobacteria has shown that the C-terminus is well conserved while the N-terminus shows more variability (Figure 5) (Ishijima *et al.* 1985, Luinenburg and Coleman 1992 and Smith and Plazas 2011). The C-terminus is conserved because it is where the catalytic domain of PEPc protein is located. The regulatory mechanism of this enzyme is located at the amino-terminus and it shows low homology due to the diversity of mechanisms that regulate this enzyme (Ishijima *et al.* 1985).

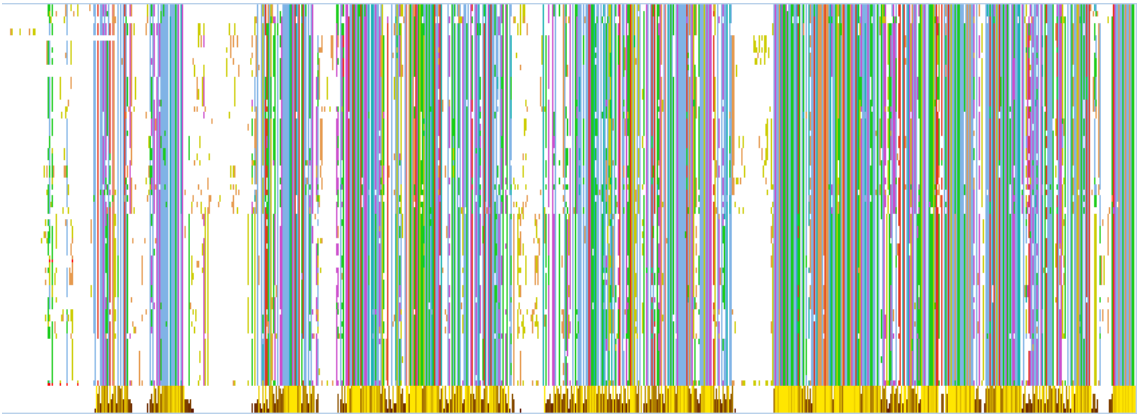

Figure 5. Overview of PEPc amino acid alignment of 63 cyanobacteria which are noted in Appendix 1. The different colors correspond to the different amino acids. There are 9 gaps which could be explained by insertions in some sequences along the evolution. As it is mentioned above the N-terminus is not conserved while C-terminus is well conserved. It seems to be a single isoform in cyanobacteria (Integrated Microbial Genome).

It seems that allosteric regulation of PEPc is common in higher plants (O'Leary 1982). Also, it has been observed in green algae (Rivoal *et al*. 1988), many cyanobacteria and some bacteria (Ishijima *et al.* 1985). Different substances are involved in this kind of regulation (Rivoal *et al.* 1998). For instance, bacterial PEPc depends on acetyl-CoA but it is inhibited by aspartate or malate (Chen *et al.*2002). In green algae, PEPc is activated by glutamine and dihydroxyacetone phosphate while it is inhibited by glutamate, aspartate, malate and 2-oxoglutarate (Rivoal *et al.* 1998). In higher plants, PEPc is regulated by two manners. The first one corresponds to activation by glucose 6-phosphate and glycine or inhibition by malate or aspartate. The second one corresponds to a phosphorylation on a serine residue at the N-terminus that activates the enzyme in a reversible manner. It is known that prokaryotes are not regulated by phosphorylation because they lack the phosphorylation domain (Lepiniec *et al.* 1994 and Chen *et al.* 2002).

In some cyanobacteria and plants it has been shown that oxaloacetate can inhibit PEPc activity (O'Leray 1982 and Owttrim and Colman 1986). Nevertheless, oxaloacetate levels *in vivo* are

low because when this compound is formed it is rapidly converted into aspartate or malate, which are more stable and repress PEPc activity as well (O'Leray 1982 and Luinenburg and Coleman 1992).

Phosphoenolpyruvate carboxylase amino acid sequence and possible structure in Synechocystis PCC 6803:

Even though the amino acid sequence of PEPc from *E. coli* and *Synechocystis* PCC 6803 has shown low similarity, there are several residues and domains which are conserved. The structure of PEPc from *E. coli* (ePEPc) has been deeply studied and crystallized (Kai *et al.* 1999) and its features have been used to compare PEPc from *Synechocystis* PCC 6803.

MNLAVPAFGLSTNWSGNGNGSNSEEESVLYQRLKMVEELWERVLQSECGQELVDLLTELRLQGTHEAITSEISEEVIMGITQRIEHLEL
NDAIRAARAFALYFQLINIVEQHYEQNEQQRNRWEASQETNFYEQAGNEEEMVPPSRLGASTEPLPVGIDQNELQASVGTFHWLM
RELKRLNVPPQHIQNLLDHLDIRLVITAHPTEIVRHTIRRKQRRVDRILRKLDQLQGSVTGRDWLNTWDAKTAIAQLTEEIRFWWRTD
ELHQFKPTVLDEVDYSLHYFDEVLFDAVPELSKRLGQAIKETFPHLRAPRANFCYFGSWVGGDRDGNPSVTPEVTWQTACYQRGLVL
GKYLFSLGELVAILSPSLHWCKVSQELLDSLERDRIQLPEIYEELSLRYRQEPYRMKLAYVTKRLENTLRRNNRLANPEERQTMITMPAE
NHYRTGEELLEELRLIQRNLTETGLTCLELENLITQLEVYGFNLAQLDFRQESSRHAEAIAEIAEYMGVLTTPYEEMAEEDKLAWLGVEL
QTRRPLIPQEMPFSERTRETIETLRTLRHLQMEFGVDICQTYIISMTNDASDVLEVLLLAKEAGLYDPATASNSLRIVPLFETVEDLKNAP
GIMDSLFSLPFYRATLAGSYHSLKELQNQPPDYYQIPTTTALLNPGNLQEIMVGYSDSNKDSGFLSSNWEIHKAQKSLQAVAQSHRVIL
RLFHGRGGSVGRGGGPAYKAILAQPAGTVDGRIKITEQGEVLASKYSLPELALYNLETLTTAVIQASLLKSSFDFIEPWNRIMEELACTA
RRAYRSLIYEEPDFLDFFLTVTPIPEISELQISSRPARRKGGKADLSSLRAIPWVFSWTQTRFLLPAWYGVGTALKSFVDQDPVKNMKLL
RYFYFKWPFFNMVISKVEMTLSKVDLTIASHYVQELSKPEDRERFDRLFQQIKQEYQLTRDFAMEITAHPHLLDGDRSLQRSVLLRNRT
IVPLGLLQISLLKRLRQVTQEAETSGVRYRRYSKEELLRGALLTINGIAAGMRNTG

Figure 6. Amino acid sequence of PEPc from *Synechocystis* PCC 6803. The residues that are marked in colors are important for the structure or activity of the enzyme according to PEPc from *E. coli*. The purple residues are involved in the stabilization of the structure, possibly a tetramer. The red residues are involved in the active site. The green color corresponds to a mobile loop. The blue color shows the domains which are involved in aspartate binding. The sequence was extracted from uniprot.org.

It has been showed that ePEPc is a homotetramer (Kai *et al.* 1999). It seems that two residues are involved in the stabilization of the tetramer (R 438 and E433 *E.coli* numbering) (Kai *et al.* 1999 and Smith and Plazas 2011) which are present in all cyanobacteria PEPc amino acid sequences described so far (Smith and Plazas 2011), including *Synechocystis* PCC 6803 (Figure 6-purple residues). This suggests that PEPc of *Synechocystis* PCC 6803 could be a homotetramer as well.

The active site of ePEPc has shown that there are seven crucial residues which are present in *Synechocystis* PCC 6803 sequence (H138, R396, K546, H579, R581, R587 and R699 *E.coli* numbering and red color in Figure 6 for *Synechocystis* PCC 6803). Lysine 546, Arginine 581 and 699 seem to bind to bicarbonate while Arginine 396 seems to be essential to PEPc function (Kai *et al.* 1999).

A glycine rich loop is a feature of PEPc enzyme (Figure 6-green color) which is involved in catalysis and the binding of Aspartate. One of the Glycines located in this loop (Underlined residue in Figure 6) helps to position the substrate into the active site. Besides, this loop forms

a lid which protects the intermediate products of the reaction from the water molecules that are surrounding the enzyme (Smith *et al.* 2008). When aspartate binds to PEPc immobilizes the loop away from the active site not permitting the activity of PEPc. In addition, there are three domains which are responsible to bind to Aspartate (Smith *et al*. 2008 and Smith *et al.* 2011) (Figure 6-blue color). Nonetheless, it has been shown that PEPc from *Anacystis nidulans* possesses these domains but it is not inhibited by aspartate (Ishijima *et al.*1985 and Smith *et al.* 2008). Therefore, it cannot be assumed that aspartate inhibits PEPc from *Synechocystis* PCC 6803.

Smith and partners tried to model the PEPc structure of different cyanobacteria using *Zea mays* as a template*. Zea mays* and *Synechococcus* PCC 7002 PEPc amino acid sequences show 30 % similarity, approaching *E. coli* (32%) (Smith *et al.* 2011). In spite of the low similarity among amino acids sequences, it seems enough to produce models (Smith *et al.* 2011). The models mainly present a characteristic β-barrel and on the C-terminus many α-helices. This is in agreement with ePEPc crystal structure which has many bundles of α-helices on the C-terminal of the β-barrel. The bundle of α-helices seems to be characteristic of PEPc protein and they have importance in the stabilization of the tetramer, bicarbonate and aspartate binding and also it is there where the active site is located (Kai *et al.* 1999 and Smith *et al.* 2011). Thus, it may be assumed that these structural features are present in all PEPc of cyanobacteria, including *Synechocystis* PCC 6803.

In addition, two extra beta sheets seem to be present in *Synechococcus* PCC 7002 PEPc modeling structure (Smith *et al.* 2011). Although *Syenchococcus* PCC 7002 and *Syenchocystis* PCC 6803 show quite high similarity (62%) (Smith *et al.* 2011) it cannot be attributed that these additional beta sheets are present in *Syenochocystis* PCC 6803. Further structural studies are required in *Syenchocystis* PCC 6803 PEPc in order to prove the reliable structure.

Malate dehydrogenase (MDH):

Malate dehydrogenase belongs to the NAD-dependent dehydrogenases family (Minarik *et al.* 2002). This enzyme catalyzes the interconversion of malate and oxaloacetate using NAD/H or NADP/H as a cofactor (Figure 7). MDH is widely distributed in prokaryotes and eukaryotes (Ocheretnia *et al.* 2000). In eukaryotes, it seems to participate in exchanging substrates and reducing equivalents between organelles and cytoplasm. Thus, there are different isoforms since this enzyme is present in chloroplasts, cytoplasm, mitochondria, microbodies and plastidis (Goward *et al.* 1994, Musrati *et al.* 1998 and Ocheretnia *et al.* 2000). The main difference among them is that the isoform located in the chloroplast uses NADP/H as a cofactor while all the other ones use NAD/H (Gietl 1992 and Musrati *et al.* 1998).

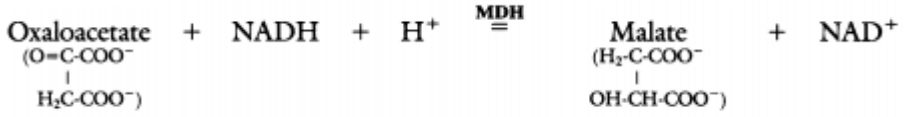

Figure 7. Reaction that malate dehydrogenase catalyzes. The reaction is reversible and therefore malate and oxaloacetate can be substrate or product of the reaction (Wilks *et al.* 1988).

In prokaryotes there is, mainly, a single isoform. At least one exception has been observed in cyanobacterium *Coccochloris peniocystis,* which possesses two isoforms (Norman and Colman 1991). When MDH from *E. coli* has been compared to all eukaryote MDH, it has shown that this MDH is similar to mitochondria isoform (Hall *et al.* 1992). This similarity could be explained by the endosymbiosis theory (Mc Alister-Heen 1988 and Minarik *et al.* 2002). However, MDH from cyanobacteria are all NAD/H isoforms and this does not agree with the theory because chloroplast isoform is NADP/H dependent (Ocheretina *et al.* 2000). In the cyanobacterium *Anacystis nidulans,* it has been demonstrated that MDH is located inside and on the surface of the thylakoids (Sallal and Nimer 1988). This enzyme plays an important role in the TCA cycle. The cycle produces different substrates for different metabolic pathways including amino acid synthesis (Norman and Colman 1991).

Usually, the reaction that MDH catalyzes favors malate as a product (Honka *et al.* 1990 and Gietl 1992). However, in the extreme thermophilic *Methanothermus fervidus* and *Streptomyces aureofaciens* the reduction of oxaloacetate is preferred (Honka *et al.* 1990 and Mikulasova *et al.* 1997).

In general, MDH gene is composed by approximately 1000 nucleotides. The amino acid sequence contains around 300 amino acids in all MDH studied so far (Hall *et al.* 1992, Cendrin *et al.* 1993 and Minarik *et al*. 2002). Most of MDH studied are active in form of a homodimer (Iijima *et al.* 1979, McAlister 1988, Nishiyama *et al.* 1990, Hall *et al.* 1992, Mikulasova *et al.* 1997, Minarik *et al.* 2002). However, in some bacteria, MDH protein is a homotetramer (Iijima *et al.* 1979 and Musrati 1998). The molecular weight of a monomer is roughly 30 KDa in most organisms (Iijima *et al.* 1979, McAlister 1988, Honka *et al.* 1990, Nicholls *et al.* 1992, Cendrin *et al.* 1993 and Mikulasova *et al.* 1998). Nonetheless, the MDH of cyanobacterium *Coccochloris peniocystis* has shown a molecular weight of 90 KDa but the active form of this particular MDH is unknown (Norman and Colman 1991).

There are several substances which inhibit the activity of this enzyme. Oxaloacetate inhibits MDH activity in mitochondria isoform from animals, Pseudomonas, in the extreme thermophile *Thermus flavus*, in the extreme halophilic archeabacterium *Haloarcula marismortuim* and the cyanobacterium *Coccochloris peniocystis* (Iijima *et al.* 1980, Norman and Colman 1991, Hall *et al.* 1992 and Cendrin *et al.* 1993). Nevertheless, this inhibition phenomenon is not observed in *Streptomyces aureofaciens* (Mikulasova *et al.* 1997). In the cyanobacterium *C.peniocystis* ATP, citrate acetyl-CoA and CoA seems to inhibit MDH activity as well (Norman and Colman 1991). Despite the inhibition effect, it has been demonstrated that in *E.coli* high concentrations of malate or citrate stimulate the activity of this enzyme in the direction of oxaloacetate formation (Hall *et al.* 1992).

Malate dehydrogenase amino acid sequence and possible structure in Synechocystis PCC 6803:

In general, the enzyme is divided in two domains, the NAD/H and the catalytic domain (Minarik *et al.* 2002). The nucleotide domain is located at the N-terminus of the protein while the C-terminal is where the substrate binds (catalytic domain) and therefore, where the amino acids which are involved in catalysis are located (Minarik *et al.* 2002). Although there are several MDH isoforms, it is known that the amino acids involved in catalysis and nucleotide binding are well conserved among MDH (Gietl 1992).

Three Arginine residues (R93, R99 and R162) are crucial for the substrate binding and catalysis (Figure 8-yellow residues) (Gietl *et al.* 1991 and Goward *et al*. 1994). Two of them (R93 and R99) are located in a well-conserved loop in most of MDH (Figure 8-green residues) (Goward *et al.* 1994 Musrati *et al.* 1998). R93 and R162 bind to the substrate and help to orientate it in the proper manner in order to lead to the catalysis (Musrati *et al*. 1998). Also, an Aspartic acid in position 44 plays an important role in NAD/H binding (Figure 8-red residue) (Goward *et al.* 1994 and Musrati *et al.* 1998). In addition, two residues are important during catalysis (Aspargine (N131) and the Histidine (H186) (Figure 8-Blue residues), which are involved in the proton relay system (Musrati *et al.* 1998).

MNILEYAPIACQSWQVTVVGAGNVGRTLAQRLVQQNVANVVLLDIVPGLPQGIALDLMAAQSVEEYDSKII
GTNEYEATAGSDVVVITAGLPRRPGMSRDDLLGKNANIVAQGAREALRYSPNAILIVVTNPLDVMTYLAWK
VTGLPSQRVMGMAGVLDSARLKAFIAMKLGACPSDINTLVLGGHGDLMLPLPRYCTVSGVPITELIPPQTIE
ELVERTRNGGAEIAALLQTGTAYYAPASSAAVMVESILRNQSRILPAATYLDGAYGLKDIFLGVPCRLGCRGV
EDILEVQLTPEEKAALHLSAEAVRLNIDVALAMVSDG

Figure 8. Amino acid sequence of MDH. The length of the sequence is 324 amino acids. The colors correspond to different remarkable residues which are involved in NAD/H binding (red), substrate binding (yellow) proton relay system during catalysis (blue) and stabilization of a possible dimer structure (grey).

Even though MDH amino acid sequences from different sources have low similarity, it has been observed that the three-dimensional structures are similar (Goward and Nicholls 1994). MDH structure from *E.coli* (eMDH) (Hall et al. 1992) has been studied and it may be attributed to MDH structure from *Syenchocystis* PCC 6803.

Also, it has been shown that malate dehydrogenase and lactate dehydrogenase have similar catalytic mechanism and therefore they are close related (Hall *et al.* 1992, Nicholls *et al*. 1992, Goward and Nicholls 1994 and Minarik *et al.* 2002).

The active site of MDH is a hydrophobic vacuole. The nucleotide binds firstly, followed by the substrate. When the complex is bound a conformational change occurs, thus the external loop closes the active site and the catalysis occurs (Hall *et al.* 1992 and Minarik *et al.* 2002).

The active form of eMDH is a homodimer. Some interactions with the solvent help to the interaction between subunits but an Aspartic acid (D45 *E.coli* numbering) is involved in the stabilization of the dimer (Minarik *et al.* 2002). This residue is also present in *Synechosytis* amino acid sequence (Figure 8- grey residue (D56). Thus, it might be possible that MDH from *Synechocystis* is active as a homodimer but more studies should be carried out in order to know the real structure of MDH in *Synechocystis* PCC 6803.

Material and Methods:

Cloning and transformations

Amplification of the construction

PCR reactions were performed with Phusion High-Fidelity Hot Spot II DNA polymerase protocol from Finnzymes. When plasmids were amplified from 1 pg to 10 ng were used as a template. In order to amplify genes, DNA sequence of *Syenchocystis* PCC 6803 genome was used as template. The amount of genomic DNA was in the range of 50-250 ng. All the DNA and RNA concentrations in this study were measured by Nanodrop 2000 Spectrophotometer from Thermo Scientific.

All the PCR products were purified using Gene JET Purification kit from Thermo Scientific. However, when unspecific products were present in the PCR product, gel purification was performed using Gene JET Gel Extraction kit from Thermo Scientific.

Gibson assembly

Gibson method was accomplished using 5 µL of DNA (from 90 pmol to 50 ng) and 15 µL of master mixture (or with 15 µL of dH_2O in case of negative control). The master mixture composed of 82.57 µL dH_2O, 16.5 µL of 5x isothermal buffer, 1.65 µL of T5 exonuclease (0.2 U· μL^{-1}), 8.25 µL of Taq DNA ligase (40 U · μL^{-1}) and 1.03 µL Phusion DNA polymerase (2 U· μL^{-1}). The DNA with the Gibson master mixture was blended and incubated for 60 minutes at 50°C.

Transformation of Escherichia coli DH5α strain

Approximately 5 µL of the Gibson mixture was placed in 1.5 microcentrifuge tube. *Escherichia coli* DH5α strain was previously prepared to be competent and it was stored at -80°C. Then, 100 µL of competent cells were first thawed on ice and added into the microcentrifuge tube. The mixture was incubated on ice for 30 minutes, heat shocked for 1 minute at 42°C and chilled on ice for 5 more minutes. 900 µL of LB media at room temperature was subsequently added into the microcentrifugetube and incubated 60 minutes at 37°C. The tube was centrifuged at 13000 rpm for 2 minutes. 900 µL of the supernatant were discarded while the 100 µL remaining were used to resuspend the cells. Finally, the resuspended solution was spreaded onto LB agar plate with Chloramphenicol [20 µg ·mL^{-1}]. The plate was incubated overnight at 37 °C.

The colonies that appeared after the incubation overnight were used to verify the incorporation the desired construction. Dream *Taq* DNA polymerase protocol from Fermentas was used to do PCRs and screen some crucial parts of the designed construction (Sp1 and Sp2). The colonies which were positive in both screening parts were grown in LB media overnight. Then, the plasmid was extracted using Gene JET Plasmid Miniprep kit from Thermo Scientific. Once the plasmid was purified it was sent to sequencing (Macrogen Inc 2013). When the sequencing results demonstrated that the plasmid was correct, *Synechocystis* PCC 6803 was transformed.

Transformation of Synechocystis PCC 6803

Synechocystis PCC 6803 (wild type) cultures were grown in BG11 medium for 2-3 days. When the OD_{750} was approximately 0.3 the cells were harvested by centrifugation (5000 rpm for 10 minutes at 20°C) and resuspended in 250 μL of fresh BG11. 1 μg of plasmid was used to transform 100 μL of cell suspension which was next incubated at 25°C in low light for 4-6 hours. The suspension was spreaded on a nitrocellulose filter on top of BG11 agar plates. After 24 hours the filters were moved to BG11 agar plates with Chloramphenicol [20 $\mu g \cdot ml^{-1}$]. When colonies appeared, they were sent to sequencing in order to confirm that they had incorporated the desired construction.

Semiquantitative RT-PCR

Light/Darkness samples

Cultures of cyanobacteria were grown in 0.5 liter bottle with 300 mL of BG11 or BG11 with Chloramphenicol [20 $\mu g \cdot mL^{-1}$], under light conditions (27 $\mu E \cdot m^{-2} \cdot sec^{-1}$) at 30 °C with air bubbles. When the OD_{750} reached 0.3 the first duplicate samples were taken (L). Afterwards, the bottles were wrapped with aluminium foil in order to create dark environment. After 1 and 24 hours duplicate samples were taken (D1 and D24). The samples were taken by pipetting 10 mL of culture and transferred into falcon tube. The tubes were centrifuged at 5000 rpm for 10 min at 20 °C and the supernatants were discarded. The pellets were frozen with liquid nitrogen and stored at -80°C.

RNA extraction

The pellets were resuspended in 500 μL of TRIzol. They were transferred into screw-cap tubes which contained 0.2 g of glass beads. The tubes were shacked in the bead beater machine (Precellys 24- Bertin technologies) at 6800 rpm for 30 seconds. The procedure was repeated 3 times and between them the tubes were kept for 2 min on ice. 100 μL of chloroform was added in each tube and mixed gently by inversion. The tubes were incubated at room temperature for 10 min. After that, a centrifugation was performed at 14000 rpm for 15 min at 4°C. An aqueous phase appeared and it was transferred into new tubes. Then, 250 μL of isopropanol were added into the tubes, mixed and incubated at room temperature for 10 min. Centrifugation was again done at 14000 rpm for 10 min at 4°C. The ensuing pellets were kept and washed with 1mL of 75% ethanol. The ethanol was removed carefully and the pellets were air dried. The pellets were subsequently dissolved in 60 μL of autoclave distillated water. The suspensions were transferred into new tubes and the concentration of RNA was measured. In order to avoid DNA contamination, DNAse treatment was performed by DNase I, RNase-free from Fermentas. The treatment was modified from the standard protocol. It was accomplished by adding in a RNase-free tube: 6 μg of RNA, 10 μL of 10x reaction buffer with $MgCl_2$, 10 μL of DNase I, RNase free and up to 100 μL of autoclave distillated water. The mixture was incubated at 37°C for 30 min followed by the addition of 20 μL of EDTA (25mM). Thus, 10 min of incubation was done at 65°C. The samples were transferred into new. Finally, the RNA concentration was measured again.

RT-PCR

The reverse transcription reaction was performed using cDNA synthesis kit from Thermo Scientific. For each sample approximately 0.5 µg of RNA was taken, 1 µL and up to 12 µL of random hexamer primer and Water nuclease-free were added, respectively. Then, 4 µL of 5X reaction buffer, 1 µL RiboLock RNase Inhibitor (20u · μL^{-1}), 2 µL of 10 mM dNTP Mix and 1 µL of RevertAid Minus M- MuLV Reverse Transcriptase (200u · μL^{-1}) were added. The samples were incubated 5 min at 25°C, 60 min at 42°C and 5 min at 70°C.

Afterwards, approximately 1 µL of cDNA (the amount of cDNA to perform this step depended on the amount of RNA that was used to perform cDNA), 2 µL of 10X Dream Taq Buffer, 0.4 µL of dNTP Mix, 2 µL of each primer, 0.1 µL of DreamTaq DNA Polymerase and 12.5 µL of water, nuclease-free were used to perform a PCR. Four different pair of primers (23S, universal, native and engineered (Table 1) were used in order to test if overexpression of PEPc or MDH was accomplished. The negative control was performed using 0.5 µg of the RNA tube after DNAse treatment.

Table 1. Primers used in the semiquantitative RT-PCR. For means forward primer while rev means reverse primer. G+C refers to Guanine and Cytosine content in the primer. The temperature corresponds to the ones used in the PCR reaction.

Primers	Sequence	Number of nucleotides	G+C content (%)	Temperature (°C)
23S_for	CTGATCTCCGCCAAGAGTTC	20	55	62
23S_rev	TTACCGTTGGCACGATAACA	20	45	62
PEPc_universal_for	GGTCTGGTAATGGCAATGGTTC	22	50	62
PEPc_universal_rev	CCCGGATGGCATCATTGAGT	20	55	62
PEPc_for	GACGGCGATCGCTCTTTGCAA	21	57	55
PEPc_native_rev	TGGCCGCGGTGTTCGTTC	18	67	61
PEPc_engineered_rev	TCACAGTAAGCAGGGTCTAGGCA	23	52	52
MDH_universal_for	GCTGAAATTGCCGCCTTACT	20	50	60
MDH_universal_rev	AGAAAGATGGAGGGCAGCTT	20	50	60
MDH_for	CGGGCACAGCCTATTATGCG	20	60	60
MDH_native_rev	GTATTGGCACTGTCCGTTATCGTG	24	50	60
MDH_engineered_rev	CTGCAGCGGCCGCTACTAGT	20	65	60

Western Blot

Light/Darkness samples

Cultures of cyanobacteria were grown in 1 liter bottle with 750 mL of BG11 (WT) or BG11 with Chloramphenicol (Engineered cells) [20 μg ·mL^{-1}], under light conditions (27 μE·m^{-2}·sec^{-1}) at 30 °C with air bubbles. When the OD_{750} reached 0.5 the first duplicate samples were taken (L). Then the bottles were wrapped with aluminium foil in order to create dark environment. After 24 hours duplicate samples were taken (D). The samples were taken by pipetting 100 mL of culture and transferred into falcon tube. The tubes were centrifuged at 5000 rpm for 10 min at 4 °C and the supernatants were discarded. The pellets were frozen in liquid nitrogen and stored at -80°C.

Protein extraction

The pellets were suspended with 2 mL of PBS buffer. The samples were centrifuged (5000 rpm for 10 min) and the pellets were resuspended in 200 μL of PBS. The cells were frozen at -80 C and quickly thawed at 37° C. At that moment, 2 μL of protease inhibitor was added. After this step, all the samples were kept on ice.

The cells were broken using the glass beater (5500 rpm for 30 seconds, three times, keeping the samples on ice 2 min between). Then, 100 μL of PBS buffer was added followed by 2 minutes of centrifugation at 14000 rpm at 4 °C. The supernatant was collected and centrifuged again for 2 minutes. The green supernatant fraction corresponded to the crude cell extract. The protein concentration was measured using RC DC protein assay from BioRad. The standard was made using Albumin from bovine serum provided by Sigma.

Protein gel

Two protein gels with exactly the same samples and amounts were run. The gels used (Min-PROTEAN TGX™) and the running buffer were provided by BioRad. The gels were run for 40 min at 200 V.

One gel was stained with 20 mL of Page Blue TM protein staining solution (shacked, at room temperature, for one hour). Then, Page Blue was replaced by distillated water, changing the water 5 times during 1 hour. The other gel was used to transfer the proteins from the gel to a nitrocellulose membrane (Western Blot).

Western blot

The gel with all the proteins was inserted in a sandwich which contained a nitrocellulose membrane (Amersham™ Gybond-ECL from GE Healthcare), two filter papers and two sponges. All these components were previously wet with transfer buffer for at least 10 min. The blotting procedure was performed in TE 22-tank transfer unit from Amershan Bioscience at 4°C at 30 V overnight. Following this, the gel was stained with Coomassie Blue (data not shown) while the membrane was blocked with T-TBS buffer solution with 5% of BSA protein. It was shacked for 1 hour at room temperature. Then, the membrane was washed 3 times (15 minutes each) with T-TBS buffer.

The dilution of the primary antibodies generated against PEPc plant and purified from rabbit (Agisera Company) was 1:1000, which was added on the membrane and shacked for 1 hour at room temperature. The washes were repeated and were followed by the addition of diluted secondary antibody (generated against rabbit antibody and purified from goat (Biorad Company)(1:5000) on the membrane. It was incubated under gently shaking for 1 hour, at room temperature and washed another three times.

The membrane was thus incubated with Immuno-Star^{TM} HRP- substrate kit from Biorad and the chemiluminescence reaction was detected by Chemi Doc XRS machine from Biorad.

Results:

PEPc-pBlueScript and MDH-pEERM

The first step in order to overexpress PEPc and MDH was to design the strategy and the primers. Since the PEPc and MDH genes are present in *Synechocystis* PCC 6803 genome, the strategy was to insert an extra copy of the native genes under the control of a strong promoter. PEPc sequence was introduced in a plasmid based on pBlue Script (Figure 9-B1) while MDH sequence was introduced in pEERM plasmid (Figure 9-B2). These plasmids contain the upstream and downstream regions of psbA2 gene of *Synechocystis* PCC 6803, a transcription terminator and the Chloramphenicol resistant cassette as a selection marker. The sequence is showed in Figure 9A and it was designed in order to replace the psbA2 gene in *Syenochocystis* PCC 6803 genome.

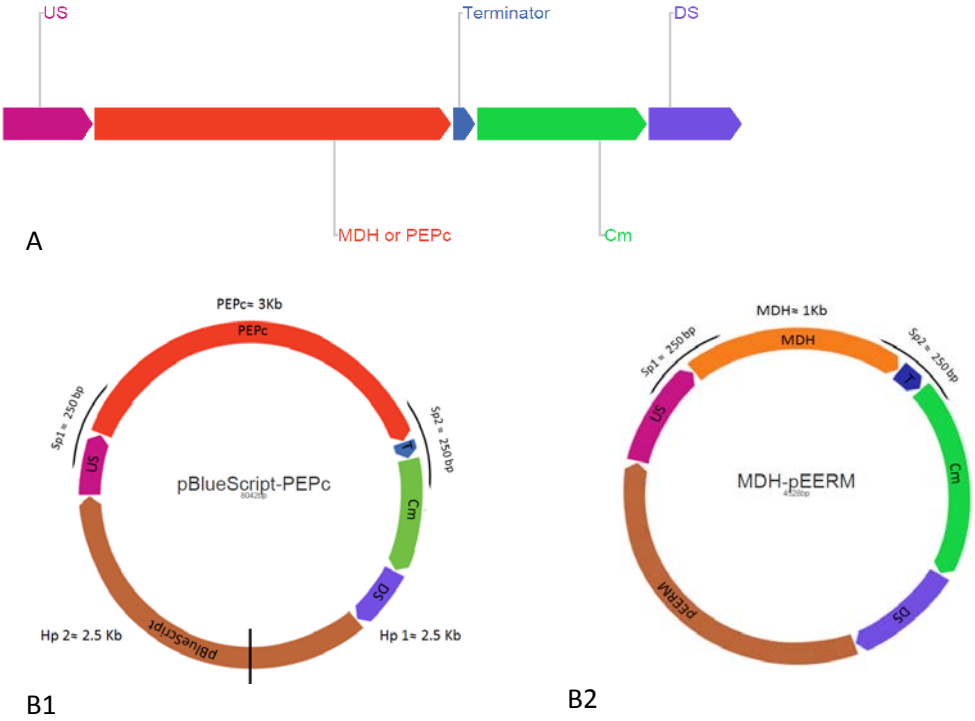

Figure 9. Plasmid design. A) Constructed sequence. Abbreviations: Cm- Chloramphenicol resistance cassette, , DS- downstream region of psbA2 gene, MDH-MDH gene, PEPc- PEPc gene, US- upstream region of psbA2 gene. B) Designed plasmids. Abbreviations: Cm- Chloramphenicol resistance cassette, , DS- downstream region of psbA2 gene, MDH-MDH gene, PEPc- PEPc gene, Sp- Screening parts, T- transcriptional terminator, US- upstream region of psbA2 gene. B1) pBlueScript plasmid with PEPc gene. Hp1 and Hp2 correspond to the two halves of the plasmid. B2) pEERM plasmid with MDH gene.

The different parts were amplified for further assembly (Figure 10). In PEPc cloning, PEPc sequence was successfully amplified. However, the amplification of complete plasmid did not work properly (Figure 10A- band number 3). An undesired band at 1.5 Kb appeared while an expected band is present at 5 Kb, but it was really weak. The amplification of one half of the plasmid (Hp2- Figure B1) was successfully achieved while the other half (Hp1- Figure B1) was

accompanied with unspecific products. In order to achieve the desired amplification in the whole plasmid and Hp1, different annealing temperatures were tested (data not shown) but it did not result in better results.

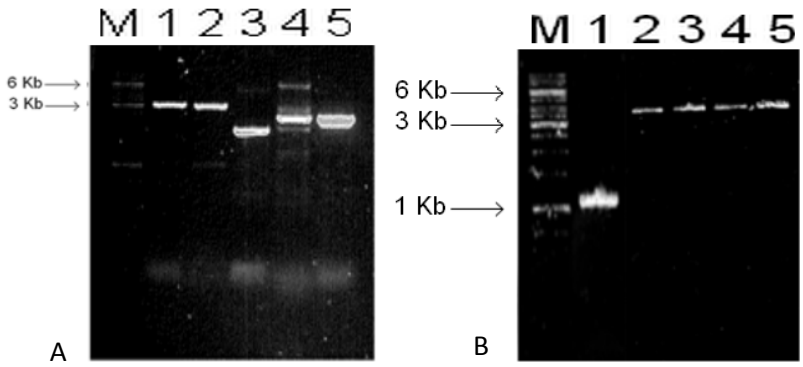

Figure 10. Agarose gel loaded with PCR products. A) Amplification of PEPc construction. The different lines correspond to: M- 1 Kb marker. Line 1 and 2- PEPc gene. Line 3- whole plasmid. Line 4- Hp1 and line 5 - Hp2. B) Amplification of MDH construction. The lines correspond to: M- 1 kb Marker, Line 1- MDH gene and lines 2-5 pEERM plasmid.

Another attempt was to amplify the whole plasmid and Hp1 using 2 step PCR. Even though it did not work for the amplification of the whole plasmid, it resulted on a better amplification of Hp1 (Figure 11).

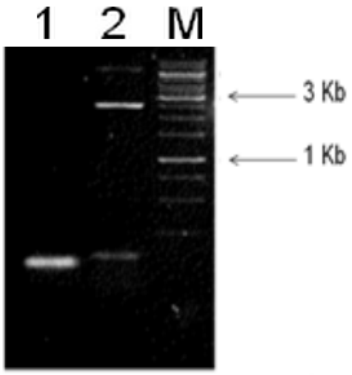

Figure 11. The 2-step PCR products in PEPc construction. M – 1 kb marker. Line 1 - whole plasmid. Line 2- Hp1.

In MDH cloning, MDH was amplified successfully while pEERM amplification showed some unspecific bindings at the temperature used according to Oligo Calc: Oligonucleotide Properties Calculator (data not shown). Subsequently, different annealing temperatures were used (data not shown). At 58.4 °C the primers work better and the amplification of pEERM was accomplished (Figure 10B).

After Gibson assembly and transformation of *E. coli DH5α* strain, a number of colonies appeared on the LB plate with Chloramphenicol. In the PEPc experiment, 22 colonies appeared while just 2 colonies were present in the MDH experiment. PCR was performed in order to confirm the incorporation of the correct plasmid with the correct sequence (amplifying Sp1 and Sp2 (Figure 9B1 and B2). In amplification of Sp1 and Sp2 in PEPc construction, unspecific products were present (Figure 12A) but colonies number 1, 4 and 8 had brilliant bands. On the other hand, both MDH colonies showed the correct bands (Figure 12B).

Selected colonies were sequenced in order to confirm that they had incorporated the desired genetic construction. The sequencing results showed that for PEPc only colony number 8 had incorporated the construction (data not shown) while both MDH colonies were correct. Therefore, colony number 8 (PEPc) and colony number 2 (MDH) were selected and used to transform into the cells of *Synechocystis* PCC 6803.

After several weeks, some transformed *Synechocystis* colonies appeared with an additional copy of the PEPc or MDH. Selected colonies were sequenced in order to check that the sequences were correct and the results demonstrated that the transformation with extra PEPc or MDH gene were successfully introduced into the genome replacing the psbA2 gene.

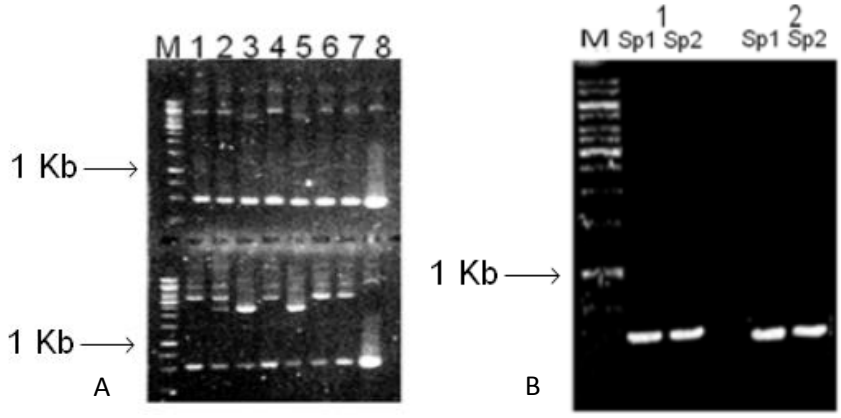

Figure 12. Agarose gels after screening PCR. M corresponds to 1 Kb Marker. The numbers correspond to the colonies. A) The gel shows both PCR screening parts (Sp1 and Sp2) for PEPc construction. The upper part corresponds to screening PCR- Sp1. The downer part of the gel corresponds to the screening PCR -Sp2. B) The gel shows both PCR screening parts of the two colonies in case of MDH experiment.

Relative levels of PEPc or MDH transcript in WT and transformed cells

To perform Semiquantitative RT-PCR different primers were designed (Table 1 and Figure 13). The first pair of primers (23S) was designed to amplify 23S cDNA. The 23S mRNA was used as template in order to obtain 23S cDNA, therefore, it represents the expression of 23S gene. The second pair of primers (Native) was designed to amplify the PEPc or MDH gene which is present in *Syenchocystis* PCC 6803 wild-type genome. The forward primer annealed at the end of the MDH or PEPc sequence, while the reverse primer annealed to the gene next to PEPc or MDH gene in *Synechocystis* genome (sll0892 and ssl1762 respectively). The third pair (Engineering) can detect the additional native copy of PEPc or MDH which has been inserted. It has replaced psbA2 gene and it is under control of psbA2 promoter. The forward primer is in fact, the same as the native primer forward, so it annealed at the end of PEPc or MDH gene. The reverse one annealed to the spacer and terminator sequence. These two sequences were introduced together with PEPc or MDH additional copy (Figure 9A). The last pair of primers (Universal) annealed in the middle of PEPc or MDH sequence. With these primers, total PEPc or MDH expression was detected but it was not possible to distinguish from which PEPc or MDH sequence was expressed. All the PCR products were designed to amplify 250 bp of the gene.

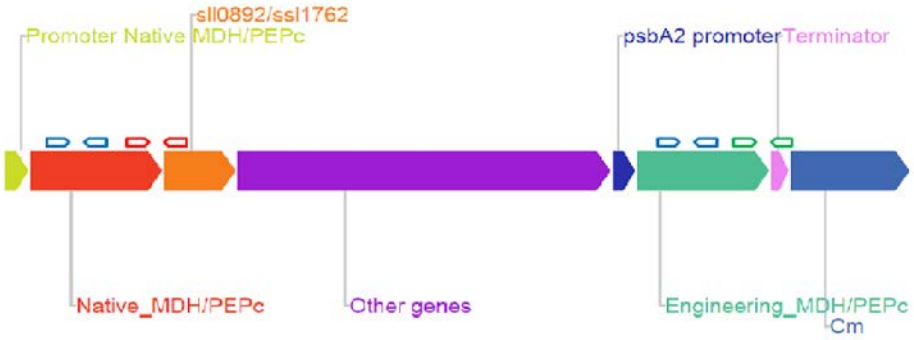

Figure 13. Primers designed for the semiquantitative RT-PCR technique. The DNA sequence and the different parts are represented by the arrows in different colours. The primers are the triangles over the DNA sequence. Red primers anneal to the Native gene (Native), Green primers to the inserted extra copy of the native gene (Engineering) and the blue ones can anneal to both genes, native and the inserted extra copy of the native gene (Universal).

Table 2. PCR cycles used in semiquantitative RT-PCR.

Primers	PCR Cycle
23S	17
PEPc_Native	30
PEPc_Engineering	26
PEPc_Universal	28
MDH_Native	27
MDH_Enginnering	22
MDH_Universal	26

As it can be noticed in Table 2, different cycles have been chosen for the different primers. These cycles correspond to the beginning of the exponential phase in the PCR reactions. Lower cycle means higher amount of cDNA and therefore more expression of the gene. 23S primers present the lower cycle due to the amount of transcript is high. MDH or PEPc present in WT genome (MDH_native or PEPc_native) seems to have a low expression due to the cycle which starts the exponential phase is high. The inserted extra copy of MDH or PEPc gene (MDH_engineering or PEPc_engineering) is more expressed than the native gene (lower cycle number). The universal primers (MDH_universal or PEPc_universal) have shown an intermediate cycle between engineering and native MDH and PEPc.

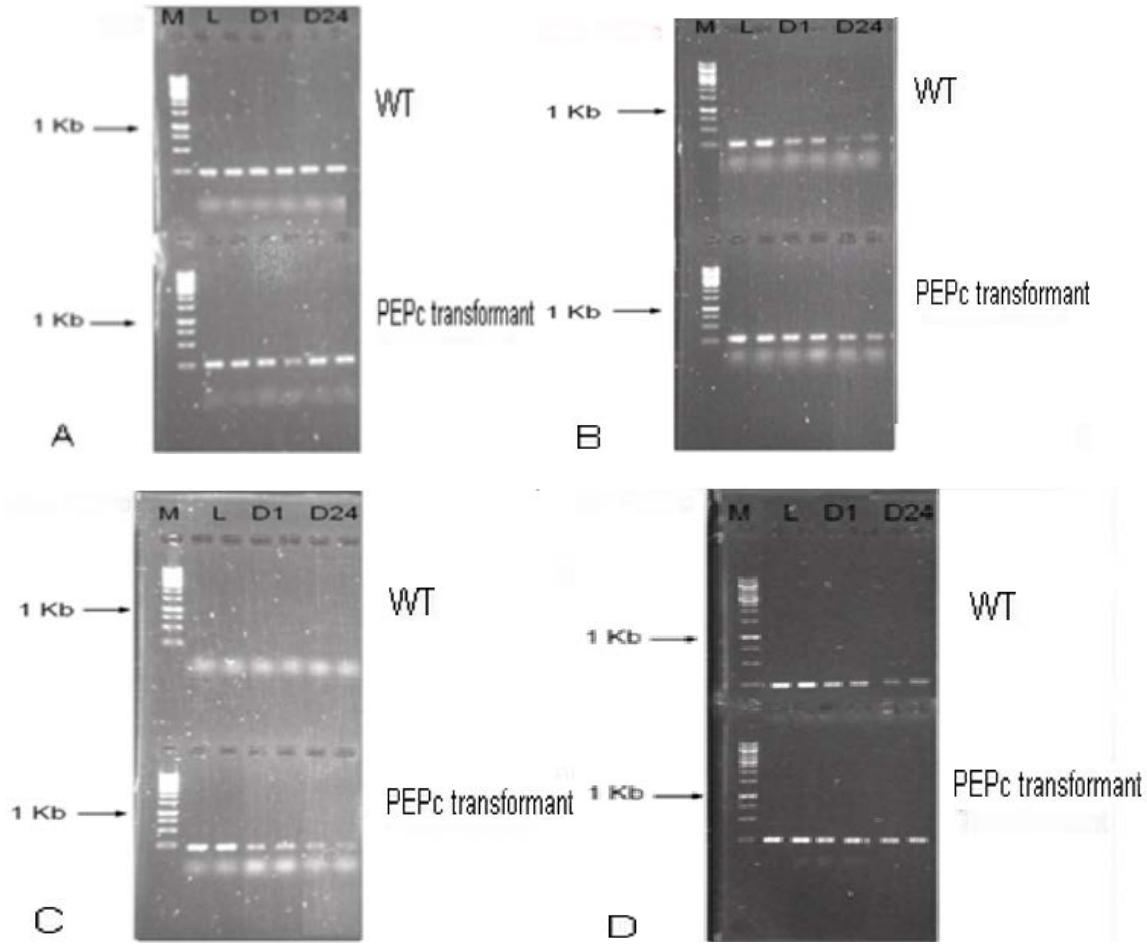

Figure 14 . Agarose gels showing the products of semiquantitative RT-PCR in PEPc experiment. A) Expression of 23S gene (reference gene). B) Expression of PEPc gene which is present in *Syenchocystis* PCC 6803 genome (Native_PEPc). C) Expression of PEPc gene which has been inserted in *Syenchocystis* PCC 6803 genome replacing psbA2 gene (Enginerring_PEPc). D) Expression of total PEPc gene (Universal_PEPc). WT means wild-type *Synechocystis* PCC 6803. PEPc Transformant corresponds to *Syenchocystis* PCC 6803 with the additional copy of the native PEPc gene. The upper part of the four gels correspond to wild type Synechocystis PCC 6803 while the lower part to PEPc transformant. Also all the gels show the three different conditions tested with two replicates each, light (L), 1 hour darkness (D1) and 24 hour darkness (D24).

In the PEPc experiment, the results of the 23S gene expression (Figure 14) were quite constant in both cases WT and PEPc transformant, except for one sample (PEPc transformant D1-second replicate) (Figure 14 A) which seems to be weaker. The native PEPc gene showed brilliant bands under light in WT and PEPc transformant. However, they were weaker during darkness at both time points (D1 and D24) (Figure 14 B). The inserted PEPc gene did not (as expected) show any band in the WT. Nevertheless, in PEPc transformant inserted PEPc gene was present. Under light conditions the bands were more shining than in darkness (Figure 14 C). When universal PEPc primers were used, the same pattern was observed as native PEPc gene in WT. In PEPc transformant the bands were more abundant compared to the WT but it could be noticed a reduction of the expression in darkness (Figure 14 D).

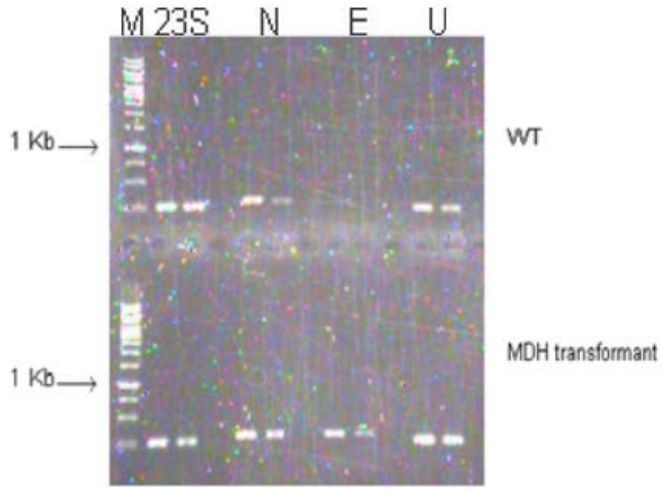

Figure 15. Agarose gel that shows the products of semiquantitative RT-PCR in MDH experiment. The upper part of the gel corresponds to WT *Synechocystis* PCC 6803. The lower part of the gel corresponds to MDH transformant. 23S corresponds to the expression of 23S gene. N corresponds to MDH gene present in *Syenchocystis* PCC 6803 genome. E corresponds to the additional native copy of MDH gene which has replaced psbA2 gene in *Syenchocystis* genome. U corresponds to the total expression of MDH gene. The cells were grown under light at 30°C and two replicates were used to perform the experiment.

In MDH experiment, the expression of 23S was constant in all cases. MDH gene present in *Synechocystis* PCC 6803 genome showed expression in all cases (Native_MDH) even though in WT replicate 2 (Figure 15) the expression was lower. The extra copy of MDH gene (Engineered_MDH) showed expression in the MDH transformant while no expression can be seen in the WT. When the total expression of MDH gene (Universal_MDH) was checked, the bands present in MDH transformant were at a higher level than in the WT.

Relative levels of PEPc protein in WT and transformed cells

The protein gel shows that the same amount of total protein was loaded in both gels (see material and methods) (Figure 16). After the Western Blot was done, a band was visible which correspond to the PEPc protein (105-110 KDa) as it can be seen in Figure 17. The amount of PEPc protein in the PEPc transformant *Syenchocystis* is higher than WT *Syenchocystis* PCC 6803 (Figure 17 and 18-L).

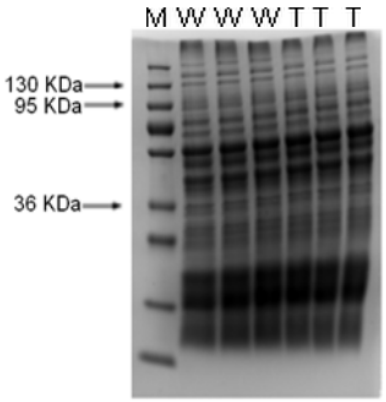

Figure 16. Protein gel loaded with protein extraction (approximately 1.5 mg·mL $^{-1}$) of WT *Synechocystis* PCC 6803 and the PEPc transformant *Syenchocystis*. M corresponds to marker. W corresponds to WT *Syenchocystis* PCC 6803 protein extraction. T corresponds to protein extraction of PEPc transformant.

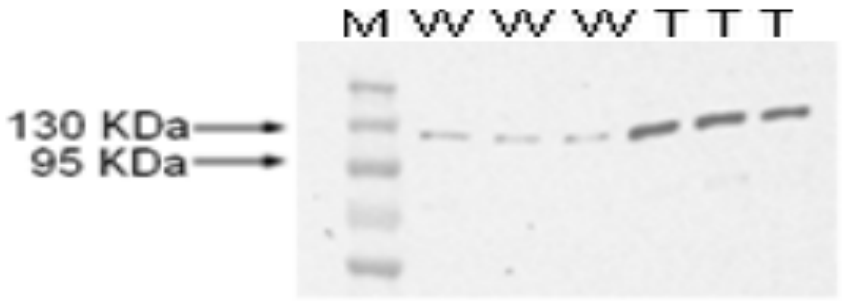

Figure 17. Nitrocellulose membrane after Western Blot when light condition was tested. M corresponds to Marker. W corresponds PEPc protein of *Syenchocystis* PCC 6803 (WT). T corresponds to PEPc protein of the PEPc transformant.

As it was done in semiquantitative RT-PCR light and darkness (24 hours) were tested. In both conditions, the amount of PEPc protein was higher in PEPc transformed than in WT (Figure 18). However, neither WT nor PEPc transformant showed any difference in level of PEPc protein when comparing cells in light and darkness (Figure 18).

Figure 18. Nitrocellulose membrane after Western Blot when light and darkness conditions were tested. M corresponds to Marker. W corresponds to PEPc protein of *Syenchocystis* PCC 6803 PCC (WT). p corresponds to PEPc protein of PEPc transformant. L corresponds to light while D corresponds to 24 hours of darkness. Two replicates were used in each condition.

Discussion:

PEPc-pBlueScript and MDH-pEERM

The overexpression of the genes has been carried out by inserting an extra copy of the native genes encoding PEPc and MDH in the chromosome of the cyanobacterium *Synechocystis* PCC 6803. The plasmids used had the upstream and downstream region of the psbA2 gene and by homologous recombination these additional copies replaced the psbA2 gene in cyanobacterial cells. Thus, these additional native copies are under control of the psbA2 promoter. PsbA2 gene belongs to the psbA family in which there are three members, psbA1, psbA2 and psbA3. PsbA2 and PsbA3 encode for the same protein, D1, which is one of the main proteins in the photosystem II (Mohamed *et al*. 1993). PsbA1 gene encodes for a protein similar to D1 (Salih and Jansson 1997). Even though psbA2 and 3 encode for the same protein, they differ from the 5' non-coding sequence (Mohamed *et al.* 1993). The psbA2 promoter seems to be shorter and stronger than psbA3 because psbA2 is more expressed than psbA3 under light conditions. Also, it has been demonstrated that psbA2 promoter is light intensity regulated. When psbA2 protein is non-functional, the expression of psbA3 gene is upregulated (Mohamed *et al*. 1993). For that reason, the replacement of psbA2 with another gene seems to do not affect growth of *Syenchocystis* considering that psbA3 will provide the necessary D1 protein (Mohamed *et al.* 1993 and Lindberg *et al.* 2010). In another study, psbA1 and psbA2 promoters were tested (Iwaki *et al.* 2006). It seems that psbA1 promoter is stronger than psbA2 despite de fact that no test was performed to see whether the replacement of psbA1 gene could have any negative effect on the cells. Therefore, the overexpression of MDH or PEPc by replacing psbA2 gene is suitable because: (*i*) light is required for both, the psbA2 promoter activity and *Syenchocystis* growth, and(*ii*) the replacement of psbA2 gene does not give any affect on the cellular growth.

Different problems have appeared during the amplification of the different parts of PEPc construction (PEPc gene, whole plasmid, Hp1 and Hp2). It is possible to amplify 5 Kb (whole plasmid) using Phusion Hot Spot II DNA Polymerase, even though in this study it has not been possible. It seems that the primers anneal to another part of the plasmid or to *E.coli* chromosome with more affinity than to the desired region of the plasmid, leading to obtain a 1.5 Kb product (Figure 10A). Some other primers were designed in order to amplify the plasmid in two halves. One half (Hp2) was easy to amplify while the other half (Hp1) presented some problems. There were different unspecific products which mean that the primers could anneal to unspecific parts of the plasmid or to *E.coli* chromosome as well. Two steps PCR was performed in order to try to improve the PCR product. This kind of PCR includes the annealing step into the extension one. Thereby, this PCR avoids the hybridization of the primers with unspecific DNA. As it can be observed (Figure 11) after 2- step PCR the desired band (around 2.5 Kb) was more brilliant than the others leading to obtain the desired product. In the case of MDH construction, the plasmid was changed to pEERM plasmid. That change is due to the complicate amplification procedure that the plasmid used for PEPc construction showed. Therefore, it was easier to change the plasmid to a smaller one which also has the sequences to work with BioBricks (pEERM). The amplification of the different parts for MDH construction was easier. Only some annealing temperatures had to be tested in order to amplify plasmid pEERM properly.

Different techniques such as BioBricks, overlap extension PCR or Gibson method (Shetty *et al.* 2008, Gibson *et al.* 2009 and Bryksin and Matsumura 2010) can be used to assemble different parts of a vector construction. Nevertheless, in this study Gibson method was chosen. This method was published in 2009 and it involves the assembly of DNA molecules in a single isothermal step. It requires an overlap between 20- 150bp which can be obtained by designing proper primers and doing a PCR (Gibson *et al.* 2009 and Collins *et al.* 2010). The primers that have been used for the amplifications contain a tail in order to get an overlap of 40 bp. The Master mix used for Gibson method contains three enzymes; 5' T5 exonuclease, Phusion DNA polymerase and *Taq* DNA ligase. The process is illustrated in Figure 19.

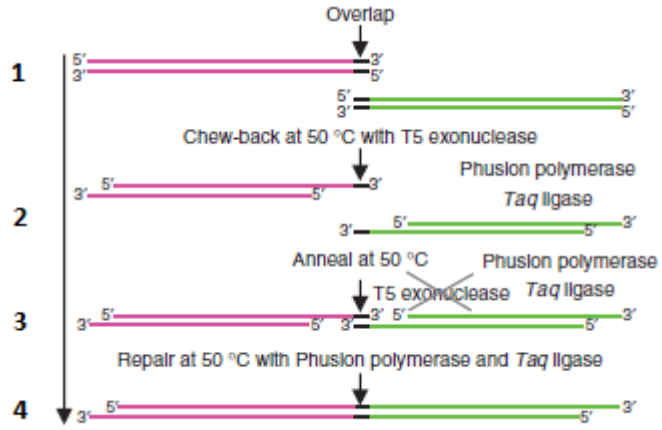

Figure 19. It represents the Gibson method which joins two genes. Step number 1: PCR where the DNA amplification is performed and therefore the overlap region is formed. Step number 2: T5 exonuclease remove nucleotides from 5' extreme of the double DNA strands. Small sequences of single strand DNA are formed which are complementary between genes. Step number 3 and 4: Phusion polymerase fills the gaps and *Taq* ligase joins the genes (Gibson *et al.* 2009).

The product of Gibson reaction was used to transform *E. coli* D5α strain. The cells were spreaded on LB agar plate which contains Chloramphenicol antibiotic. The cells which have incorporated the plasmid (with Chloramphenicol resistance cassette) are supposed to be the only ones that can grow in this media. However, it can be possible that during Gibson reaction, the assembly of the fragments was wrong or not all the plasmids had incorporated the PEPc or MDH sequence, leading to empty plasmids. Another possibility is that the cells had incorporated Hp2 part (which is where the Chloramphenicol resistant cassette is present) in PEPc experiment. These wrong constructions leaded to false positive colonies. This is what it could have happened to those colonies that grew on LB plates with Chloramphenicol but the plasmid construction was wrong when their sequences were verified.

Relative levels of PEPc or MDH transcript in WT and transformed cells

First, it was interesting to study the overexpression of PEPc and MDH at transcript level. There are several techniques (Northen blot, microarrays, semiquantitative RT-PCR and quantitative real time RT-PCR) which can be used to study the expression of the genes and therefore the amount of mRNA (Duggan *et al.* 1999, Marone *et al*. 2001, Bustin *et al.* 2005, Alvarez and Nourbakhsh 2011). In this study semiquantitative RT-PCR technique was used in order to study the level of expression of MDH or PEPc gene. RT-PCR could not be used because it can detect

the transcript but it cannot quantify the amount of mRNA. Quantitative real time RT-PCR could have been used, however, the technique is more complex. This technique shows in which number of PCR cycle the DNA amplification begins the exponential phase in the PCR reaction. Hence, the values can be analysed statistically leading to an accurate comparison among samples.

The functionality of designed primers was previously tested. Following, the PCR cycle which exponential phase began was checked (Table 2). The number of cycle is an important parameter in this technique because it allows to semiquantify the cDNA present in the samples (Marone *et al.* 2001). Using random primers during the reverse transcriptase (RT) reaction, the amount of a specific cDNA is depending on the amount of the specific mRNA. If a conventional PCR (30 cycles) was performed, the plateau phase would be achieved obtaining always approximately the same amount of DNA. Thus, the cDNA synthesized from mRNA could be detected but not the amount present in the samples.

The reference gene chosen in this study is 23S gene, which encodes the bigger ribosomal subunit. It has been reported that 16S gene (the small ribosomal subunit) is a proper reference gene in *Synechocystis* PCC 6803 when light and dark conditions are tested (Pinto *et al.* 2012). For that reason, it is assumed that 23S gene works properly as well under these conditions. This is in agreement with the results of the semiquantitative RT-PCR, because 23S is always constant while the other genes showed changes under the conditions tested in this study.

When the PCR cycles were tested, the lower cycle was from 23S, which means that the mRNA present in the cell is high (Table 2). This is explained by the fact that 23S is a constitutional gene. The native MDH or PEPc presents a higher cycle compared to 23S because they have lower expression than 23S. The additional copy of the native genes (engineering) presents a lower cycle than native MDH or PEPc gene. It can be caused by the psbA2 promoter regulation which is stronger when it is light exposed (Mohamed *et al*. 1993). Universal primers showed an intermediate cycle between native and engineering probably because in WT these primers detect the same transcript as native primers do, consequently lower cycle could not detect native MDH or PEPc transcript. In transformants, the lower cycle could detect the extra copy of MDH or PEPc gene expression but they could not detect expression of the native MDH or PEPc gene. In addition, higher cycle could yield results that do not correspond to the beginning of the exponential phase.

Before using the primers with the cDNA, the negative control was performed in order to detect any DNA present in the samples after the DNAse treatment which could lead to misunderstanding of results. The negative control was negative in all cases (data not shown), therefore all DNA present in the semiquantitative RT-PCR was amplified from cDNA.

On the one hand, the semiquantitative RT-PCR in PEPc experiment demonstrates constant 23S amplification (except for one sample). This means that the amount of cDNA used in the other PCRs with different primers was the same. Any variation in the results of the other PCRs observed was caused by the different conditions tested in this study. Since the results of the Universal_PEPc in WT and Native_PEPc in WT and PEPc transformant (Figure 14) show that after 24 hours in darkness the transcript level is lower compared to the light condition; it can lead me to the conclusion that the expression of PEPc gene present in *Syenchocystis* PCC 6803

genome (Native PEPc) is light regulated. It has been observed, that the gene expression of one PEPc isoform involved in leaf greening in sorghum plants (C_4 plant) is controlled by a phytochrome and follows the circadian rhythm (Thomas *et al.* 1990). Also, during light period in greening leaf of C_3 monocots of rice and wheat, a moderate increase in mRNA levels and activity of one PEPc isoform has been noticed. The same pattern was observed during the induction of CAM plants which the phytochrome seems to induce the PEPc activity. The latter activity is related to the expression of PEPc gene and therefore higher amount of PEPc transcpript was observed (Lepiniec *et al.* 1994). Since phytochromes are present in *Syenchocystis* PCC 6803 (Hübschmann *et al.* 2001) it might be possible that these photoreceptors play a role in the regulation of PEPc gene expression in this cyanobacterium.

Regarding the results for the inserted PEPc copy (Figure 14C); no expression of this gene was observed in WT. However, this gene was present in the PEPc transformant, as expected. This result confirms that the PEPc transformant incorporated the additional PEPc gene copy. The light regulation of the extra copy of PEPc gene inserted into the cyanobacterium (Engineering_PEPc) (Figure 14C) was due to the inserted extra copy of PEPc gene which was attached with the psbA2 promoter and it replaced the psbA2 gene (Figure 9A). Thus, although the promoter of the native PEPc is not known how strong is, it can be assumed that the new copy of the PEPc gene is more expressed under the presence of light than the native PEPc since the psbA2 promoter is a strong promoter (Mohamed *et al*. 1993).

The comparison of the total amount of transcript PEPc gene in WT and PEPc transformant (Figure 14 D) shows that under light's presence the amount of PEPc is not that high in the PEPc transformant compared to WT. However, in WT it can be noticed the remarkable decrease of transcript in darkness while in PEPc transformant that kind of decrease was not that noticeable. This can be explained by the total amount of PEPc transcript is higher in the PEPc transformant compared to WT. This was expected since the PEPc transformant has the additional PEPc copy under the control of a strong promoter.

On the other hand, in MDH experiment, the same results as in PEPc experiment were observed when under light condition. This result was also expected since the same cloning was performed. However, when the total amount of MDH was compared between the WT and the MDH transformant (Figure 15- U), it is noticeable that the level of the MDH transcript is higher in the transformant compared to the one in the natural strain. This evidence was not noticed under light in PEPc experiment.

Thus, thanks to the Semiquantitative RT-PCR results, it can be demonstrated that the insertion of an extra copy of the genes (MDH or PEPc) is transcriptional active in both cases.

Relative levels of PEPc protein in WT and transformed cells

The stained gel was a useful tool in order to make sure that the amount of proteins loaded from the different samples was the same (Figure 16). Therefore, if any difference in the amount of PEPc protein was observed among the different samples after the immunodetection, it would not be attributed to the fact that there are differences in the protein concentration among the samples. Thus, stained gel serves as a control. Another useful tool was the staining of the gel after the proteins were transferred into the nitrocellulose

membrane. This second gel control shows that the transfer of the protein in every sample was equal and therefore, the same amount of protein was transferred to the nitrocellulose membrane in every sample.

The results of the Western Blot under light presence show that the amount of PEPc protein in the PEPc transformant was higher than the one in the WT (Figure 17 and Figure 18-L). This result was expected since an extra copy of the native PEPc gene was inserted in PEPc transformant and the expression of this extra copy gene was active under light as it was demonstrated in the semiquantitative RT-PCR previously.

Even though the expression of PEPc gene in both cases, Native and Engineered PEPc, was down regulated under darkness (Figure 14-B and C), the amount of protein remained constant during darkness (Figure 18) meaning that the protein is stable in the both conditions. It could also be possible that the transcription of the gene is repressed under darkness but a longer exposure under darkness is required in order to observe effects at the protein level.

It is common that when the level of a protein is high, it can inhibit the expression of the gene leading to the repression of the transcription (Heintzen *et al.* 1997). Nevertheless, this cannot be attributed to PEPc because the high level of PEPc protein has not affected the transcription of the native PEPc gene in the PEPc transformant.

It is also known that the activity of PEPc is insignificant in darkness (Coleman and Colman 1980). This kind of reduction of the activity can be caused by several factors. Firstly, fluctuations have been observed in the internal pH in cyanobacteria under light and darkness. The internal pH under light is higher than under darkness (Coleman and Colman 1981). Secondly, it has also been shown that in green algae the PEPc activity is inhibited by aspartate, malate and other substances (Rivoal *et al.* 1998), in higher plants by malate and aspartate and in some plants by oxaloacetate as well (Lepiniec *et al*. 1994). In the thermophilic bacteria *Synechococcus volcanus* the PEPc activity is inhibited by malate and aspartate (Chen *et al*. 2002) and in *Coccochloris peniocystis* cyanobacterium PEPc activity seems to be repressed by citrate, oxaloacetate, malate and aspartate (Owttrim and Colman 1986). It is unlikely that oxaloacetate could contribute to repress the activity of PEPc in darkness since this intermediate metabolite is rapidly converted into other compounds such as aspartate or malate (O'Leray 1982 and Luinenburg and Coleman 1992). However, it seems that in *C.peniocystis* cyanobacterium, aspartate is the major stable compound formed after β-carboxylation by PEPc and the level of this compound remains high during darkness (Owttrim and Colman 1986). Also, the latter study has been demonstrated that the inhibition by aspartate, malate and oxaloacetate was higher when the pH was lower. Thus, the low pH that cyanobacteria, in general, have shown during darkness, the high level of aspartate in the same condition and the enhanced effect of aspartate, malate and oxaloacetate in low pH, could contribute to the repression of the activity of the PEPc protein under darkness in *C.peniocystis* and possibly in *Syenchocystis* PCC 6803.

In addition, some other factors could influence the activity of this enzyme under darkness. In *C.peniocystis* the Km of PEPc protein for PEP and HCO_3^- is high which agrees with PEPc from C_4 plants (Owttrim and Colman 1986). The fact that PEPc has high Km is because of the fact that cyanobacteria are able (during light) to accumulate inorganic carbon into the cell (CCM) (Price

2011) and therefore low Km value is not needed. Nevertheless, the accumulation of inorganic carbon in the cell is decreased under darkness (Miller and Colman 1980 and Coleman and Colman 1981) and therefore, the high Km value for this compound could regulates the PEPc activity under darkness. Also, it seems that the other substrate, PEP, is decreased as well during darkness (Owttrim and Colman 1986).

Consequently, a combination of several factors such as low pH, presence of inhibitors and the absence of substrates during darkness could explain the reduced activity of the PEPc protein in cyanobacteria. The repression of the activity of the protein might play a role in the down regulation of PEPc transcription under dark periods.

Even though darkness has not been tested in MDH experiment, it seems that the MDH enzyme is light regulated as well but in a different way than PEPc does. Recently, an unusual TCA cycle has been discovered in cyanobacteria (Steinhauser *et al.* 2012). The presence of isocitrate lyase and malate synthase enzymes has allowed that the TCA cycle is involved in the glyoxylate cycle and therefore it contributes to the respiratory pathway (Norman and Colman 1991). The glyoxylate cycle is important because it can regenerate oxaloacetate under darkness, when PEPc activity is repressed (Norman and Colman 1991). MDH could play an important role redirecting the carbon flux under light and darkness. It is suggested that under light presence, when PEPc is active and produces oxaloacetate, MDH does not oxidize malate into oxaloacetate. This effect is due to MDH is repressed by ATP (result from photophosphorylation) and high levels of NADH produced during photosynthesis contribute to the reduction of oxaloacetate (Norman and Colman 1991). However, during darkness the production of oxaloacetate via PEPc is decreased and then is when MDH oxidizes malate in order to form oxaloacetate which can be converted into other compounds needed.

It can be concluded that the insertion of an extra copy of the PEPc native gene (PEPc experiment) has led to obtain more protein. However, more time is required in order to verify if this protein is not active during darkness and if in the MDH transformant more protein is present.

Additional experiments:

Segregation

In this study, it has been demonstrated that an additional copy of the native MDH or PEPc gene was inserted. However, it is known that cyanobacteria have multiple chromosomes (Schneider *et al.* 2007) so it was unlikely that all the inserted copies were segregated in all the chromosomes. A PCR was performed in order to find if psbA2 gene was still present in any chromosome. Two PCR were performed amplifying two different fragments in the sequence of psbA2 gene. As it can be observed in figure 20, psbA2 gene was still present in the PEPc transformant. This means that psbA2 gene has not been completely replaced by the extra copy of PEPc gene in all chromosomes. In order to obtain 100% segregation, colonies are being continuously spreaded on BG11 with Chloramphenicol (the same procedure is being done in MDH transformant). Nevertheless, the colonies do not grown fast enough and the 100% segregation has not been achieved yet. More time is required but a further comparison of MDH and PEPc expression between the transformant (not 100% segregated) and the one that is 100% segregated would be an interesting point to examine.

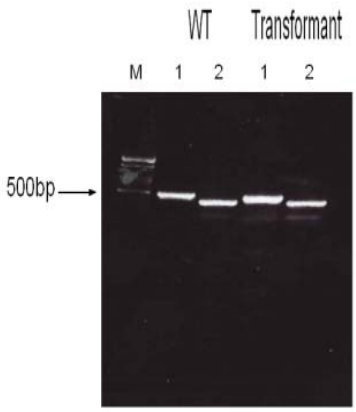

Figure 20. Agarose gel showing the PCR product of psbA2 gene. WT- Wild type *Syenchocystis* PCC 6803, transfromant- PEPc transformant, M- Marker, 1- one fragment of psbA2 gene, 2- another fragment of psbA2 gene.

Triple construct

The three first enzymes from the MOG pathway (*Phosphoenolpyruvate synthase* (PPSA), *Phosphoenolpyruvate carboxylase* (PEPc) and *Malate dehydrogenase* (MDH) were inserted into pEERM' plasmid which contains the upstream and downstream region of the psbA2 gene (Figure 21). The three genes will replace the psbA2 gene in *Synechocystis* PCC 6803.

All the gene amplifications were done using Phusion High-Fidelity Hot Spot II DNA polymerase protocol from Finnzymes while the assembly of the different parts was achieved using the Biobrick method. The restriction enzymes used were XbaI, SpeI and PstI.

PPSA was amplified using primers which contained XbaI (forward primer), SpeI and PstI (reverse primer) sequences (Appendix 2). Then, the plasmid pEERM´and the PPSA gene sequences were cut using restriction enzymes; one tube was mixed with 3 µg of pEERM', 3 µL of XbaI, 3 µL of PstI, 3 µL of SAP, 6 µL of fast restriction buffer and up to 60 µL dH_20. The other tube was mixed with 3 µg of PPSA gene (previously purified), 3 µL of XbaI, 3 µL of PstI and up to 60 µL dH_2O. Both tubes were then incubated for, at least, 30 min at 37°C. After that, PCR purification kit was used. The Gibthon: ligation calculator was used to calculate the 3:1 ratio in order to perform the ligase reaction. The ligase reaction was carried out using the Quick ligation kit from Biolabs. Then 10 µl of the ligation mix were used to transform *E.coli* as described in materials and methods.

The procedure to amplify the other two genes, PEPc and MDH was more complicated. Both genes contained the sequences which could be digested by the restriction enzymes used in this method. PEPc gene contains two sequences which can be cut by PstI enzyme while MDH gene contains the sequences that can be cut by XbaI and PstI. Then, if the amplification of the genes were carried out without any modification, the genes would have been cut in these sequences by the restriction enzyme digestion, leading to the obtaining of incomplete gene sequences. Hence, two mutagenesis were carried out in both genes. The mutagenesis allowed to replace one nucleotide without affecting the amino acid codon. Thus, the restriction enzymes could have not cut the gene sequences.

The mutagenesis were carried out using Phusion High-Fidelity Hot Spot II DNA polymerase protocol from Finnzymes but the PCR was run 18 cycles instead. The templates for the mutagenesis PCR were the plasmids mentioned in material and methods (PEPc-pBlueScript and MDH-pEERM). After each mutagenesis-PCR reaction, 40 µL of the PCR product was incubated with 1 µL of DpnI enzyme for 1 hour at 37 °C while the other 10 µL were kept on ice. Then, an agarose gel was run with both samples (with and without DpnI samples). DpnI digests sequences which are methylated. The plasmids used as templates were previously been introduced into *E.coli* and they were methylated; and after DpnI digestion the template used in the PCR was cut. Thus, when the agarose gel was processed, the appearance of the same weight bands were observed (with and without DpnI treatment), meant successful mutagenesis.

When both mutagenesis in PEPc and MDH genes were done, the genes were amplified using primers which were designed to have the sequences of recognition of the restriction enzymes (XbaI sequence in forward primer while SpeI and PstI in the reverse one) (Appendix 2) for further cut by restriction enzymes and assembly by ligation.

When PPSA was already in a vector (pEERM´) firstly the PEPc gene was inserted and latter the MDH gene. The insertions were carried out as described in the assembly of PPSA and pEERM'. After the insertion of a gene into the vector, *E.coli* was transformed. The plasmids from *E.coli* were each time obtained and digested with the proper restriction enzymes and an agarose gel with the digested products was run in order to make sure that the desired insertion had been incorporated. After all genes were inserted, the plasmid was sent for sequencing in order to verify that the three genes were present and the desired construction was correct. The last step was the transformation of *Syenchocystis* PCC 6803 (procedure described in material and

methods) but colonies have not appeared yet and more time was required in order to evaluate the overexpression of the three genes in this microorganism.

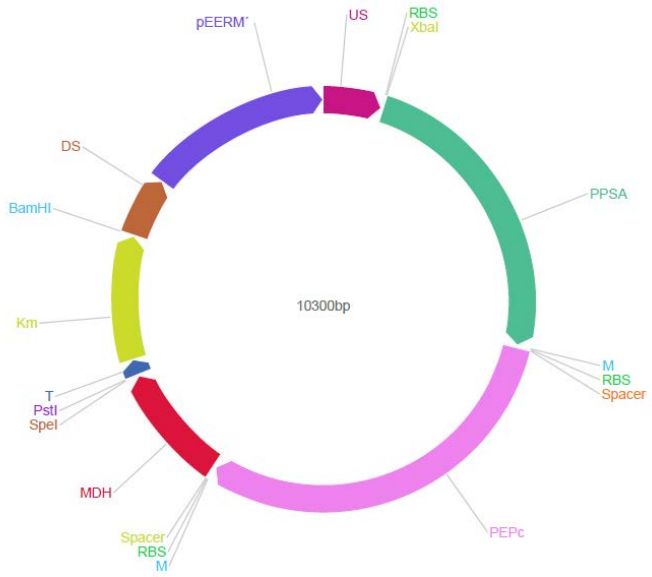

Figure 21. pEERM´plasmid with the three first genes in MOG pathway (PPSA, PEPc and MDH). The abbreviations are: BamHI- sequence which can be cut by BamHI enzyme, DS- downstream sequence of psbA2 gene, , Km- kanamycin resistant cassette, MDH- malate dehydrogenase gene, M- Mixed sequence (XbaI and SpeI), pEERM´- plasmid, PEPc- phosphoenolpyruvate carboxylase gene, PPSA- phosphoenolpyruvate synthase gene, PstI- sequence which can be cut by PStI enzyme, RBS- Ribosome binging site, SpeI- sequence which can be cut by SpeI enzyme , T-transcriptional terminator, US- upstream sequence of psbA2 gene and XbaI- sequence which can be cut by XbaI enzyme.

Further experiments:

In the case of the PEPc experiment, further experiments are required in order to examine if the PEPc protein in the PEPc transformant is active and therefore if it is able in fixing more inorganic carbon. Also, the activity of PEPc protein will be examined under different levels of CO_2.

In the case of the MDH transformant, it could be interesting to further evaluate if there is a higher amount of MDH protein present in the MDH transformant by using the Western Blot technique. However, due to the several isoforms present in nature, it is unknown if the antibodies available in the market would detect the isoform present in *Synechocystis* PCC 6803. Also, it would be interesting to study if the overexpressed protein is active and therefore if the catalysis of the reaction is more efficient.

When colonies with the triple construction will appear (see additional experiments) further experiments will be carried out. Gas chromatography mass spectroscopy would be a proper technique to use, in order to be able to examine if the three overexpressed proteins are active and therefore if more metabolites are produced compared to WT. In addition, this triple construction will be introduced into a strain which has a biofuel pathway. Then the amount of the final product of the biofuel pathway will be compared with the strain from the biofuel pathway and the strain with the triple construction and the biofuel pathway.

Finally, more enzymes from the MOG pathway will be introduced into *Syenchocystis* PCC 6803. Since the rest of the enzymes are not present in the *Synechocystis* PCC 6803 genome, the genes will be introduced individually into this microorganism. Then, the effect and activity of these enzymes will be evaluated while they are present into this microorganism. If the enzymes are active inside this cyanobacterium, the different genes will be grouped as operons and introduced into genome of *Synechocystis* PCC 6803. When all the enzymes will be cloned in different operons, they will be placed into different locations in the genome and the overexpression of the genes will be evaluated. At last, if all the construction can be achieved and more carbon fixation is occurring, it will be introduced into a strain which has incorporated a biofuel pathway and the production of the interest product will be examined.

Acknowledgments:

I would like to thank everyone in the cyanogroup in the Chemistry department-Ångström laboratory in Uppsala University. I especially appreciate the efforts of Elias Englund who has supervised almost every step of this project. I would also like to thank my supervisor Peter Lindblad who gave me the opportunity to work in this interesting project.

Finally I would like to gratitude my family and friends for their support, but primarily Elia Fernandez, Dimitra Peirasmaki and Alberto Rodriguez to report objective criticism and help me to write and improve this thesis.

Appendix:

Appendix 1. List of 63 cyanobacteria which their PEPc sequence was used to do the amino acid alignment represented in Figure 5.

Acaryochloris marina MBIC 11017, *Anabaena variabilis* ATCC 29413, *Arthrospira maxima* CS-328, *Arthospira platensis* NIES-39, *Arthospira platensis* str.paraca, *Arthrospira sp.* PCC 8005, *Crocosphaera watsonii* WH 8501, *Cyanobium sp.* PCC 7001, *Cyanothece sp.* ATCC 51142, *Cyanothece sp.* CCY0110, *Cyanothece sp.* PCC 7424, *Cyanothece sp.* PCC 7425, *Cyanothece sp.* PCC 8801, *Cyanothece sp.* PCC 8802, *Cyanothece sp.* PCC 7822, *Cylindrospermopsis raciborskii* CS-505, *Gloeobacter violaceus* PCC 7421, *Lyngbya sp.* PCC 8106, *Microcoleus chthonoplastes* PCC 7420, *Microcystis aeruginosa* NIES- 843, *Nodularia spumigena* CCY 9414, *Nostoc azollae* 0708, *Nostoc punctiforme* PCC 73102, *Nostoc sp.* PCC 7120, *Oscillatoria sp.* PCC 6505, *Prochlorococcus marinus str.* AS9601, *Prochlorococcus marinus str.* MIT 9202, *Prochlorococcus marinus str*. MIT 9211, *Prochlorococcus marinus str.* MIT 9215, *Prochlorococcus marinus str.* MIT 9301, *Prochlorococcus marinus str.* MIT 9303, *Prochlorococcus marinus str.* MIT 9312, *Prochlorococcus marinus str.* MIT 9313, *Prochlorococcus marinus str.* MIT 9515, *Prochlorococcus marinus str.* NATL1A, *Prochlorococcus marinus str.* NATL2A, *Prochlorococcus marinus subsp.pastoris, Prochlorococcus marinus subsp, Raphidiopsis brookii* D9, *Synechococcus elongatus* PCC 6301, *Synechococcus elongates* PCC 7942, *Synechococcus sp.* BL107, *Synechococcus sp.* CB0101, *Synechococcus sp.* CB 0205, *Synechococcus sp.* CC9311, *Synechococcus sp*. CC 9605, *Synechococcus sp.* CC9902, *Synechococcus sp.* JA-2-3B'a(2-13), *Synechococcus sp.* JA-3-3ab, *Synechococcus sp.* PCC 7002, *Synechococcus sp.* PCC 7335, *Synechococcus sp*. RCC307, *Synechococcus sp.* RS9916, *Synechococcus sp.* RS9917, *Synechococcus sp.* WH 5701, *Synechococcus sp.* WH 7803, *Synechococcus sp.* WH 7805, *Synechococcus sp*. WH 8102, *Synechococcus sp.* WH 8109, *Synechocystis sp.* PCC 6803, *Thermosynechococcus elongates* BP-1, *Trichodesmium erythraeum* IMS 101.

Appendix 2. Primers used to create the triple construction. Abbreviations correspond to: cod1-change of a nucleotide without changing the codon, cod2-change of the second nucleotide, for- forward primer, MDH- Malate dehydrogenase gene, PEPc- Phosphoenolpyruvate carboxylase gene, PPSA-Phosphoenolpyruvate synthase gene, rev- reverse. Colors correspond to: Green- sequence which can be cut by PstI restriction enzyme, Purple- Amino acids which have been changed (mutagenesis) in the genes, Red- Sequence which can be cut by SpeI restriction enzyme Yellow- Sequence which can be cut by XbaI restriction enzyme.

Primer	Sequence
PPSA_for	CTGAACATCGACAAATACATTAGTGGAGGTTCTAGAATGG
PPSA_rev	TCTCTACTGCAGACTAGTCTAGCCTAGGGCTTTTTCC
PEPc_for	TAGAGATCTAGATAGTGGAGGTTAGAGAATGAACTTGGCAGTTCCTG
PEPc_rev	TCTCTACTGCAGACTAGTTCAACCAGTATTACGCATTCC
MDH_for	TAGAGATCTAGATAGTGGAGGTTAGAGAATGAATATTTTGGAGTATGCTCCG
MDH_rev	TCTCTACTGCAGACTAGTTTAACCGTCGCTAACCATGG
PEPc_cod1_for	CCACTGCCTGCAATGATTTTTGGGC
PEPc_cod1_rev	GCCCAAAAATCATTGCAGGCAGTGG
PEPc_cod2_for	GGGAAATCTGCAATAGCCCCAGGGG
PEPc_cod2_rev	CCCCTGGGGCTATTGCAGATTTCCC
MDH_cod1_for	GGCGGGGAGAATCCTAGACTGATTG
MDH_cod1_rev	CAATCAGTCTAGGATTCTCCCCGCC
MDH_cod2_for	GCGAACTGCTTCCGCAGAAAGATGG
MDH_cod2_rev	CCATCTTTCTGCGGAAGCAGTTCGC

References:

Adachi K, Ohta K, Mizuno T (1994) Photocatalytic reduction of carbon dioxide to hydrocarbon using copper-loaded titanium dioxide. Sol Energ 53:187-190.

Atsumi S, Higashide W, Liao JC (2009) Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. Nat Biotechnol 27: 1177–1180.

Bar-Even A, Noor E, Lewis NE, Milo R (2010) Design and analysis of synthetic carbon fixation pathways. PNAS 107 (19) : 8889-8894.

Benschop JJ, Badger MR, Price GD (2003) Characterisation of CO_2 and HCO_3^- uptake in cyanobacterium *Synechocystis* sp. PCC 6803. Photosynth Res 77: 117-126.

Bryksin AV and Matsumura I (2010) Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. Biotech 48 (6): 463-465.

Bustin SA, V Benes V, Nolan T , Pfaffl MW (2005) Quantitative real-time RT-PCR – a perspective. J Mol Endocrinol 34:597-601.

Cendrin F, Chroboczek J, Zaccai G, Eisenberg H, Mevarech M (1993) Cloning, sequencing, and expression in Escherichia coli of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui.* Biochem 32:4308-4313.

Chen L, Omiya T, Hata S, Izui K (2002) Molecular characterization of a phosphoenolpyruvate carboxylase from a thermophilic cyanobacterium, *Synechococcus vulcanus* with unusual allosteric properties. Plant Cell Physiol 43 (2): 159-169.

Chollet R, Vidal J, O'Leary M (1996) Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. Annu Rev Plant Physiol Plant Mol Biol 47:273-298.

ClustalW2. www.ebi.ac.uk/tools/msa/clustalw2. Retrieved May 2013.

Collins B, Copley H, Emmrich P, Handley W, Hohmann A, Knott E, Masset P, Reeve B and Sanderson T (2010) BBF RFC 57: Assembly of bioBricks by the gibson method. The BioBricks Foundation.

Coleman JR and Colman B (1980) Demonstration of C_3-Photosynthesis in a bluegreen alga, *Coccochloris peniocystis*. Planta 149: 318-320.

Coleman JR and Colman B (1981) Inorganic carbon accumulation and photosynthesis in a blue green alga as a function of external pH. Plant Physiol 67: 917-921.

Deng M and Coleman JR (1999) Ethanol synthesis by genetic engineering in cyanobacteria. Appl Environ Microbiol 65 (2): 523-528.

Dresselhaus MS and Thomas IL (2001) Alternative energy technologies. Nature 414: 332-337.

Duggan DJ, Bittner M, Chen Y, Meltzer P ,Trent JM (1999) Expression profiling using cDNA microarrays. Nat Genet 21: 10-14.

Dvornyk V, Vinogradova O , Nevo E (2003) Origin and evolution of circadian clock genes in prokaryotes. PNAS 100 (5): 2495-2500.

Eisenhut M, Wolfgang R, Haimovich, Bauwe H, Kaplan A, Hagemann M (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. PNAS 105 (44): 17199-17204.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Sci 281: 237-240.

Figge RM, Cassier-Chauvat C, Chauvat F, Cerff R (2001). Characterization and analysis of an NAD(P)H dehydrogenase transcriptional regulator critical for the survival of cyanobacteria facing inorganic carbon starvation and osmotic stress. Mol Microbiol 39: 455–468.

Gibthon ligation calculator. www.django.gibthin.org/tools/ligcalc. Retrieved May 2013.

Gibson DG, Young L, Chuang RY, Venter JV, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6: 343-344.

Gietl C (1992) Malate dehydrogenase isoenzymes: Cellular locations and role in the flow of metabolites between the cytoplasm and cell organelles. Biochim Biophys Acta 1100: 217-234.

Goward CR and Nicholls D (1994) Malate dehydrogenase: A model of structure, evolution, and catalysis. Protein Sci 3: 1883-1888.

Hall MD, Levitt DG, Banaszak LJ (1992) Crystal structure of *Escherichia coli* malate dehydrogenase. J Mol Biol 226:867-882.

Heintzen C, Nater M, Apel K, Staiger D (1997) AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in Arabidopsis thaliana. Proc Natl Acad Sci USA 94: 8515-8520.

Honka E, Fabry S, Niermann T, Palm P, Hensel R (1990) Properties and primary structure of the L-malate dehydrogenase from the extremely thermophilic archaebacterium *Methanothermus fervidus*. Eur J Biochem 188: 623-632.

Hüebschmann T, Boerner T, Hartmann E, Lamparter T (2001) Characterization of the Cph1 holo-phytochrome from Synechocystis sp. PCC 6803. Eur J Biochem 268: 2055-2063.

Iijima S, Saiki T, Beppu T (1980) Physicochemical and catalytic properties of thermostable malate dehydrogenase from an extreme thermophile *Thermus flavus* AT-62. Biochim Biophys Acta 613: 1-9.

Ikeuchi M, Tabata S (2001) *Synechocystis* sp. PCC 6803 – a useful tool in the study of the genetics of cyanobacteria. Photosynth Res 70: 73–83.

Integrated microbial genome. www.img.igi.doe.gov/ Retrieved September 2012.

Ishijima S, Katagiri F, Kodaki T, Izui K,Katsukl H, Nishikawa K, Nakashima H, Ooi T (1985) Comparison of amino acid sequences between phosphoenolpyruvate carboxylases from *Escherichia coli* (allosteric) and *Anacystis nidulans* (non-allosteric):identification of conserved and variable regions. Biochem Biophys Res Commun 133: 436-441.

Iwaki T, Haranoh K, Inoue N, Kojima K, Satoh R, Nishino T, Wada S, Ihara H, Tsuyama S, Kobayashi H, Wadano A (2006) Expression of foreign type I ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39) stimulates photosynthesis in cyanobacterium *Synechococcus* PCC 7942 cells. Photosynth Res 88: 287–297.

Kai Y, Matsumura H, Inoue T, Terada K, Nagara Y, Yoshinaga T, Kihara A, Tsumuera K, Izui K (1999)Three-dimensional structure of phoshoenolpyruvate carboxylase: A proposed mechanism of allosteric inhibition. Proc Natl Acad Sci 96: 823-828.

Kasting JF and Siefert JL (2002) Life and the evolution of earth's atmosphere. Sci 296: 1066-1068.

Katagiri F, Kodaki T, Fujita N, Izui K, Katsuki H. (1985) Nucleotide sequence of the phosphoenolpyruvate carhoxylase gene of the cyanohacterium *Anacystis Nidulans*. Gene 38:265-269.

Lepiniec L, Vidal J, Chollet R, Gadal P, Cretin C (1994) Phosphoenolpyruvate carboxylase: structure, regulation and evolution. Plant Sci 99: 111-124.

Lindberg P, Park S, Melis A (2010) Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. Metab Eng 12: 70-79.

Lindblad P, Lindberg P, Oliveira P, Stensjo K, Heidorn T (2012) Design, Engineering, and Construction of Photosynthetic Microbial Cell Factories for Renewable Solar Fuel Production. AMBIO 41 (Supplement 2): 163–168.

Lucrecia Alvarez M and Nourbakhsh M (2011) RNA mapping protocols: Northern Blot and amplification of cDNA ends. Meth Mol Biol 700:199-220.

Luinenburg I and Coleman JR (1990) A requirement of phosphoenolpyruvate carboxylase in the cyanobacterium *Synechococcus* PCC 7942. Arch Mircobiol 154: 471-474.

Luinenburg I, Coleman JR (1992) Identification, characterization and sequence analysis of the gene encoding phosphoenolpyruvate carboxylase in *Anabaena* sp. PCC 7120. J Gen Microbiol 138: 685-691.

Macrogen Inc 2013. www.macrogen.com. Retrieved May 2013.

Maeda S, Badger MR, Price GD (2002) Novel gene products associated with $NdhD_3/D_4$-containing NDH-1 complexes are involved in photosynthetic CO_2 hydration in the cyanobacterium, *Synechococcus* sp. PCC 7942. Mol Microbiol 43: 425-435.

Mares I, Barthova I, Leblova S (1979) Purification and properties of phosphoenolpyruvate carboxylase from green leaves of maize. Collect Czech Chern Commun 44:1835-1840.

Mares I, Leblova S (1980) Phosphoenolpyruvate carboxylase from leaves of maize, sorghum and millet. Photosynthetica 14:25-31.

Marone M, Mozzetti S, De Ritis D, Pierelli L, Scambia G (2001) Semiquantitative RT-PCR analysis to assess the expression levels of multiple transcripts from the same sample. Biol Proced Online 3(1): 19-25.

McAlister-Henn L (1988) Evolutionary relationships among the malate dehydrogenases. TIBs 13:178-181.

Mikulasova D, Kollarova M, Miginiac-Maslow M, Decottignies P, Jacquot JP, Kutejova E, Mernik N, Egyudova I, Musrati R, Horecka T (1998) Purification and characterization of the malate dehydrogenase from *Streptomyces aureofaciens*. FEMS Microbiol Lett 159: 299-305.

Miller AG and Colman B (1980) Active transport and accumulation of bicarbonate by a unicellular cyanobacterium. J bacteriol 143 (3): 1253-1259.

Minarik P, Tomaskova N, Kollarova M, Antalik M (2002) Malate dehydrigenases- structure and function. Gen Physiol Biophys 21: 257-265.

Mohamed A, Eriksson J, Osiewacz HD, Jansson C (1993) Differential expression of the *psbA* genes in the cyanobacterium *Synechocystis* 6803. Mol Gen Genet 238:161-168.

Musrati RA, Kollarova M, Mernik N, Mikulasova D (1998) Malate dehydrogenases: distribution, function and properties. Gen Physiol Biophys 17: 193-210.

Nicholls DJ, Miller J, Scawen MD, Clarke AR, Holbrook JJ, Atkinson T, Goward CR (1992) The importance of arginine 102 for the substrate specificity of *Escherichia coli* malate dehydrogenase. Biochem Biophys Res Commun 189(2): 1057-1062.

Nishiyama M, Shimada K, Horinouchi S, Beppu T (1991) Role of threonine 190 in modulating the catalytic function of malate dehydrogenase from a thermophile *Thermus flavus*. Biol Chem 266 (22): 14294-14299.

Norman EG and Colman B (1991) Characterization of a malate dehydrogrenase in the cyanobacterium *Coccochloris peniocystis*. Arch Microbiol 156: 28-33.

Oligo Calc. www.basic.northwestern.edu/biotools/OligoCalc.html. Retrieved November 2012.

O'Leary MH (1982) Phosphoenolpyruvate carboxylase: an enzymologist's view. Plant Physiol 33:297-315.

Omata T, Price GD, Badger MR, Okamura M, Gohta S, Ogawa T (1999) Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. Strain PCC 7942. Proc Natl Acad Sci USA 96: 13571-13576.

Ocheretina O, Haferkamp I, Tellioglu H, Scheibe R (2000) Light-modulated NADP-malate dehydrogenases from mossfern and green algae: insights into evolution of the enzyme's regulation. Gene 258: 147-154.

Owttrim GW, Colman B (1986) Purification and Characterization of phosphoenolpyruvate carboxylase from a cyanobacterium. J Bacteriol 168:207-212.

Parmar A, Kumar Singh N, Pandey A, Gnansounou E, Madamwar (2011) Cyanobacteria and microalgae: A positive prospect for biofuels. Bioresour Technol 102 (11): 10163–10172.

Peralta-Yahya PP, Zhang F, Del Cardayre SB, Keasling (2012) Microbial engineering for the production of advanced biofuels. Nat 488: 320-328.

Pinto F, Pacheco CC, Ferreira D, Ferreira PM, Tamagnini P (2012) Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. PLoS ONE 7: Issue 4.

Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L (2004) Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl Acad Sci USA 101:18228-18233.

Price GD (2011) Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. Photosynth Res 109: 47-57.

Rivoal J, Plaxton WC, Turpin DH (1998) Purification and characterization of high- and low-molecular-mass isoforms of phosphoenolpyruvate carboxylase from *Chlamydomonas reinhardtii.* J Biochem 331: 201-209.

Rosgaard L, Jara de Porcellinis A, Jacobsen JH, Frigaard NU, Sakuragi Y (2012) Bioengineering of carbon fixation, biofuels, and biochemical in cyanobacteria and plants. J Biotech 162: 134-147.

Sakata T , Kawai T (1981) Heterogeneous photocatalytic production of hydrogen and methane from ethanol and water. Chem Phys Lett 8:341- 344.

Salih GF, Jansson C (1997) Activation of the silent *psbAf g*ene in the cyanobacterium *Synechocystis* sp Strain *6803 p*roduces a novel and functional D1 protein. The Plant cell 9:869-878.

Sallal AKJ and Nimer NA (1988) The intracellular localization of malate dehydrogenase in *Anacystis nidulans*. FEMS Microbiol Lett 50:151-155.

Shetty RP, Endy D, Knight TF (2008) Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2(5).

Shibata M, Ohkawa H, Kaneko T, Fukuzawa H, Tabata S, Kaplan A, Ogawa T (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. Proc Natl Acad Sci USA 98:11789-11794.

Shibata M, Ohkawa H, Katoh H, Shimoyama M, Ogawa T (2002a) Two CO_2 uptake systems in cyanobacteria: four systems for inorganic carbon acquisition in *Synechocystis* sp.strain PCC 6803. Funct Plant Biol 29: 123-129.

Shibata M, Katoh H, Sonoda M, Ohkawa H, Shimoyama M, Fukuzawa H, Kaplan A, Ogawa T (2002b) Genes essential to sodium-dependent bicarbonate transport in cyanobacteria- function and phylogenetic analysis. J Biol Chem 277: 18658-18664.

Smith AJ, London J, Stanier RY (1967) Biochemical basis of obligate autotrophy in blue-green algae and thiobacilli. J Bacteriol 94: 972-983.

Smith AA, Coomes MW, Smith TE (2008) Isolation and sequence of the phosphoenolpyruvate carboxylase gene of the marine cyanobacterium *Syenchococcus* PCC 7002. J Biol Sci 8 (8): 1261-1270.

Smith AA, Plazas MC (2011) *In silico* characterization and homology modeling of cyanobacterial phosphoenolpyruvate carboxylase enzymes with computational tools and bioinformatics servers. Am J Biochem Mol Biol 1 (4): 319-336.

So AKC and Espie GS (1998) Cloning, characterization and expression of carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. Plant Mol Biol 37: 205-215.

So AKC, MKay MJ, Espie GS (2002) Characterization of a mutant lacking carboxysomal carbonic anhydrase from the cyanobacterium *Synechocystis* PCC 6803. Plant 214: 456-467.

Solomon S, Quin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H (2007) Contribution of working group I to the fourth assessment report of the intergovernment panel on climate change. Cambridge University Press.

Soltani N, Zarrini G, Ghasemi Y, Shokravi S and Baftehchi L (2007) Characterization of a soil cyanobacterium *Fischerella* sp. FS 18 under NaCl stress. J Biol Sci 7: 931-936.

Steinhauser D, Fernie AR, Araujo WL (2012) Unusual cyanobacterial TCA cycles: not broken just different. Trends Plant Sci 9 (17): 503-509.

Tamoi M, Ishikawa T, Takeda T, Shigeoka S (1996) Molecular characterization and resistance to hydrogen peroxide of two fructose-1,6-bisphosphatases from *Synechococcus* PCC 7942. Arch. Biochem. Biophys 334: 27–36.

Thomas T, Cretin C, Vidal J, Keryer E, Gadal P, Monsinger E (1990) Light-regulation of phosphoenolpyruvate carboxylase mRNA in leaves of C4 plants: evidence for phytochrome control on transcription during greening and for rhythmicity. Plant Sci 69: 65-78.

Ting IP, Osmond CB (1973) Photosynthetic phosphoenolpyruvate carboxylases. Characteristics of alloenzymes from leaves of C_3 and C_4 plants. Plant Physiol 51:439-447.

Tomitani A, Knoll AH, Cavanaug CM, Ohno T (2006) The evolutionary diversification of cyanobacteria: Molecular-phylogenetic and paleontological perspectives. Proc Natl Acad Sci 103: 5442-5447.

Uedan K, Sugiyama T (1976) Purification and characterization of phosphoenolpyruvate carboxylase from maize leaves. Plant Physiol 57:906-910.

Wilks HM, Hart KW, Feeney R, Dunn CR, Muirhead H, Chia WN, Barstow DA, Atkinson T, Clarke AR, Holbrook JJ (1988) A specific, highly active malate dehydrogenase by redesign of lactate dehydrigense framework. Sci 242:1541-1544.