Faculty of Biological Sciences
University of South Bohemia in České Budějovice

Bachelor thesis

Construction of expression vector for cyanobacterium *Synechocystis* sp. PCC 6803

Jitka Kručinská

Supervisor: **Ing. Roman Sobotka, PhD.**

In České Budějovice, May 2007

Kručinská, J., 2007: Construction of expression vector for cyanobacterium *Synechocystis* sp. PCC 6803. Bc. thesis in English – 37 pages. Faculty of Biological Sciences, University of south Bohemia, České Budějovice, Czech Republic.

Annotation

The cyanobacterium *Synechocystis* sp. PCC 6803 is an important model organism used in the study of photosynthesis and related processes. The only disadvantage of *Synechocystis* for molecular biology studies is that it lacks a cloning/expression system. The aim of this work was to evaluate whether it is possible to use a mutated high copy version of the native *Synechocystis* plasmid for expression of a chosen gene in this organism.

Anotace

Sinice *Synechocystis* sp. PCC 6803 je významný modelový organismus pro studium fotosyntézy a s ní souvisejících porcesů. Pro molekulárně biologický výzkum zůstává hlavní nevýhodou *Synechocystis* nepřítomnost systému pro klonování a expresi. Cílem této práce je zjistit, zda je možné použít mutovanou verzi přirozeného plazmidu s vysokým počtem kopií pro expresi vybraného genu v tomto organismu.

I hereby declare that I have worked out this bachelor thesis by myself, using the cited literature.

In České Budějovice, 11 May, 2007

……

Jitka Kručinská

Acknowledgement

In the first place I owe a great deal to my parents and my sister who supported me in study with huge patience and love.

I'd like to thank all the great people around me who showed me that doing science doesn't mean working in a curious place with strange people, but that it can also be kind of fun and that scientists are ordinary people. Thanks to my supervisor for being in a good mood almost all of the time, to the head of our laboratory Martin Tichý, to the best technician Eva Prachová, to Panter (who had to listen to all my effusions on the late state of my work in the lab) and to all of the other people from Opaťák.

A big "thanks" also to Alison who corrected my English.

And of course I thank all of my friends that keep me smiling ☺

University of South Bohemia,
Faculty of Biological Sciences,
České Budějovice

Academy of Sciences of the Czech Republic
Institute of Microbiology,
Department of Autotrophic Microorganisms, Laboratory of Photosynthesis,
Třeboň

CONTENT

1. INTRODUCTION

Cyanobacteria

Cyanobacteria are amongst the most ancient organisms on Earth. The oldest known fossils, in fact, are cyanobacteria from *Archaean* rocks of Western Australia, dated 3.5 billion years old (Fig. 1.1). This may be somewhat surprising, since the oldest rocks are only a little older: 3.8 billion years old. This group of organisms has developed oxygenic photosynthesis in a manner similar to that in chloroplasts of higher plants. It is believed that their production of oxygen has been tremendously important in shaping the course of evolution and ecological change throughout earth's history. The oxygen atmosphere that we depend on was generated by numerous cyanobacteria photosynthesizing during the *Archaean* and *Proterozoic* Era.

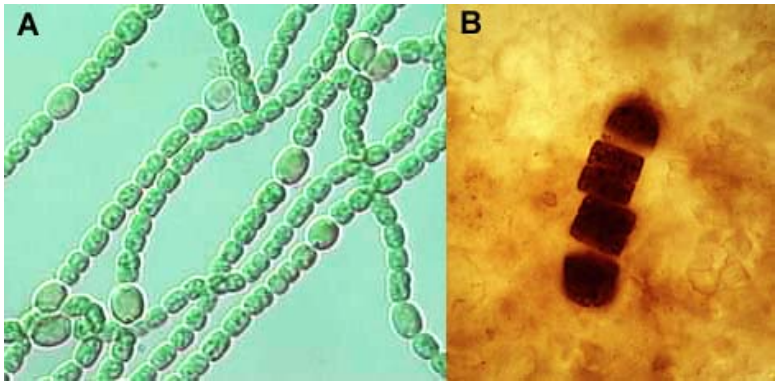

Fig. 1.1: Picture of modern and an ancient fossil of cyanobacteria: **A)** filamentous nitrogen fixing *Anabaena* cells **B)** fossil of proterozoic filamentous cyanobacteria found in central Australia (http://www.ucmp.berkeley.edu/precambrian/bittersprings.html).

Cyanobacteria are found in almost every conceivable habitat, from oceans to fresh water to bare rock to soil. They typically inhabit all extreme biotopes except locations that are too acidic. Many species of cyanobacteria (mostly: *Nostoc*, *Anabaena*, *Scytonema etc.*) form symbiotic relationships with lichens, fungi, orchids, ferns and gymnosperms. But you can also find them at the slime, on the fur of polar bears or South American sloths to which they impart a greenish tinge.

Phylogenetic position of cyanobacteria in the bacteria kingdom is still obscure, although recent analysis of ancient genes has indicated a genetic relationship with Gram-positive bacteria (Xiong *et al.*, 2000). However, similary to other prokaryotes, cyanobacteria do not have a nucleus, mitochondria, Golgi apparatus or an endoplasmatic reticulum. The most

noticeable formations in cyanobacteria are thylakoid membranes (Fig. 1.2) which include photosynthetic apparatus and light harvesting phycobilisomes (Fig.1.3).

There is plenty of evidence to suggest that cyanobacteria are the progenitor(s) of plant plastids - estimations show that about 4500 (18%) of all nuclear genes of plants might be derived from a cyanobacterial endosymbiont. Thus the results obtained from cyanobacteria are often directly applicable to plants. Furthermore, analyses on cyanobacteria are much easier than studies with higher plants. Cyanobacteria provide an excellent cellular model for the study of photosynthesis and related processes, nitrogen fixation, heterocyst formation, circadian clock, prokaryotic differentiation and metabolism.

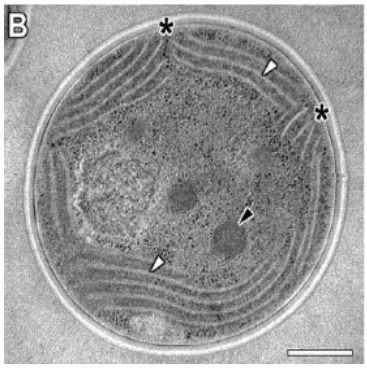

Fig.1.2: Electron microscopy picture of the cyanobacterium *Synechocystis* with visible thylakoid membranes (white arrowhead) and cytoplasmic inclusions (black arrowhead; Van de Meene *et al*., 2005).

Model cyanobacterium *Synechocystis* 6803

Synechocystis sp. PCC 6803 (*Synechocystis*) (Fig. 1.4) has become a model cyanobacterium that is used by scientists all around the world. The strain was isolated from a fresh water lake, and was deposited in the Pasteur Culture Collection (PCC) in 1968, and in the early 1980s the strain was recognized to be spontaneously transformable, to integrate foreign DNA into its genome by homologous recombination (allowing targeted gene replacement), and to be able to survive and grow in a wide range of conditions. For example, *Synechocystis* can grow in the absence of photosynthesis if a suitable fixed-carbon source such as glucose is provided (Ikeuchi and Tabata, 2001; Williams, 1988; Grigorieva and Shestakov, 1982; Rippka *et al.*, 1979).

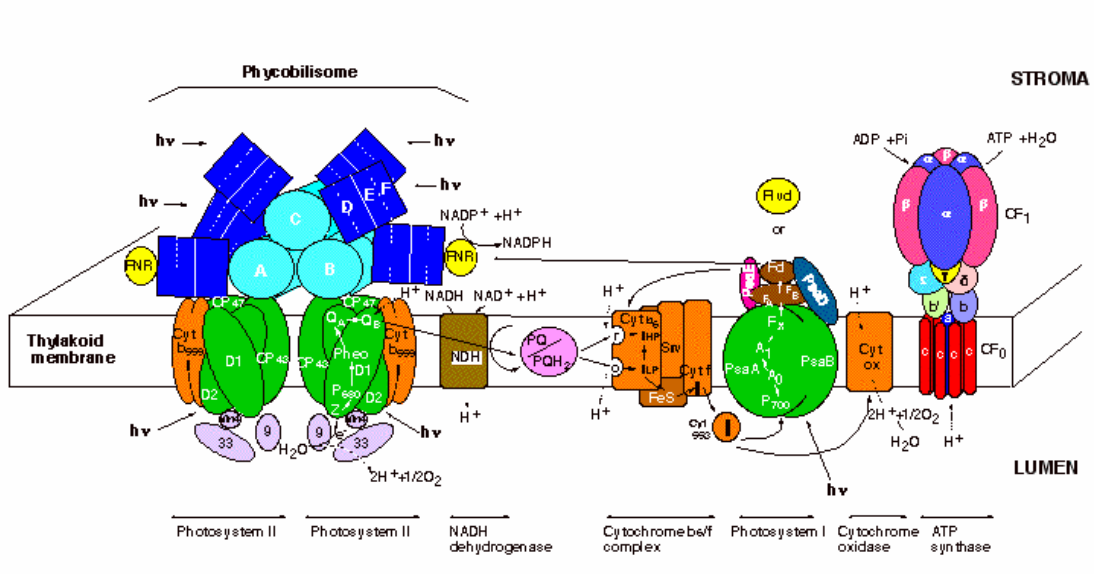

Fig. 1.3: Schematic view of the main protein complexes localized in thylakoids in cyanobacteria (Bryant, 1994).

Perhaps most significantly, *Synechocystis* was the first photosynthetic organism for which the entire genome was sequenced. In 1996, the Kazusa DNA Research Institute finished the genomic sequence of this organism (http://www.kazusa.or.jp/cyano/cyano.html), which makes it the fourth organism to have its genome completely sequenced (Kaneko *et al*., 1996).

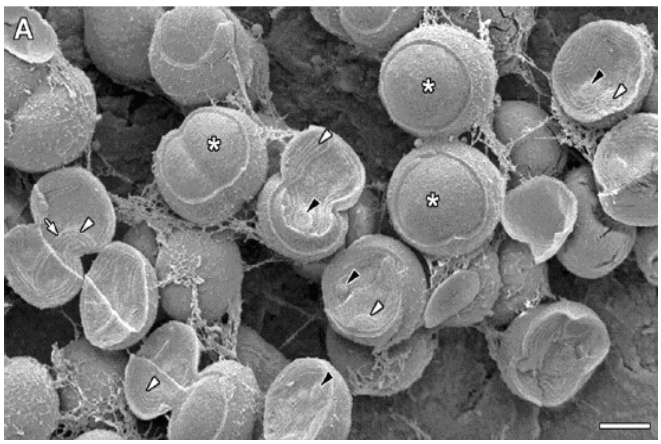

Fig. 1.4: Electron microscopy picture of high-pressure frozen and freeze-fractured *Synechocystis* cells. Outer membrane and the cytoplasmic membrane (white asterisks), thylakoid membranes (white arrowheads), fracture through the cytoplasm (white arrow), cytoplasmic inclusions (black arrowheads; Van de Meene *et al.*, 2005).

This information, coupled with the facile gene deletion strategy that is available for this organism has made *Synechocystis* the cyanobacterium of choice for research in many groups. Deletion mutants lacking a specific gene have been created and analyzed for many different genes. For example, *Synechocystis* is used to study pigment (carotenoid and chlorophyll) synthesis and its regulation, tocopherol synthesis, carbon metabolism, respiration, photosynthesis and a host of other processes (Komenda *et al.*, 2004; Raúl *et al.*, 2003; Bedu *et al.*, 1997). As was already noted, the information gained is not only valuable for understanding the physiology of cyanobacteria, but also for that of other organisms such as plants.

On the other hand, the remaining disadvantage of *Synechocystis* for molecular biology studies is the lack of cloning/expression system comparable with those used in *Escherichia coli* (*E. coli*). As this is the topic of my work, some aspects of expression systems will be described below.

Expression systems

The expression of cloned genes is one region of molecular cloning where investigators can do everything perfectly and still see the experiments fail. The reasons for this are clear. Too little is known about the folding mechanisms of proteins in different organisms to predict which host-vector system might be the best for a given protein, and few methods are available to increase the efficiency of folding or to prevent the aggregation, denaturation, and/or degradation of foreign proteins expressed in environments that are unnatural to them. In the absence of any useful guiding principles, the expression of every gene presents a unique set of problems, which must be solved empirically.

Five major expression systems have been developed to a point where they are reasonably portable from one laboratory to another, where the necessary components and vectors are readily available from commercial sources or public repositories, and where success can be documented with more than a handful of proteins. The host cells for the systems used until know are bacteria *E coli* and *Bacillus subtilis*, yeasts *Saccharomyces cerevisiae* and *Pichia pastoris*, cultured insect cells, and cultured mammalian cells (Sambrook and Russel, 2001).

.

Expression vectors

Expression systems in bacteria are based on plasmids, self-replicating, double-stranded, circular DNA molecules that are maintained as independent extra-chromosomal entities (Fig. 1.5). Some plasmids carry information for their own transfer from one cell to another (F plasmids), others encode resistance to antibiotics (R plasmids), others carry specific sets of genes for the utilisation of unusual metabolites (degradative plasmids), and some have no apparent functional genes (cryptic plasmids). Plasmids can range in size from less then 1kb to more than 500 kb. Each plasmid has a sequence that functions as the origin of DNA replication; without this site, it cannot replicate in a host cell.

Some plasmids are represented by 10 to 100 copies per host cell; these are called high-copy-number plasmids. Others maintain 1 to 4 copies per cell and are called low-copy number plasmids. Seldom does the population of plasmids in a bacterium make up more than approximately 0.1 to 5.0% of the total DNA. Some plasmids, because of the specificity of their origin of replication, can replicate in only one species of host cell. Other plasmids have less specific origins of replication and can replicate in a number of bacterial species. These plasmids are called narrow- and broad-host-range plasmids, respectively.

However, naturally occurring (unmodified or unengineered) plasmids often lack several important features that are required for a high-quality cloning vector. These features are:

1) Small size – which is necessary because the efficiency of transfer of exogenous (foreign) DNA into *E.coli* decreases significantly with plasmids that are more than 15 kb long.
2) A choice of unique (single) restriction endonuclease recognition sites into which the insert DNA can be cloned.
3) One or more selectable genetic markers for identifying recipient cells that carry the cloning vector-insert DNA-construct. These plasmid-encoded markers provide specific resistance to (i.e., the ability for growth in the presence of antibiotics such as the kanamycins, ampicillin or the tetracyclines).

In other words, plasmid cloning vectors have to be genetically engineered (Glick; Pasternak, Molecular Biotechnology, 1994).

Fig. 1.5: Electron microscopy picture of plasmids, 100 000x magnification (photo from universe-review.ca).

Expression plasmids in *Escherichia coli*

Escherichia coli is first system of choice for the expression of many heterologous proteins, it is also the best studied and has a large availability of all kinds of vector plasmids. However, the expression systems used today with amazing amount of recombinant proteins often yielded are result of extensive effort during the last three decades.

The first plasmids used in E. coli as cloning vectors either replicated poorly, were too big or they carried unsuitable selectable markers, and none of them contained more than two restriction sites that could be used for cloning (Bolivar *et al*., 1977). There are advantages in reducing the size of plasmid vectors to a minimum. Plasmid copy number, stability, and transforming efficiency all increase as the size of their DNA is reduced. Smaller plasmids can accommodate larger segments of foreign DNA before their efficiency begins to deteriorate.

In the late 1970s and early 1980s derivates of pBR322 were constructed that lacked the ancillary sequences involved in the control of copy number and mobilisation. Unfortunately, the first generation of high-copy-number plasmids suffered from a major defect: Foreign DNA sequences had been inserted only to a limited number of restriction sites located within the "natural" sequences used to construct the plasmid. These plasmids were soon replaced by a revolutionary series of pUC vectors which contain a closely arranged series of synthetic cloning sites, termed polylinkers, multiple cloning sites, or polycloning sites, which consist of banks of sequences recognised by restriction enzymes. In most cases, these restriction cloning sites are unique; i.e., they are not found elsewhere in the plasmid vector. Such arrays of recognition sequences provide a vast variety of targets that can be used singly or in combination to clone DNA fragments generated by cleavage with a large number of restriction enzymes.

The latest phase of the construction of plasmid vectors has involved the incorporation of ancillary sequences that are used for a variety of purposes, including generation of single-

stranded DNA templates for DNA sequencing, transcription of foreign DNA sequences in vitro, direct selection of recombinant clones, and expression of large amounts of foreign proteins. (Sambrook and Russel, 2001)

Plasmids used nowadays

I have chosen the pET expression system (Fig. 1.6) as an example to describe the modern cloning/expression system in *E. coli*.

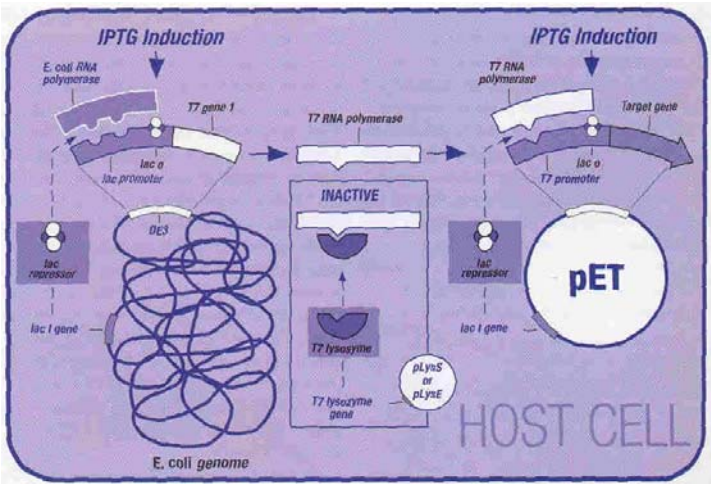

Fig. 1.6: Schema of the pET expression system (Novagen catalogue).

Within this group of vectors, the target gene is cloned under the control of a strong bacteriophage T7 promoter and the prepared vector is transferred into expression hosts containing a chromosomal copy of the T7 RNA polymerase gene under *lacUV5* control. The *lacUV5* promoter is a mutated version of the *lac* promoter whose basal activity is dramatically less sensitive to intracellular levels of the cyclic AMP¡Xa molecule known to participate in derepression of the *lac* promoter. The *lacUV5* promoter is less sensitive or less leaky at higher culture densities. Derepression or induction of the system is accomplished by the simple addition of IPTG (isopropyl-beta-D-thiogalactopyranoside), a compound used as an analogue of a lactose metabolite. Adding of IPTG results in the displacement of the Lac repressor protein (LacI) from promoters servicing both the T7 gene on the chromosome and the target gene on the pET construct. T7 RNA polymerase is so selective and active that almost all of the cell's resources are converted to target gene expression. The desired product can comprise more than 50% of the total cell protein a few hours after induction.

Alternatively, target genes are initially cloned using hosts that do not contain the T7 RNA polymerase gene, so they are virtually "off" and cannot cause plasmid instability due to the production of proteins potentially toxic to the host cell. There exist more than 40 types of the

pET system; more than 10 different *E. coli* host strains and many other companion products designed for the efficient detection and purification of target proteins.

Expression systems in cyanobacteria

The transfer of DNA to *Synechococcus* sp. by transformation, first reported over 30 years ago, heralded the beginning of genetic systems in cyanobacteria. Almost twenty years ago the first conjugation system for cyanobacteria was described (Thiel, 1994).

Unlike most bacteria with just one copy of the chromosome, cyanobacteria contain about ten copies of a single chromosome (Elhai, 1993). For this reason, some specific places on the chromosome are frequently used for protein (over)expression. In such cases a native promoters that are known to be highly expressed under specific conditions are often chosen. For example, the strong promoter of the *psbA* gene coding for the core protein of photosystem II is frequently used in *Synechocystis*. As this cyanobacterium has several copies of the *psbA* genes, replacing one of the copies by a foreign gene itself has no effect on cell growth. On the other hand, although it is possible to increase the expression of the *psbA* promoter by light, potential regulation of this promoter has not been reported (Lagarde and Vermaas, 1999). Other native promoters tested for controlled (over)expression in cyanobacteria were from the phycocyanin gene *cpcB2A2*, of *Calothrix* regulated by the colour of light (Conley *et al*., 1988); *nirA* promoter that can be switched on or off by using nitrate or ammonium as a source of nitrogen in the media or the *petJ* gene coding for plastocyanin that appears to be quite well regulated by copper availability (Qi *et al.*, 2005; Dühring *et al.*, 2006).

Use of non-native promoters in cyanobacteria has been also reported. In *Synechococcus* PCC 7942 strain, the lambda P_L promoter with a temperature-sensitive lambda *cI* repressor gene (*cI*857) was fused to the *cat* gene. At 30°C the strain produces an active *cI* repressor that inhibits transcription of *cat*. At 42°C the repressor is inactivated and more *cat* is synthesized (Friedberg, 1988). Unfortunately, cyanobacteria usually do not grow well at temperatures between 39-42°C, and that limits the potential use of this promoter. The *tac* promoter of *E.coli* (a *trp-lac* hybrid promoter), without the *lacI* repressor gene, was also able to promote high levels of transcription of the *cat* gene in *Synechocystis* (Ferino and Chauvat, 1989); however the *tac* promoter is poorly repressed and even in *E.coli* tends to be leaky (Brosius, 1988).

Despite intensive study of various cyanophages, including lysogenic strains no viral vector has been found capable of transducing any cyanobacterium (Sarma and Kaur, 1977), so the

only vectors that are currently available for construction of an expression system are plasmid. Plasmids have been observed in many species of cyanobacteria, though their biological role is not yet understood and they are often not suitable for replication in several strains of interest, such as *Synechocystis* (Houmard & Tandeau de Marsac, 1988). Therefore, promiscuous cloning vectors capable of autonomous replication both in cyanobacterial cells and *E.coli* have been prepared in the last two decades. Such vectors can be sorted into two groups: a) the first group includes shuttle vectors that contain a replicon of a cyanobacterial endogenous plasmid and a replicon of an *E.coli* plasmid (Buikema, Haselkorn 2001). An example of an indigenous plasmid is pDU1 from *Nostoc* sp. It has been used for constructing a number of shuttle vectors for transforming *Anabaena* sp. (Thiel, 1994). b) The second group is based on a promiscuous plasmid RSF1010 which could propagate in a wide range of microorganisms, including a limited number of unicellular and filamentous cyanobacteria (Houmard and Tandeau de Marsac, 1988). This makes RSF1010 the most commonly used expression vector efficiently transferred between hosts by conjugation. Some host-vector systems for gene cloning have been established by using restriction fragments from *Synechocystis* low-copy number plasmids pCA2.4 and pCB2.4 to construct shuttle vectors; however, these vectors were not stable in the transformed cells (Chauvat *et al.*, 1986).

Unfortunately, all described plasmids above are low-copy-number plasmids, and expression is generally much lower than can be achieved in *E. coli* by using high-copy-number plasmids. However, several native high-copy-number plasmids were reported in cyanobacteria, which could be good candidates for construction of cloning/expression vectors as they provide the possibility of isolation, restriction and ligation in a similar way to *E.coli*. In thermophilic cyanobacterium *Synechococcus* sp. the MA4 plasmid was found with a copy number of about 350 per cell (Miyake *et al.*, 1999) and in *Synechocystis* a high copy variant of the native plasmid pCC5.2 was reported (Xu and McFadden 1997; see below). No results from their use as vectors have been published so far.

Plasmids in *Synechocystis* 6803

Synechocystis cells are known to contain four large plasmids pSYSM (125 kb), pSYSA (119 kb), pSYSX (106 kb) and pSYSG (45 kb) in addition to three small plasmids pCC5.2 (5,2 kb), pCA2.4 (2,4 kb), pCB2.4 (2,4 kb). All these plasmid have been fully sequenced and annotated (Kaneko *et al.*, 2003). For the large plasmids (pSYSM, pSYSA, pSYSG, pSYSX) the total number of potential protein-encoding genes finally assigned was 397 and 12 potential genes

were predicted for the three plasmids. Among those potential genes 122 have homologues in different plasmids and 48 genes showed sequence similarity to chromosomal genes. A surprisingly high amount of genes from these plasmids (142, 36%) are not similar to any known genes. In pSYSM there genes for glycosyl transferase families and homologues of polysaccharide transporter were found that are suggested to be involved in the biosynthesis and transport of unknown exopolysacharides. Other revealed genes were gene *ndhK2*, which codes for a subunit of NADH dehydrogenase and four genes presumably encoding for transcriptional regulators. Interestingly, *parA* and *parB* genes involved in partitioning of plasmids during cell duplication were identified in each of the four large plasmids but not in the three small plasmids (Kaneko *et al.,* 2003).

pCC5.2 plasmid

Plasmid pCC5.2 is the object of this project and will be described more deeply. The entire pCC5.2 plasmid contains 5214bp and includes six open reading frames (ORF) starting with ATG that encode peptides larger than 90 amino acids (aa). The Shine-Dalgarno sequence and the Pribnow box ("-10" region) for a prokaryotic promoter were found before each ORF. ORF A (encoding 117 aa), ORF B (971 aa), ORF C (93 aa) and ORF D (91 aa) are located on the plus strand and ORF E (105 aa) and ORF F (97 aa) are on the minus strand (Fig. 1.7).

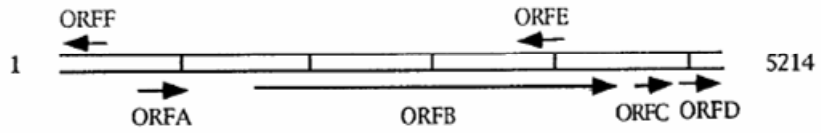

Fig. 1.7: Schema of ORFs orientation on plasmid pCC5.2.

Among the six open reading frames in pCC5.2, only the products of ORF B and ORF C were found to have homologous counterparts in the protein database. ORF B codes for a large polypeptide of 971 amino acids, and it has three known homologous proteins, all encoded by cyanobacterial plasmids from the following strains: *Agmenellum quadruplicatum*, *Anacystis nidulans*, *Microcystis aeruginosa*. ORF B was suggested to be a replication protein in *A. nidulans*, from *A. quadruplicatum* was also hypothesized to be important for plasmid replication.

ORF C encodes a smaller polypeptide of 93 aa. Two homologous peptides were detected. The first is encoded in plasmid pMA1 from cyanobacterium *Microcystis aeruginosa* and the second is encoded by pCA2.4 from the same *Synechocystis* 6803 strain as pCC5.2 comes

from. The identities between the product of ORF C of pCC5.2 and that of pMA1 and pCA2.4 are 42.1 and 41.1%, respectively. The functions of hypothetical products from these small ORFs are unknown so far (Xu and McFadden 1997).

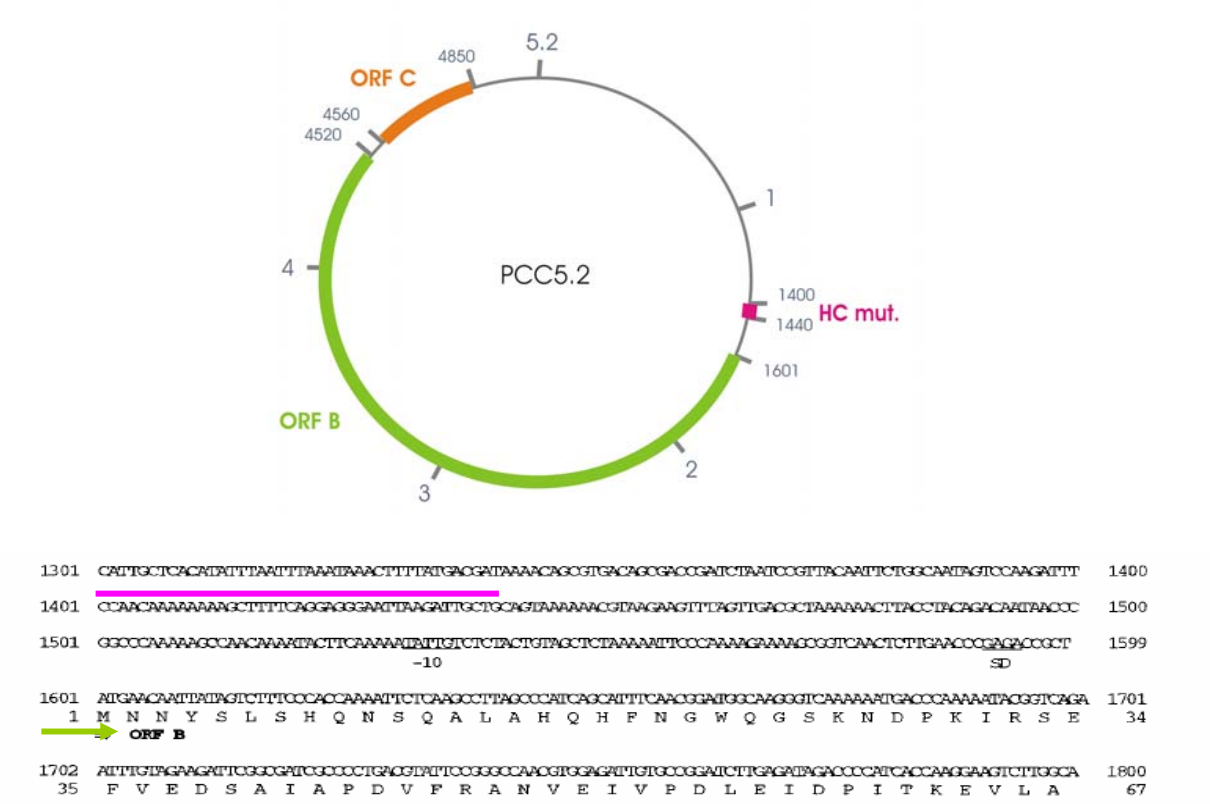

Fig. 1.8: Plasmid map with localized High copy mutation and part of the plasmid sequence with an exactly marked 40 base deletion and start of ORF B.

The laboratory of photosynthesis, IMB Trebon, obtained the *Synechocystis* strain bearing this pCC5.2 plasmid with a high number of copies per cell (HC plasmid) from Dr. Dexter A. Chisholm (Du Pont, USA). The mutant plasmid is produced in the cell in quantities comparable to the chromosomal DNA and is easily visible on the gel after purification of total DNA (Fig. 1.9). This enables simple isolation and processing of the plasmid by methods used for plasmid engineering in *E. coli*. The mutation responsible for the high copy phenotype has been localised as a deletion of 40 bp in the position of 1400 bp and 1440 bp. (Martin Tichy, unpublished, Fig. 1.8). The mutation behaves as a dominant mutation, *Synechocystis* wild type cells transformed by the HC plasmid or DNA fragment contains the HC mutation gain HC phenotype. Accidentally, a useful strain lacking the pCC5.2 plasmid with no obvious phenotype was obtained (Martin Tichy, unpublished).

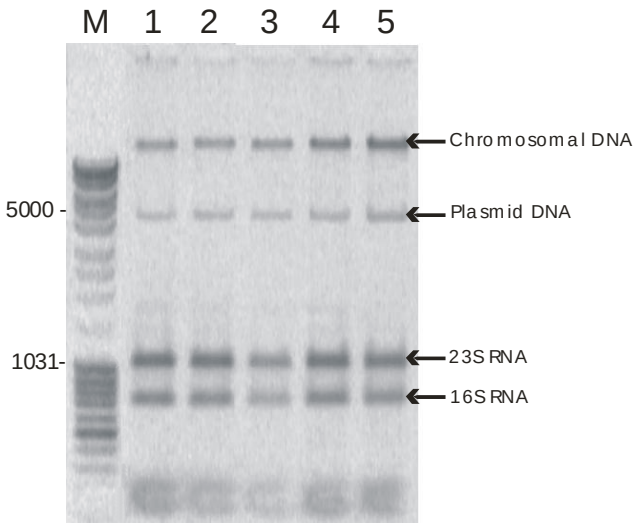

Fig 1.9: Purification of total DNA from five different colonies of HC strains *Synechocystis*.

2. AIMS OF THE PROJECT

An expression/cloning vector system for cyanobacteria would be an extremely useful tool for molecular studies concerning photosynthesis. Proteins participating in photosynthesis usually need special maintenance; some of them need thylakoid membranes in which to integrate in or cofactors like chlorophyll and carotenoids. The high copy version of the PCC5.2 plasmid is the hot candidate for a desired expression vector and there is an urgent need for its characterization.

The aim of my work has been to evaluate whether it is possible to use the pCC5.2 plasmid for the expression of chosen genes in *Synechocystis*. Particularly, it has been proposed to:

- develop protocols for manipulation with the HC plasmid
- analyze the plasmid copy number in the cell
- add a selectable antibiotic resistence gene to pCC5.2
- insert a model gene to the modified plasmid and analyze expression

3. EXPERIMENTAL PROCEDURES

Cell cultures growth

Wild type (WT) *Synechocystis* and mutants were grown in liquid BG-11 medium (Rippka *et al.* 1979) at 30°C and 40 μmol m^{-2} s^{-1} of light on a rotary shaker. Strains were grown photomixotrophicaly supplemented by 5 mM glucose, unless photoautotrophic conditions are indicated. Solid media were supplemented with 1.5% (m/v) agar, 0.3% (m/v) sodium thiosulfate, and 10 mM TES (Hydroxy-1,1-bis(hydroxymethyl)ethyl)amino-ethanesulfonic acid)/NaOH buffer, pH 8.2.

Polymerase chain reaction (PCR)

The PCR mixture (usually 25 μl volume) contains 5 μl of 5x FastStart buffer (5U/μl, Roche); 0,5 μl of 10 mM dNTPs; 0,5 μl of each primer (20 μM stock), 0,5 μl of Phusion polymerase (2U/μl, Finzyme) and 1 μl of template DNA. If not otherwise stated in the results section, these conditions were used (PTC-100 cycler, MJ Research):

Denaturation	Annealing	Polymerization	Number of cycles
94°C / 60 sec	58°C / 45 sec	72°C / 1min *	30 cycles

* for PCR fragments of about 1000 bp

PCR samples were separated on a standard horizontal electrophoresis in 1.0% TBE (45mM Tris-borate, 1mM EDTA) agarose gel and were stained with ethidium bromide.

Real-time quantitative polymerase chain reaction (RT-qPCR)

RTq-PCR was used to characterize copy number of the HC plasmid pCC5.2. The PCR product in the reaction was quantified using the Locked Nucleic Acid (LNA) fluorescent probe bought from Roche (Switzerland). These probes are only 8 or 9 bases long and the high probe-melting temperatures characteristic for longer probes are retained by changed nucleotide chemistry (Fig. 2.1). By PCR, only a short ~60 bp fragment is amplified which then hybridizes with the probe.

Primers were designed and ordered by using the Roche web tool (www.roche-applied-science.com).

The RT-qPCR reaction contains (19 µl in a 0.20 ml microfuge tube): 10 µl FastStart TaqMan ProbeMaster (Roche); 2 µl of each primer (20 µM stock); 5 µl template DNA and 0.2 µl UPL (Universal Probe Library) probe. PCR was run on Rotor-Gene RG-3000 (Corbett research) using the steps below:

Polymerase activation	Denaturation	Annealing	Polymerization	Number of cycles
95°C / 10 min	95°C / 20 sec	60°C / 30 sec	72°C / 20 sec	40 cycles

For detection of both *ORF-B* and *slr0554* genes, the same UPL probe (human # 37) was used. Primers are listed in Tab. 2.1.

Label	**Sequence**	**Probe**
slr0554F	GGGCACGGGCAAAACGGAATTT	# 37 TCTTCTTCT
slr0554R	TCGATACACAGACCAAAACCCC	
ORFB-F	GGGCACGGGAAAGACGGAAGCG	
ORFB-R	TCAATGCACAATCCGATTCCTA	

Tab. 2.1: Oligonucleotides used for RTq-PCR.

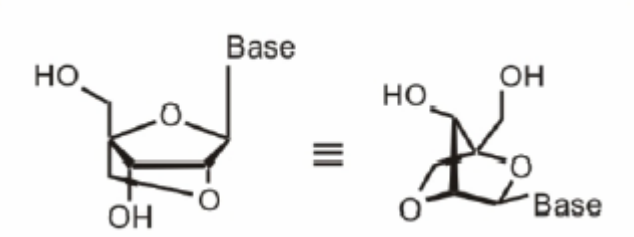

Fig. 2.1: LNA nucleoside.

Restriction and ligation

~ 100 ng of purified PCR product or plasmid DNA was digested in 20 μl with 1.0-2.0 units of the relevant restriction enzyme(s) and 2 μl of appropriate 10x buffer (Fermentas) at 37°C for 3 h. After restriction, the reaction was purified by GenElute PCR Clean-Up Kit according to the manual (Sigma, Germany). Ligation mixtures were set up as follows: ~100 ng of the insert; 10-20 ng of the plasmid DNA; 1.0 μl 10x ligation buffer; 1 μl T4 ligase (10U/μl) and 1 μl ATP (10 mM stock) in 15 μl. Ligation reactions were incubated overnight at room temperature (Sambrook and Russel, 2001).

Insertion of kanamycin resistance gene into plasmid (PCR mutagenesis)

We have adapted megaprimer PCR method (Lee *et al.,* 2004) to generate a linear synthetic construct with a kanamycin (Kan) cassette that was inserted by homologue recombination into the HC plasmid (Fig. 2.2). In this process, long fusion primers containing sequences homologous to both the HC plasmid and the Kan resistance cassette were used to separately amplify an upstream and a downstream region of the position where the Kan cassette is inserted (Tab. 2.2).

For preparation of megaprimers in the first step, PCR conditions are the same as described previously except with an increased volume of reaction (50 μl for all three steps). Megaprimers were purified by a commercial kit (GenElute PCR Clean-Up Kit; Sigma, Germany). In the second step, 10 μl of forward megaprimer was extended in PCR with 1 μl of reverse Kan primer (Kan2) and 1 μl of pUC4K plasmid harbouring Kan cassette as a template (Vieira and Messing, 1982). Analogically, for reverse megaprimer the Kan1 primer was used. For the last step, 10 μl of each PCR product from step two was mixed together with 1 μl of each 'border' primer (HC1940R and pCC5.21) and the whole construct is finished. The annealing temperature for steps two and three was set at 45 sec, 55°C and the elongation at 3 min, 72°C.

The prepared PCR product was transformed into the HC *Synechocystis* strain, where it was placed by homological recombination into the position of 1050 bp on the HC plasmid pCC5.2 and segregated (see below).

pCC5.2KAN3	**CTGTTTCCTGTGTGAAATTGTTA**TCCAAGTCAATCACCGCAAGGAT
pCC5.2KAN2	**CTTCTTCACGAGGCAGACCTCAG**CGGATTCTCAGATTTAGCCTTTG
HC1940R	CCGTCTTGTCTTCACCAAAG
pCC5.21	GTATTGTTGCGATCGCGAGC
Kan1	**TAGCCGGCTAACGTTATCGG**
Kan2	**TTGTGTCTCAAAATGTCTGATG**

Tab. 2.2: The sequence of oligonucleotides used for insertion of the Kan gene into pCC5.2 plasmid. Parts in bold are complementary to the kanamycin resistance cassette, the rest of the primers are complementary to pCC5.2.

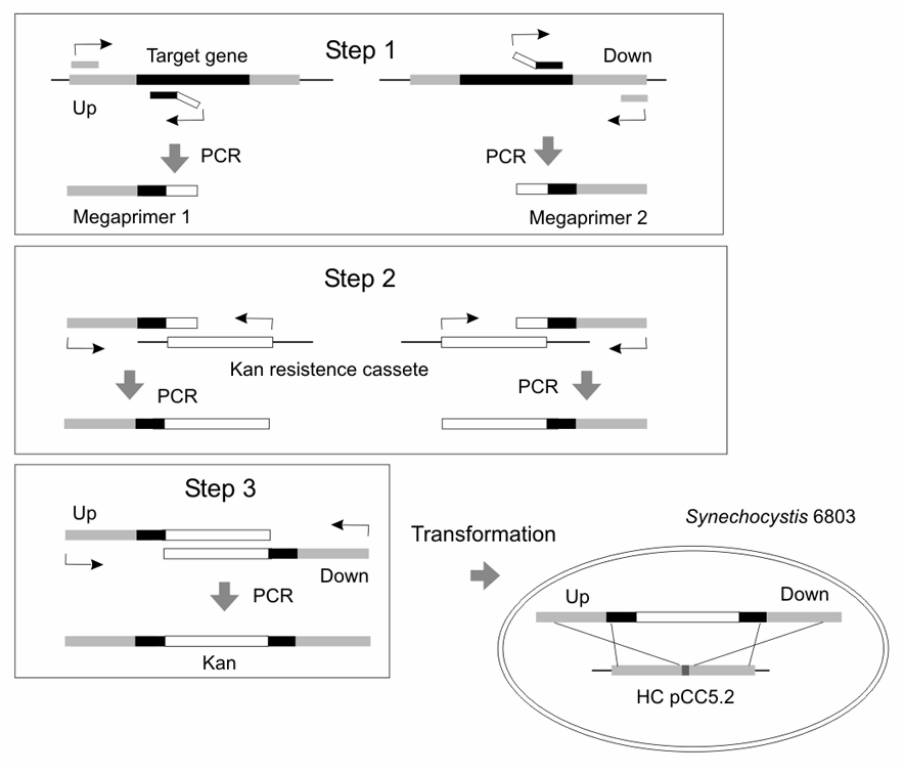

Fig. 2.2: Insertion of the kanamycin resistance gene into pCC5.2 by megaprimer PCR method.

Transformation of *Synechocystis* and segregation

A plasmid or linear PCR construct can be directly used for transformation of *Synechocystis* cells. Cells (at OD_{730} ~ 2) in liquid BG-11 are transformed by a one hour treatment with DNA at room temperature (Kufryk *et al.*, 2002). Cells are restreaked onto a BG-11 plate with glucose and incubated for 24 h, after which they are restreaked onto new a plate with a starting concentration

of antibiotic. *Synechocystis* has about 10 copies of chromosomal DNA and also plasmid DNA (see Results section). Hence, cells with a newly introduced gene are heterozygous (i.e., they have both WT and modified gene/plasmid) until replication and subsequent segregation of mutation produces a cell that is homozygous for the mutation (Currier *et al*., 1977). To promote full segregation, transformants with antibiotic marker are restreaked several times onto new plates with increasing antibiotic concentration (from 5 in the beginning to 100 μg ml^{-1} for kanamycin).

Isolation of DNA from *Synechocystis*

The full bacteriological loop of the cell culture was resuspended in 200 μl of TE buffer (100 mM Tris/HCl, pH 8.0; 10 mM EDTA, pH 8.0). An equal volume of phenol was added, the solution was agitated, left for 5 minutes at room temperature, agitated properly and spun down at 12 000 rpm for 5 min. The supernatant was pipetted into a new eppendorf tube with 150 μl of chloroform, the solution was agitated and spun at 12000 rpm for 2 min. Upper phase was carefully taken into a new eppendorf tube and mixed with 1/10 volume of Na-acetate and 400 μl of 96% ethanol. The ethanol solution was left in –70°C for 2 hours and precipitated DNA was centrifuged at 14000 rpm for 15 minutes. The ethanol was dried carefully in an exicator, and the pellet was washed in 200 μl of 70 % ethanol that was gently flushed. Pellet was spun down at 14000 rpm for 5 min and dried again in an exicator for 10 minutes. The DNA sediment was dissolved in 20 μl of TE buffer and isolated DNA was stored in –20°C (Hagemann *et al.,* 1997).

Plasmid DNA minipreparation

One *Synechocystis* colony was resuspended in 10 μl of water and mixed with 15 μl of Rusconi Mix (0.5 mg/ml lysozyme, 25 mM EDTA, 25 mM Tris/HCl pH 7.5, 0.1 mg/ml RNase, 13 % Glycerol, 2mg/10ml bromphenol blue). The mixture was incubated for 15 min at room temperature, then 2 μl of phenol was added and the mixture vortexed. The solution was extracted with 10 μl of chloroform, vortexed and spun down at 10 000 rpm for 2 min. The whole volume of supernatant was separated on a standard horizontal electrophoresis with 1 % TAE gel (40mM Tris-acetate, 1mM EDTA; Sambrook and Russel, 2001).

Plasmid purification by using a cesium chloride (CsCl) gradient

100 ml of cells in the log-phase of growth were collected at 6000 g for 10 min and then incubated with 0.5 ml of saturated NaI at 37°C for 30 min. The NaI-treated cells were washed twice with distilled water and then frozen at -70°C for 2 h (or overnight) before plasmid isolation. The frozen cells were thawed and resuspended in 0.5 ml of lysis buffer containing 50 mM glucose, 50 mM Tris-Cl (pH 8.0, 25°C), and 10 mM EDTA. Lysozyme was added to a final concentration of 15 mg/ml, the solution was gently stirred and the cells were lysed at 37°C for 1 h. 1 ml of 0.2 M NaOH / 1% SDS solution was then added, and the suspension was incubated on ice for 10 min, neutralized by 0.75 ml 3/5 M potassium acetate, and kept on ice for 5 min. The suspension was centrifuged and DNA recovered from the supernatant by 2-propanol precipitation. In the next step, DNA was dissolved in 0.125 ml TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0, at 25°C) buffer and treated with 20 μg/ml RNase A for 30 min at 37°C. Into the RNase A-treated solution 20 μl of EtBr (10 mg/ml) and CsCl (in concentration of 1 g per 1 ml of mixture) were added and dissolved at 37°C. The solution was loaded into 2 ml Beckman tubes and centrifuged at 50 000 rpm for 16 hours at 20°C (Xu and McFadden, 1997). The lower band consisting of plasmid DNA was carefully taken by syringe and analyzed on a horizontal agarose electrophoresis in 1.0 % TAE.

Purification of DNA from agarose gel

For purification of DNA from agarose gel, the excised piece of the gel containing the DNA was completely dissolved by incubation at 50°C for 10 min in NT solution from NucleoSpin Extract II kit (Macherey-Nagel, Germany). DNA was than purified from the dissolved gel by using GenElute PCR Clean-Up Kit (Sigma, USA) as though it were a PCR reaction.

4. RESULTS

Determination of the HC plasmid copy number

To estimate the copy number of the HC plasmid I used quantitative real-time PCR (RTq-PCR). Basically, a higher concentration of the template (a higher copy number of plasmid) in the reaction leads to the faster amplification of an observed gene and a lower Ct number. The Ct number means the cycle where the signal of the amplified DNA reaches the set threshold. The experiment was based on the amplification of the *ORF-B* gene located on the pCC5.2 plasmid with total DNA as a template. As a control, *slr0554* gene, a homologue of *ORF-B* located on the chromosome, was chosen, whose level would be the same in all tested strains and a course of PCR amplification should be fairly comparable. As a negative control the pCC5.2-less strain was used which lacked the plasmid entirely (see below, Fig 4.1).

For the detection of the PCR product accumulation in the reaction, I used an LNA probe designed by use of the Universal Probe Library (Roche) as described in experimental procedures. I have confirmed by a standard curve that a 10-fold higher concentration of template in the sample corresponds to 3.34 of cycle lower Ct.

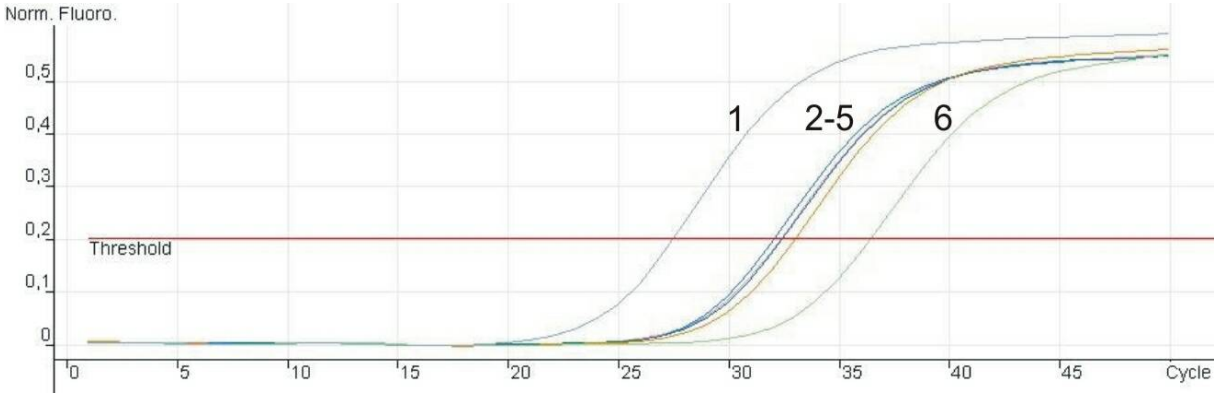

Fig. 4.1: Example of an RTq-PCR analysis for WT chromosomal (2), WT plasmid (3), HC chromosomal (5), HC plasmid (1), pCC5.2-less chromosomal (4) and pCC5.2-less plasmid (6) strain.

All the primers and the probe used for RTq-PCT are summarised in the methodology and results in Table 4.1. The WT strain of *Synechocystis* is known to have about 10 copies of chromosomal DNA per cell (Herdman *et al.*, 1979) and the copy number of pCC5.2 appears to be similar. In contrast, the HC strain has about 100-fold higher concentration of the *ORF-B* gene in the reaction than the WT. The concentration of the control chromosomal gene is

comparable suggesting that the HC strain has roughly about 1000 copies of pCC5.2 per cell. As expected, a very weak signal of *ORF-B* was detected in the pCC5.2 strain with no plasmid.

Name of sample	Dilution	Ct	pCC5.2 copy n.
WT slr0554	10^{-4}	32,48	~10
HC slr0554	10^{-4}	32,19	~10
PCC5.2-less slr0554	10^{-4}	32,55	~10
WT ORF-B	10^{-4}	33,11	~10
HC ORF-B	10^{-5}	27,61	~1000
pCC5.2-less ORF-B	10^{-3}	36,54	<0.1

Tab. 4.1.: Result of RTq – PCR and the estimation of HC and LC pCC5.2 copy numbers.
Shift of 3,34 cycle ~ 10 fold more template
Shift of 6,76 cycle ~ 100 fold more template

Insertion of kanamycin resistance cassette

Insertion of a gene for resistance to any antibiotic is usually the first step in, and is essential for preparation of a vector. Host cells that contain such a plasmid are then easily identified by screening for antibiotic resistance on a plate. *Synechocystis* is, structurally, a Gram-negative bacteria, thus it has most often been used with antibiotic resistance genes that are known to be effective for that group of bacteria (Koksharova and Wolk, 2002). These antibiotics include kanamycin, which I chose as a suitable resistance marker.

To prepare an HC plasmid with kanamycin (Kan) resistance (HC_{Km} plasmid), a method called PCR mutagenesis was employed (see experimental procedures). By this method, an approximately 2000 bp PCR fragment was prepared in three steps (Fig. 4.2) consisting of the Kan gene and ~400 bp long flanking regions to ensure homologue recombination into the position of 1050 bp in the HC plasmid (Fig. 4.3). This part of the plasmid does not contain any predicted ORF.

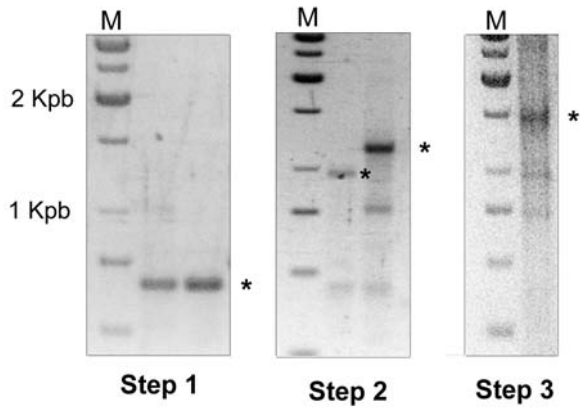

Fig. 4.2.: Products amplificated during PCR mutagenesis, each step is described in methodology.

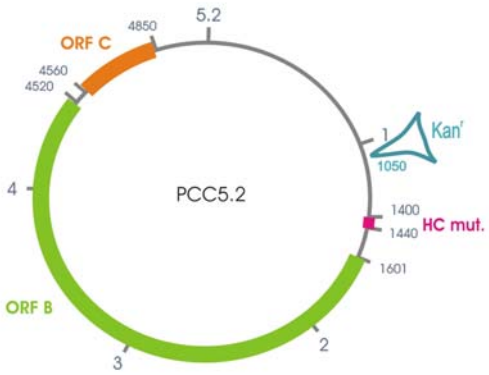

Fig. 4.3.: Map of HC pCC5.2 plasmid with inserted gene for kanamycin resistance (Kan).

This PCR construct was transformed into the HC strain and the Kan resistant colonies were selected (10 μg ml^{-1} of media). In 14 days, I obtained colonies which were then further segregated to plates with an increasing concentration of Kan. At a Kan concentration of about 80 μg ml^{-1}, the HC_{Km} plasmid becomes visible on the gel. A high copy version of pCC5.2 is easily visible on the agarose gel after DNA minipreparation (using Rusconi mix), which allows simple detection of the presence of plasmid from colonies (Fig. 4.4). Interestingly, during segregation of the HC_{Km} strain a few faster growing colonies formed, which were identified as cells that completely lacked the pCC5.2 plasmid (pCC5.2-less strain, Tichy, unpublished).

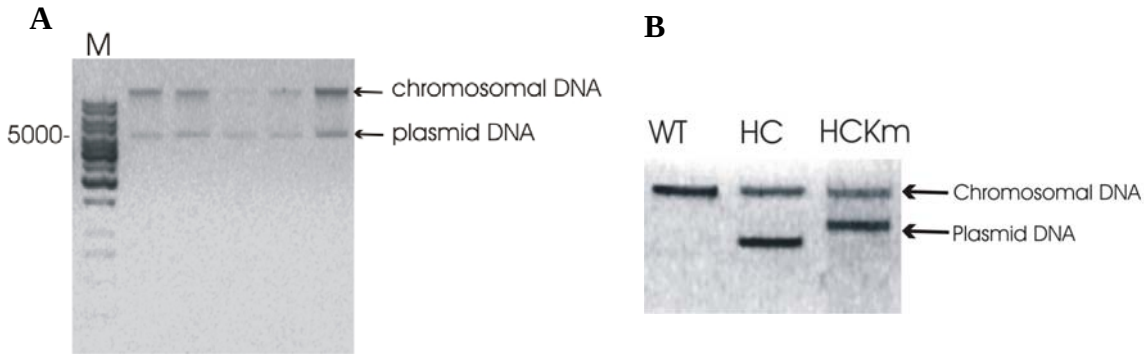

Fig. 4.4.: A) DNA minipreparation from the HC strain. **B)** Comparison of plasmid sizes in the HC and HC_{Km} strains.

Insertion of the gene coding for the CP47 protein into HC plasmid

Once I had determined that the HC plasmid has one thousand copies per cell and I had prepared the plasmid with the resistance gene, the next logical step was to try to use of the HC_{Km} plasmid for protein overexpression in *Synechocystis*.

I decided to overexpress the *Synechocystis* native CP47 protein; a large chlorophyll binding protein with six transmembrane helixes that is probably not possible to prepare in other expression systems (Fig. 4.5B). Together with D1, D2, CP43 and cytochrome *b*559 proteins, CP47 forms the core complex of Photosystem II in higher plants, algae, and cyanobacteria (Fig. 4.5A). All these proteins are assembled into Photosystem II in an ordered manner, and the inactivation of a single protein leads to the destabilization of Photosystem II and its disappearance from the thylakoids (Komenda *et al.*, 2004). However, nothing is known about the effects of a higher accumulation of any of these proteins on Photosystem II function.

Furthermore, CP47 appears to be important for the accumulation of chlorophyll, but the relationship between chlorophyll biosynthesis and the synthesis of CP47 remains unclear (Sobotka *et al.*, 2005). In addition, the mechanism for degradation of chlorophyll proteins; proteases involved in the process, and the fate of the chlorophyll released during protein degradation is unclear. Thus overexpression of CP47 could give important clues for all these topics. It should be also noted that specific CP47 antibodies are available in the Trebon laboratory.

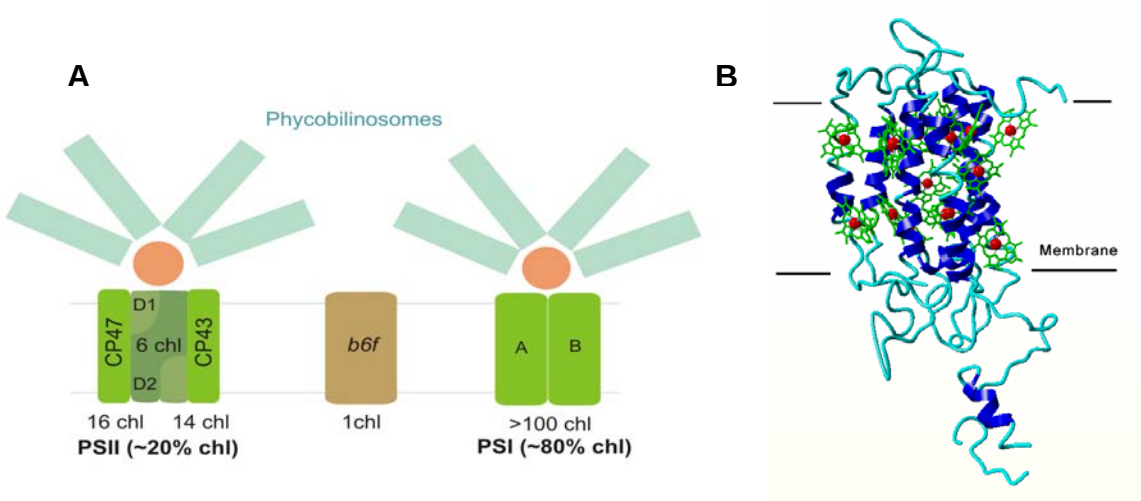

Fig. 4.5.: A) Cyanobacterial thylakoid membrane with depicted known chlorophyll binding proteins. **B)** Crystal structure of protein CP47 (Kamiya and Shen, 2003).

To insert the *psbB* gene coding for the CP47 protein into the HC_{Km} plasmid, the gene was amplified with its own promoter by primers incorporated in the *NcoI* restriction site (Fig 4.6 ; Tab. 4.2) The PCR product was purified, digested and ligated into a unique *NcoI* site in the HC_{Km} plasmid (Fig. 4.8A). The HC_{Km} plasmid was isolated from the HC_{Km} strain by the alkalic lysis method used frequently for plasmid isolation from *E. coli*. The ligation mix was transformed into the $CP47^-$ mutant where the *psbB* gene plus a several hundred base pair long downstream sequence is replaced by the spectinomycin cassette (Eaton-Rye and Vermaas, 1991). This strain is not able to grow autotrophicaly as it lacks a functional Photosystem II and so has to be maintained on medium containing glucose. It is thus possible to select autotrophic colonies with a functional CP47 protein.

Colonies containing the HC_{Km}-*psbB* plasmid were selected on kanamycin resistance combined with autotrophic growth.

psbB_NcoI1	GGCCGCG**CCATGG**GCGATC
psbB_NcoI2	GCTGTG**CCATGG**AGAAGCACCTAG

Tab. 4.2.: Sequences of primers used for the amplification of the *psbB* gene. *NcoI* restriction sites are in bold.

Fig. 4.6.: PCR amplification of the *psbB* gene coding for the CP47 protein.

A **B**

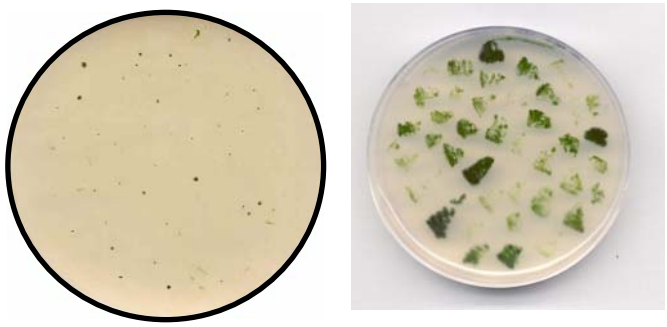

Fig. 4.7.: A) Autotrophic colonies of CP47$^-$/HC$_{km}$-*psbB* growing on a selective plate with Kan. **B)** Segregation of the same colonies on the plate with 80 μg of Kan.

After 14 days on the plate with 10 μg ml^{-1} of Kan, I obtained a number of colonies with different growth rates (Fig. 4.7). Total DNA from colonies of different sizes was isolated and tested by PCR for the presence of the plasmid pCC5.2, CP47 gene and the kanamycine gene (Fig. 4.8B). The WT DNA was used as a control. Interestingly, all these PCRs suggested that the colonies possessed a low copy version of the pCC5.2 plasmid with the incorporated *psbB* gene in the right *NcoI* site and the empty HC$_{Km}$ plasmids with no *psbB* gene.

This finding suggested that the *psbB* is expressed in the plasmid but also strongly indicated a problem with recombination between HC and native versions of plasmids in transformed cells. Simply, it looked like the cells exchanged the HC mutation between HC plasmid and the native plasmid to acquire *psbB* in low copies per cell.

To overcome this problem, the pCC5.2-less/CP47$^-$ mutant was prepared. I transformed the pCC5.2-less strain described above by chromosomal DNA of the CP47$^-$ mutant and segregated the CP47$^-$ deletion. This strain was transformed by HC$_{Km}$ + CP47 ligation mix and two types of selection were tested: **A)** I selected positive colonies for Kan to get all possible colonies bearing the HC$_{km}$ plasmid and those colonies were subsequently restreaked on a new plate with Kan combined with autotrophic growth, there I obtained just a small number of colonies. The colonies were tested by PCR for the presence of the pCC5.2 plasmid, the pCC5.2 plasmid carrying Kan and pCC5.2 carrying *psbB* (Fig. 4.9). I detected a positive

signal for all the combinations of primers. Such results show that in transformant colonies there are cells bearing two types of plasmid pCC5.2 – with and without CP47 (Fig. 4.10). By DNA minipreparation only HC_{Km} was visible, this finding demonstrates that the plasmid bearing *psbB* gene was maintained under control.

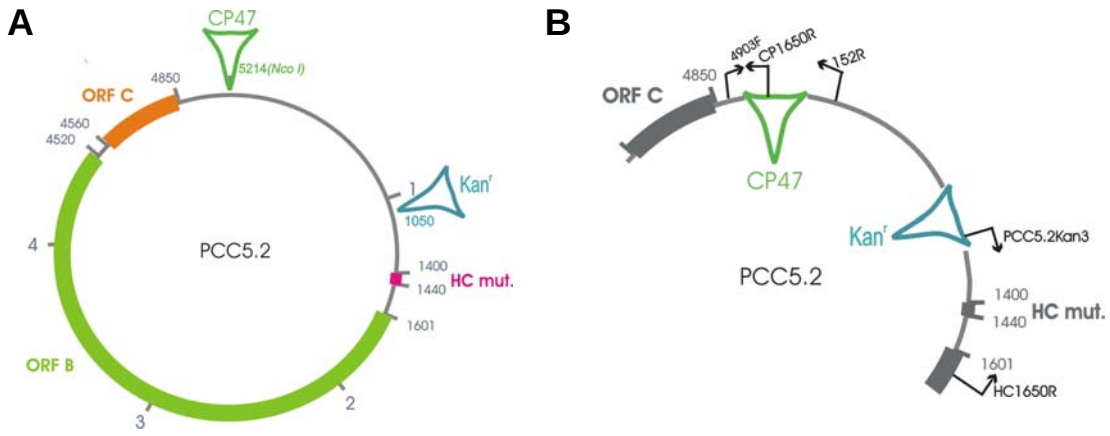

Fig. 4.8.: A) Map of the pCC5.2 plasmid with the inserted gene for CP47 protein and for kanamycin resistance. **B)** Primers used for detection of the Kan cassette and *psbB* gene inserted into the HC plasmid

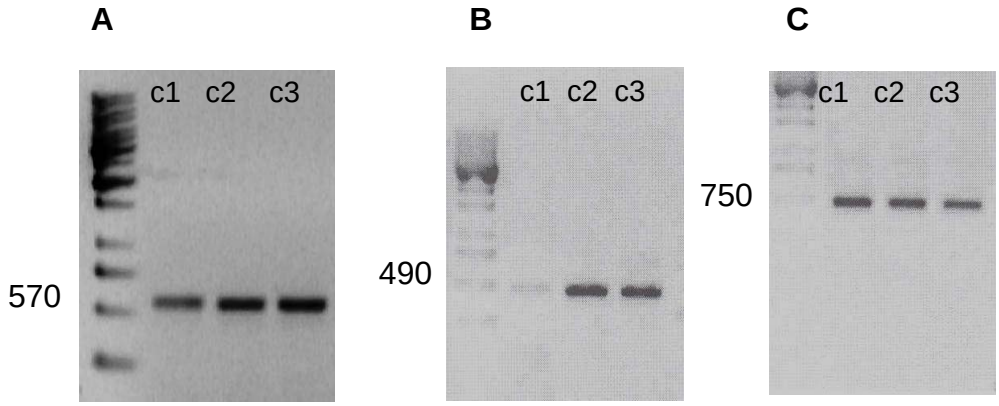

Fig. 4.9: PCR characterization of HC_{Km} + CP47 colonies; three colonies of different sizes – c1, c2, c3. **A)** PCR detection of the *psbB* gene inserted into pCC5.2 (primers CP1650R+4903F). **B)** PCR detection of the empty pCC5.2 (primers 4903F+152R). **C)** PCR detection of HC_{Km} plasmid (primers PCC5.2Kan3+HC1650R).

B) Surprisingly when colonies were selected for autotrophic growth only several tens of colonies appeared. Furthermore, no signal from the pCC5.2 plasmid bearing the *psbB* gene was detected, not even HC_{Km} was detectable. However, the *psbB* gene was apparently

presented in the cells as they grew autotrophically. The only explanation is that the *psbB* is incorporated into the chromosomal DNA. This might be caused by ligation of the insert containing the *psbB* with fragments of chromosomal DNA instead of ligation with the plasmid and subsequent homological recombination into the random places in the genome. All colonies tested were resistant for spectinomycin showing that the chromosomal copy of *psbB* is missing. In a separate experiment, I confirmed that the autotrophic growth of CP47$^-$ strain cannot be rescued by transformation with the PCR product containing the *psbB* gene.

Modified approach for the preparation of pCC5.2 vector

Isolation of pCC5.2 plasmid by a cesium-chloride (CsCl) gradient

My results indicated that the *Synechocystis* cell is able to incorporate the *psbB* gene into the genome, possibly by recombination using a *psbB* gene randomly ligated with fragments of chromosomal DNA. In the first part of the project I used an isolated plasmid based on the standard plasmid isolation from *E. coli* (Sambrook and Russel, 2001). Such isolation is highly contaminated by chromosomal DNA (Fig. 4.10B). To solve this problem I prepared pure plasmid DNA by a combined approach. In the first step, the HC plasmid and chromosomal DNA were separated on a CsCl gradient in order to discard most of the chromosomal DNA (Fig. 4.10A). In the next step, the HC plasmid from the gradient was further separated on the TAE agarose gel. The plasmid band was cut from the gel and purified by use of commercial kits (see experimental procedures).

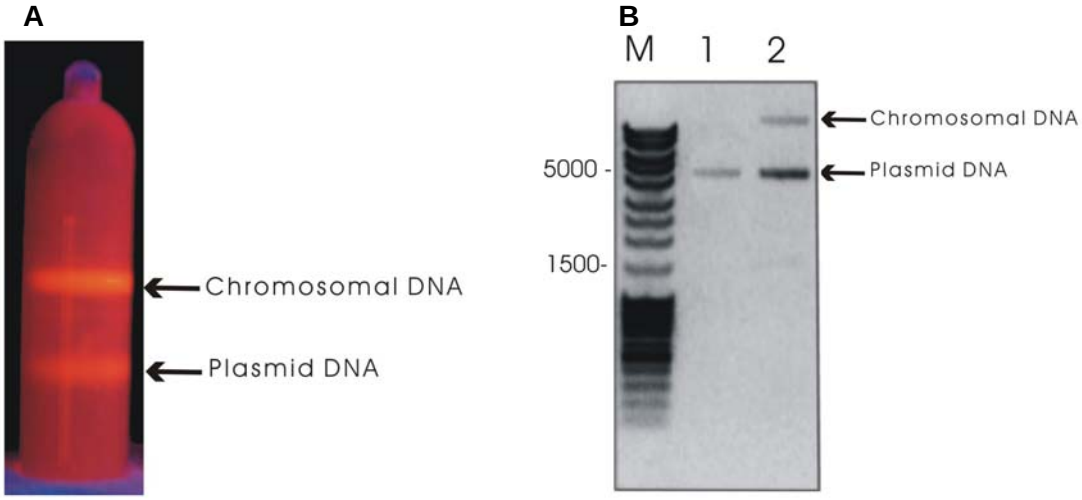

Fig. 4.10.: A) DNA isolated from *Synechocystis* HC strain and separated using CsCl gradient. **B)** Comparison of the plasmid DNA isolated by CsCl gradient with a plasmid isolation based on the standard alkalic lysis method.

Fusion of the gene of interest with erythromycin cassette

By preparation of a chromosomal DNA-free plasmid I solved a potential problem with homologue recombination of randomly ligated fragments of chromosomal DNA with the insert. Another modification I decided to do was fusion of the gene for overexpression with the resistance gene followed by ligation of the whole construct into the HC plasmid. By this approach I hoped to avoid the false selection on antibiotic due to the empty plasmid with the resistance gene.

In addition, aside from the *psbB* gene, I also decided to test overexpression of the *gfp* gene encoding for the light-emitting green fluorescent protein of the jellyfish *Aequorea victoria*. The great advantage of this system is that it works without an additional substrate and, unlike other reporter proteins, does not suffer from background problems (Kunert *et al.*, 2000). Moreover, GFP is a soluble protein and therefore I would not expect the same potential problems for cell viability as for CP47 which is a membrane protein whose high accumulation can be toxic. The construct containing the *gfp* gene under the *Synechocystis* native *psbAII* promoter (described in the introduction) has already been prepared by Dr. Martin Tichy and was available in our laboratory.

Finally, for this modified approach I used the erythromycin (Ery) resistance gene instead of Kan for screening of colonies. The reason for this was the significantly affected growth rate observed in the HC_{Km} strain bearing the HC plasmid with the Kan cassette. Whilst the doubling time of the HC strain was about 10 h, which is comparable with the WT, the HC_{Km} strain on higher concentration of Kan has doubling time about 33 h. It should be noted that such concentration of Kan itself has no effect on growth of different Kan resistant strains of *Synechocystis* (Tichy, unpublished data). It is also believed that erythromycin (Ery) offers very fine selection in cyanobacteria (Shen *et al.*, 1993).

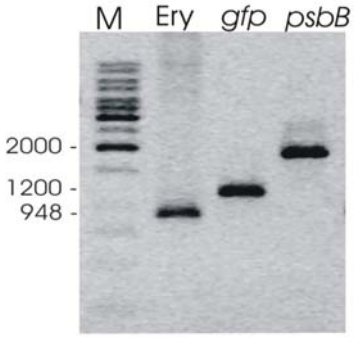

Fig. 4.11.: PCR amplification of Ery cassette, *gfp* and *psbB* genes, respectively.

Fig. 4.12.: *BanII*, *NcoI* and *BspHI* restriction sites, respectively.

The gene and the resistance cassette were ligated together by *BanII* sites and then inserted into the vector through the *NcoI* site (Fig. 4.15). I amplified both *psbB* and *gfp* genes together with their promoters by one primer containing an *NcoI* site (or *bspHI* site for *gfp*) and a second primer with an incorporated *BanII* site (Tab. 4.3.; Fig. 4.11, 4.12). The *BanII* containing primer was forward for the Ery and reverse for the *psbB*/*gfp* genes, which should ensures that only the gene and the cassette can be ligated together and no long unspecific chains of fragment would form (Fig. 4.15).

A

primer	Sequence	direction
PsbABanIIF	CGGAACCCA**CCCGAG**ACAATGTGACCA	forward
GfpBspH1R	CCAAGGAA**TCATGA**GGATCTATTTGTAGAGTTCATCCATG	reverse

B

Primer	Sequence	direction
psbBNcoIF	GGCCGCG**CCATGG**GCGATC	forward
psbBBanIIR	GCTGTGCAAGAAGC**GAGCCC**TAGGC	reverse

C

Primer	Sequence	direction
Ery-BanIIF	GTTCATATTTATCA**GAGCCC**GTGCTAT	forward
EryNcoIR	AACAAGTTAAGGGATG**CCATGG**TTTATG	reverse
EryBspHIR	AA**TCATGA**GTTAAGGGATGCAGTTTATG	reverse

Tab. 4.3.: Sequence of primers used for the amplification of *gfp* **(A)**, *psbB* **(B)** and Ery **(C)** genes; bold case letters indicate restriction sites.

The amplified PCR fragments were mixed: *psbB* plus Ery (or *gfp* plus Ery) in equal concentration and digested by the *BanII* restriction enzyme. The digested fragments were purified by a commercial kit (see experimental procedures) and ligated together. As the result of *psbB* - Ery ligation was weak, I amplified the fuzed construct by PCR using ‘border‘ *NcoI* primers and ligation mixture as a template (Fig. 4.13). In the second step, the ligated construct was digested by the *NcoI* (or *bspHI* for the *gfp* gene) enzyme on both sites and the restriction enzyme was subsequently deactivated by heat. When chilled, the HC plasmid digested by the *NcoI* and ligase enzyme was added to prepare the final overexpression vector (Fig. 4.15).

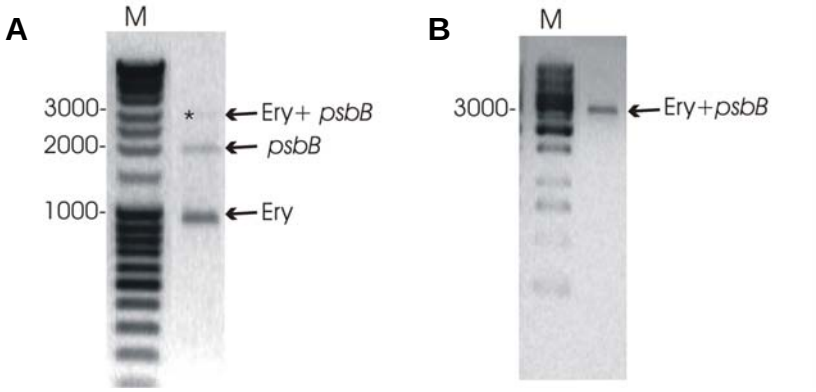

Fig. 4.13: A) Analysis of the Ery and *psbB* gene after ligation. **B)** PCR amplification of Ery-*psbB* construct.

For the CP47 overexpression, the plasmid was transformed into the pCC5.2-less/CP47$^-$ strain and colonies were screened for resistance to Ery or for the combination of Ery resistance and autotrophic growth. ~20 days after transformation several colonies appeared. On the agar plate supplemented with Ery (10 μg ml^{-1}) and 5 mM glucose I observed colonies with different growing rates. In contrast, on the plate with Ery (10 μg ml^{-1}) and no glucose, a high number of slowly growing colonies was obtained. Nevertheless, these colonies were also able to grow autotrophicaly in the presence of Ery confirming the presence of the fused Ery-*psbB* construct. Surprisingly, after DNA minipreparation using the Rusconi mix, no plasmids were visible (Fig. 4.14). In addition, 20 colonies of all types were tested by PCR for the presence of the pCC5.2 plasmid and the plasmid with Ery-*psbB* construct. By this method, I did not detect any pCC5.2 plasmid in colonies.

The ligation mix with the HC plasmid and the *gfp*-Ery construct was transformed into the pCC5.2-less strain but, unfortunately, screening of colonies is still in progress at the time of thesis preparation.

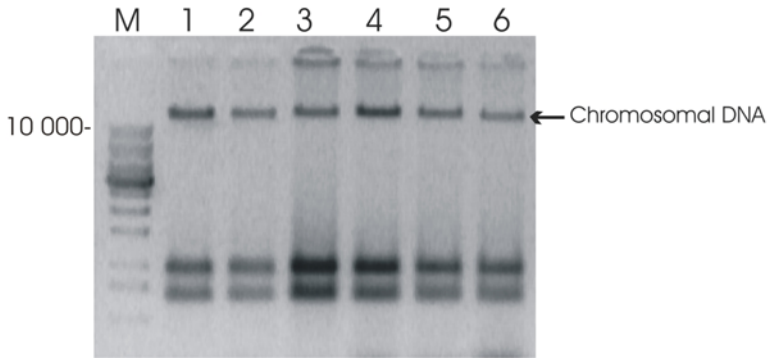

Fig. 4.14: DNA isolation from colonies CP47-/HC-Ery-*psbB* (1-6) resistant to Ery and able to grow autotrophically. Predicted mobility of the HC plasmid with Ery-*pbsB* is about 8000 pb.

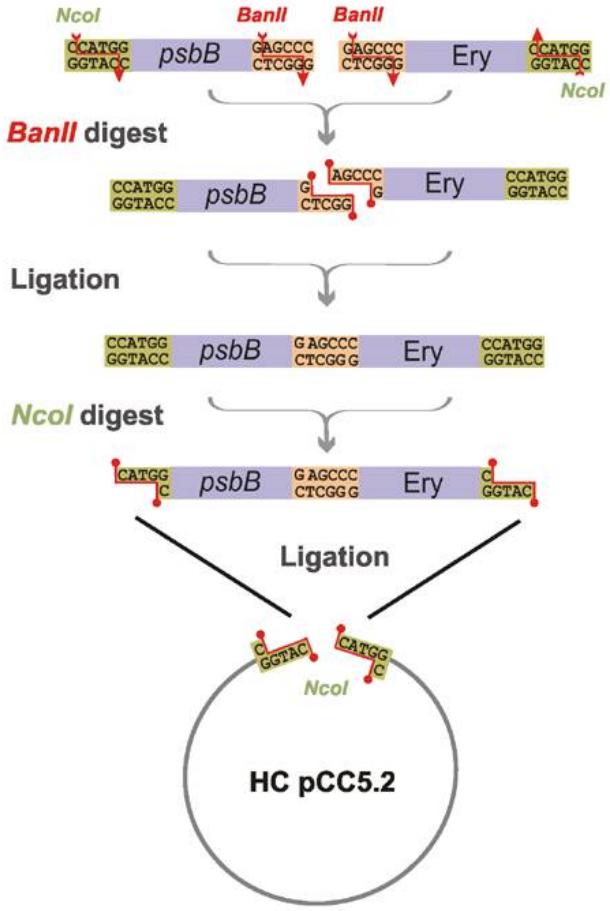

Fig. 4.15: Strategy of producing HC pCC5.2 bearing Ery and *psbB* fused gene in *NcoI*.

5. DISCUSSION

The plasmid pCC5.2 belongs to an un-promiscuous group of plasmids, however its function is still unknown, even though both this plasmid and its host organism have been completely sequenced. Nevertheless, the sequences of the potential proteins on the plasmids allow the finding of homologues on other plasmids and chromosomal DNA from the host or different organisms. Some plasmids bear homologues of enzymes participating in important biochemical pathways, for example the cryptic gene *ndhK2* on the *Synechocystis* pSYSG plasmid, which probably codes for a subunit of NADH dehydrogenase (Kaneko *et al.*, 2003). Such genes may work as a backup for those genes encoded in the chromosome and may be expressed in some specific growth/stress conditions. However, in the case of the pCC5.2 plasmid, there is no such ORF, whose role in the cell metabolism would be obvious. So, the function of this plasmid is unknown, however it does not seem to be important for *Synechocystis* growth under the conditions I have employed. The doubling time of the HC strain was comparable with the WT as well as with the pCC5.2-less strain suggesting that there is no significant effect of the pCC5.2 plasmid on *Synechocystis* viability. This corresponds with my finding that different strains cultivated for a longer time (months) on agar plates spontaneously lost the pCC5.2 plasmid.

From the sequencing of the pCC5.2 six potential ORFs have been identified but only *ORF B* and *ORF C* have some homologues. *ORF C* consists of 93 amino acids (aa) and it is similar to a group of hypothetical proteins whose function is unknown. *ORF B* consists of 971 aa and forms an important part of the whole pCC5.2 plasmid. This protein has homologues in five other cyanobacteria, where the genes are suggested to have an important role in plasmid replication. This also indicates that *ORF B* encodes a protein for pCC5.2 plasmid replication. Furthermore, sequence analysis of the chromosome of *Synechocystis* 6803 shows that there is a protein that has 31% sequence identity to the product of *ORF B* (Xu and McFadden, 1997). Maybe this homologue protein plays a role in chromosomal replication in *Synechocystis* or even somehow controls the copy number of chromosomal DNA.

By RTq-PCR I have determined that the pCC5.2 copy-number is approximately 10 copies per cell but increased by ~100 times for a plasmid that carries the HC mutation. It is known that plasmid replication initiates in a predetermined *cis*-site called *ori* and can proceed either by a rolling circle or a theta replication mechanism. Some of the plasmid-encoded elements required for their replication, such as antisense RNA molecules and DNA repeated sequences

located close to *ori*, determine plasmid attributes like copy number and incompatibility (Actis *et al.*, 1998). Plasmid pCC5.2 has been shown to replicate by rolling-circle mechanism but there is no apparent *ori* region (Xu and McFadden, 1997). However, it is very interesting that the mutation responsible for the high copy of pCC5.2 was localised upstream of *ORF B* (1000-1040 bp) in its putative promoter. It is attractive to speculate that *ORF B* controls the replication of the plasmid and also its own transcription and the mutation in the promoter causes uncontrolled accumulation of this protein and consequently the plasmid. I tested antibodies against *ORF B* prepared on a commercial basis but, unfortunately, no specific cross-reactions were observed.

An important part of my project was the preparation of a pCC5.2 bearing kanamycine cassette. By using PCR mutagenesis I prepared a construct and obtained a fully segregated HC_{Km} strain. However, to maintain the high accumulation of the HC_{Km} plasmid, HC_{Km} cells must be maintained on agar plates with a moderate concentration of kanamycine (50 $\mu g\ ml^{-1}$). This demonstrates that in this case cells can somehow reduce their plasmid copy number. At such a concentration of antibiotics, the HC_{Km} strain had also significantly decreased growth rate in comparison with the WT. The HC strain has about 1000 copies per cell, so logically, when HC has kanamycine inserted, there will also be 1000 copies of the Kan gene. It is possible that high accumulation of the protein promoting Kan resistance can cause problems for cell viability.

In spite of the problem with reduced HC_{Km} strain growth rate, the plasmid with the Kan cassette has been chosen for subsequent work, because it was nicely selectable. In the first trial, I simply inserted the *psbB* gene coding for the CP47 protein into the HC_{Km} plasmid and transformed the WT cells. I got hundreds of colonies which were resistant to kanamycine and able to grow autotrophicaly but the *psbB* gene was always detected in a low copy version of the plasmid, whereas the HC_{Km} plasmid was empty. As CP47 is a membrane protein, which may be somehow involved in chlorophyll synthesis (Sobotka *et al.*, 2005), it is probable that its high accumulation is toxic for the cell. So, a way for *Synechocystis* to survive an uncontrolled accumulation of CP47 could be a recombination between the plasmids to get the *psbB* gene on the low copy version of the plasmid (or the native *ORF B* promoter on the plasmid with the *psbB* gene).

To avoid possible recombination between the HC and the LC plasmid by homologue recombination, the pCC5.2-less strain was used for the next transformation. Screening for the combination of Kan resistance and autotrophic growth produced colonies with two types of HC_{Km} plasmid – with and without the *psbB* gene, which indicated double transformation of

these cells by both the self-ligated HC_{Km} and the *psbB* HC_{Km} plasmid. Only a few colonies obtained in this case would correspond to the expectation that the double trasformation is a rare event. Cells selected only for autotrophic growth had no HC plasmid at all. However, native *psbB* remained interrupted suggesting that the *psbB* gene has to be in another place in the genome, where its transcription can be controlled. This finding indicated to me that the PCR product with *psbB* could make random fragments with chromosomal DNA in a ligation reaction and some fragments may be able to promote homological recombination. Chromosomal DNA remains in the solution when the plasmid was isolated by standard plasmid isolation based on alkalic lysis method. I have tried to solve this problem by using as precise a plasmid isolation method as possible to avoid unspecific recombination in *Synechocystis*. I combined purification through the CsCl density gradient with separation of the isolated plasmid on agarose gel, which resulted in the preparation of chromosomal DNA free plasmid isolation.

To improve the selection of positive colonies and overcome the false selection of self-ligated empty HC_{Km}, I made a plasmid with the Ery resistance gene fused with the *psbB* again. Tens of transformants were able to grow autotrophicaly in the presence of erythromycin but, surprisingly, they did not contain a pCC5.2 plasmid. Cells must apparently have both these genes, as all were able to grow autotophically when selected only on Ery resistance. The question is, how it is possible that the PCR construct was inserted into some place in the genome when there was very probably no chromosomal DNA in the ligation mixture? I do not know the answer yet, however, these results might indicate some new mechanism in *Synechocystis* that allows the incorporation of a foreign fragment of DNA into the genome perhaps via the pCC5.2 plasmid. For future work it would probably be very useful to employ the *recA*$^-$ strain. The *recA* gene is essential for homologue recombination and a fully segregated *Synehocystis* mutant is described on web (http://bacteria.kazusa.or.jp/cyano/Synechocystis/mutants/).

Preparation of an HC pCC5.2 consisting of an Ery cassette fused with the *gfp* gene instead of *psbB* is now under progress. The *Gfp* protein may be better for overexpression than CP47 as it is a soluble protein and thus its high accumulation should be less toxic for the cell. Another possibility for overcoming toxicity caused by thousands of copies of inserted gene could be with the help of inducible promoters. In this case I could get better results than with the native CP47 promoter or even with the *psbA2* light regulated promoter now being tested for GFP expression (Lagarde and Vermaas, 1999). Examples of well-regulated promoters are the copper dependent *petE* promoter (Buikema and Haselkorn, 2001) or the recently described

nirA nitrate/ammonium regulated promoter (Desplancq *et al.*, 2005; Qi *et al.*, 2005). Both have already been used for controlled expression in cyanobacteria. Using some of these promoters for the expression of a pCC5.2-based vector could be controlled in a similar way to how IPTG is used for the modern expression systems in *E. coli*. Moreover, overexpression is not its only use; such a vector is suitable for other applications such as down-regulation of a gene in a particular genetic background or RNA silencing.

6. CONCLUSIONS

Synechocystis 6803 is the most studied cyanobacteria and it is completely sequenced including all native plasmids. During this project I have characterized a mutated version of a small *Synechocystis* plasmid, which has a noteworthily high copy number. By using RTq-PCR I have estimated the number of copies as approximately 1000 per cell. As an important step for the preparation of a usable expressing vector I inserted the gene for kanamycine resistance into this plasmid.

The main part of my work was based on transformation experiments which showed some features of *Synechocystis* which complicate the use of this plasmid for protein expression. To eliminate overexpression of a given gene, cells are able to change plasmid copy number by recombination with the native version of plasmid. Cells also seem to use fragments of chromosomal DNA contaminating the ligation reaction to insert a gene into genome. If this is true, it demonstrates that transformation and recombination efficiency of *Synechocystis* is amazing.

To prevent these problems I have prepared a new approach for plasmid isolation from *Synechocystis* and developed a method for the insertion of a gene fused directly with the gene for resistance into the plasmid. Surprisingly, even when using this approach some *Synechocystis* cells were able to incorporate the fusion construct into their genome and thus again overcome screening conditions. Taking all results into account, it is probable that overexpression of the CP47 protein is simply not compatible with cell viability and use of the *gfp* gene would be more successful. Nevertheless, experimenting with the CP47 protein I have confirmed that *Synechocystis* is the organism that always finds a way to survive.

REFERENCES

Actis, L.A., Tolmasky,M.E., Crosa, J.H. (1998) Bacterial plasmids: replication of extrachromosomal genetic elements encoding resistance to antimicrobial compounds. Front Biosci 3 **4:** D43-62.

Bolivar, F., Rodriguez, R.L., Green, P.J., Betlach, M.C., Boyer, H.W. (1997) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2:** 95-113.

Brosius, J. (1988) Expression vectors employing lambda-, trp-, lac-, and lpp-derived promoters. In: Rodruguez, R.L., Denhardt, D.T, (Eds) Vectors. A survey of Molecular Cloning Vectors and Their Uses, pp 205-255. Butterworth, Boston.

Bryant, D. B. (1994) The Molecular Biology of Cyanobacteria. Kluwer Academic Publishers, Dordrecht, Boston.

Buikema, W.J., Haselkorn, R. (2001) Expression of the Ananbaena hetR gene from a copper-regulated promoter leads to heterocysts differentiation under repressiong conditions. Proc Natl Acad Sci USA **5:** 2729-34.

Conley, P.B., Lemaux, P.G., Grossman, A. (1988) Molecular characterization and evolution of sequences encoding light-harvesting components in the chromatically adapting cyanobacterium Fremyella diplosiphon. J Mol Biol **199:** 447-465.

Currier, T.C., Haury, J.F., Wolk, C.P. (1977) Isolation and preliminary characterization of auxotrophs of a filamentous cyanobacterium. J Bacteriol **129:** 1556-1562.

Desplancq, D., Bernard, C., Sibler, A.P., Kieffer, B., Miguet, L., Potier, A., Van Dorsselaer, A., Weiss, E.(2005) Combining inducible portein overexpression with NMR-grade triple isotope labeling in the cyanobacterium Anabaena sp. PCC 7120. BioTEchniques **39**: 405-411.

Dühring, U., Axmann, I. M., Hess, W. R., Wilde, A. (2006) An internal antisense RNA regulates expression of the photosynthesis gene isiA. Proc. Natl. Acad. Sci. USA **103:** 7054–7058.

Durán, R.V., Hervás, M., De la Rosa, M.A., Navarro, J.A. (2003). The efficient functioning ofphotosynthesis and respiration in Synechocystis sp. PCC 6803 strictly requires the presence of either cytochorme c6 or plastocyanin. J Biol Chem Papers in Press.

Eaton-Rye, J.J., Vermaas, W.F.J. (1991) Oligonucleotide-directed mutagenesis of psbB, the gene encoding CP47, employing a deletion mutant strain of the cyanobacterium Synechocystis sp. PCC 6803. Plant Mol Biol **17:** 1165–1177.

Elhai, J. (1993) Strong and regulated promoters in cyanobacterium Anabaena PCC 7120. FEMS Microbiol Lett **2:** 179-84.

Ferino, F., Chauvat, F. (1989) A pormotor-probe vecotr-host systém for the cyanobacterium Synechocystis PCC 6803. Gene **84:** 257-266.

Friedberg, D. (1988) Use of reporter genes in cyanobacteria. Meth Enzymol **167:** 736-747.

Glick, B.R., Pasternak, J.J.(2003) Molecular Biotechnology, pp 47-90, ASM Press, Washington, D.C.

Grigorieva, G., Shestakov, S. (1982) Transformation in the cyanobacterium Synechocystis sp. 6803. FEMS Microbiol Lett **13:** 367-370.

Hagemann, M., Schoor, A., Jeanjean, R., Zuther, E., Joset, F. (1997) The stpA gene from Synechocystis sp. strain PCC 6803 encodes the glucosylglycerol-phosphate phosphate involved in cyanobacterial osmotic response to salt shock. J Bacteriol **179:** 1727-1733.

Herdmann, M., Javier, M., Waterbury, J.B., Rippka, R., Stanier, R.Y., Mandel, M. (1979) Deoxyribonucleic acid base composition of cyanobacteria. J GenMicrobiol **111:** 63-71.

Houmard, J., Tandeau de Marsac, N. (1988) Cyanobacterial genetic tools: Current status. Meth Enzymol **167:** 808-847.

Chauvat, F., De Vries, L., Van der Ende, A., Van Arkel, G.A. (1986) A host-vector systém for gene cloning in the cyanobacterium Synechocystis PCC 6803. Mol Gen Genet **204:** 185-191.

Ikeuchi, M., Tabata, S. (2001) Synechocystis sp. PCC 6803 - a useful tool in the study of the genetics of cyanobacteria. Photosynthesis Res **3:** 109-136.

Kamiya, N.,Shen, J.R. (2003) Crystal structure of oxygen-evolving photosystem II from Thermosynechococcus vulcanus at 3.7-A resolution. Proc Natl Acad Sci U S A **100(1):**98-103.

Kaneko, T., Nakamura, Y., Sasamoto, S., Watanabe, A., Kohara, M., Matsumoto, M., Shimpo, S., Yamada, M., Tabata, S. (2003) Structural analysis of four large plasmids Harboring in a unicellular cyanobacterium, Synechocystis sp. PCC 6803. DNA Res **10:** 221-228.

Koksharova, O.A., Wolk, C.P. (2002) Genetic tools for cyanobacteria. Appl Microbiol Biotechnol **58:** 123-137.

Komenda, J., Reisinger, V., Müller, B.C., Dobáková, M., Granvogl, B., Eichacker, L.A. (2004) Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in SynechocystisPCC 6803. J Biol Chem **253:** 48620-48629.

Kufryk, G.I., Sachet, M., Schmetterer, G., Vermaas, W.F.J. (2002) Transformation of the cyanobacterium Synechocystis sp PCC 6803 as a tool for genetic mapping: optimization of efficiency. FEMS Microbiol Lett **206 (2):** 215-219.

Kunert, A., Hagemann, M., Erdmann, N, (2000) Constructions of promoter probe vectors for Synechocystis sp. PCC 6803 using the light-emiting reporter systém Gfp and LuxAB. J Microbiol Methods **3:** 185-94.

Lagarde,D., Vermaas, W. (1999) The zeaxantin biosynthesis enzyme beta-carotene hydroxylase is ivolved in myxoxanthophyll synthesis in Synechocystis sp. PCC 6803. FEBS Lett **3:** 247-51.

Lee, J., Lee, H.J., Shin, M.K., and Ryu, W.S. (2004) Versatile PCR-mediated insertion or deletion mutagenesis. Biotechniques **36:** 398-399.

Miyake, M., Nagai, H., Shirai, M., Kurane, R., Asada, Y. (1999) A high-copy-number plasmid capable of replication in thermophilic cyanobacteria. Appl Biochem Biotechnol **77-79:** 267-75.

Poncelet, M., Cassier-Chauvat, C., Leschelle, X., Bottin, H., Chauvat, F. (1998) Targeted deletion and mutational analysis of the esential (2Fe-2S) plant-like ferredoxin in Synechocystis PCC 6803 by plasmid shuffling. Mol Microbiology **28(4):** 813-821.

Porra, R.J., Thomson, W.A., Kriedemann, P.E. (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochemica et Biophysica Acta **975 (3):** 384-394.

Qi, Q.G., Hao, M., Ng, W.O., Slater, S.C., Baszis, S.R., Weiss, J.D., Valentin, H.E. (2005) Application of the Synechococcus nirA promoter to estabilish an inducible expression systém for engineering the Synechocystis tocopherol pathway. Appl Env Microbiol **71:** 5678-5684.

Rippka, R., Deruelles, J., waterbury, J.B., Herdman, M., Stanier, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J Gen Microbiol **111:** 1-61.

Sambrook, J., Russell, D. (2001) Molecular Cloning: a Laboratory Manual, 3rd ed., Cold Spring Harbor, N.Y.

Sarma, T.A., Kaur. B. (1997) Characterization of host-range mutants of cyanophage N-1. Acta Virol **41:** 245-250.

Shen, G., Boussiba, S., Vermaas, W.F.J. (1993) Synechocystis sp. PCC 6803 strains lacking photosystem I and phycobilisome function . Plant Cell **5:** 1853-1863.

Sobotka, R., Komenda, J., Bumba, L., Tichy, M. (2005) Photosystem II assembly in CP47 mutant of Synechocystis sp. PCC 6803 is dependent on the level of chlorophyll precursors regulated by ferocheletase. J Biol Chem **280:** 31595-31602.

Thiel,T. (1994) The Molecular Biology of Cyanobacteria, pp. 581-611, Kluwer Academic Publishers, Dordrecht, Boston.

Van de Meene, A.M.L., Hohmann-Marriott, M.F., Vermaas, W.F.J., Roberson, R.W. (2005) The three-dimensional structure of the cyanobacterium Synechocystis sp. PCC 6803. Arch Microbiol **184:** 259-270.

Vieira, J., Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene**19(3):**259–268.

Wiliams, J.G.K. (1988) Construction of specific mutation in photosystem II photosynthetic reaction center by genetic engineering methods in Synechocystis 6803. Methods Enzymol **167:** 766-778.

Xiong, J., Fischer, W.M., Inoue, K., Nakahara, M., Bauer, C.E. (2000) Molecular evidence for the early evolution of photosynthesis. Science **289:** 1724-30.

Xu, W., McFadden, B.A. (1997) Sequence analyssis of plasmid pCC5.2 from cyanobacterium Synechocystis PCC 6803 that replicates by a rolling circle mechanism. Plasmid **2:** 95-104.