UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1149

Engineering Transcriptional Systems for Cyanobacterial Biotechnology

DANIEL CAMSUND

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

ISSN 1651-6214
ISBN 978-91-554-8954-0
urn:nbn:se:uu:diva-223599

Dissertation presented at Uppsala University to be publicly examined in Häggsalen, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala, Thursday, 5 June 2014 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Brian Pfleger (University of Wisconsin, USA).

Abstract
Camsund, D. 2014. Engineering Transcriptional Systems for Cyanobacterial Biotechnology. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1149. 63 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-8954-0.

Cyanobacteria are solar-powered cell factories that can be engineered to supply us with renewable fuels and chemicals. To do so robust and well-working biological parts and tools are necessary. Parts for controlling gene expression are of special importance in living systems, and specifically promoters are needed for enabling and simplifying rational design. Synthetic biology is an engineering science that incorporates principles such as decoupling, standardization and modularity to enable the design and construction of more advanced systems from simpler parts and the re-use of parts in new contexts. For these principles to work, cross-talk must be avoided and therefore orthogonal parts and systems are important as they are decoupled by definition. This work concerns the design and development of biological parts and tools that can enable synthetic biology in cyanobacteria. This encompasses parts necessary for the development of other systems, such as vectors and translational elements, but with a focus on transcriptional regulation. First, to enable the development and characterization of promoters in different cyanobacterial chassis, a broad-host-range BioBrick plasmid, pPMQAK1, was constructed and confirmed to function in several cyanobacterial strains. Then, ribosome binding sites, protease degradation tags and constitutive, orthogonal promoters were characterized in the model strain *Synechocystis* PCC 6803. These tools were then used to design LacI-regulated promoter libraries for studying DNA-looping and the behaviour of LacI-mediated loops in *Synechocystis*. Ultimately, this lead to the design of completely repressed LacI-regulated promoters that could be used for e.g. cyanobacterial genetic switches, and was used to design a destabilized version of the repressed promoter that could be induced to higher levels. Further, this promoter was used to implement an orthogonal transcriptional system based on T7 RNAP that was shown to drive different levels of T7 promoter transcription depending on regulation. Also, Gal4-repressed promoters for bacteria were engineered and examined in *Escherichia coli* as an initial step towards transferring them to cyanobacteria. Attempts were also made to implement a light-regulated one-component transcription factor based on Gal4. This work provides a background for engineering transcription and provides suggestions for how to develop the parts further.

Keywords: Cyanobacteria, Synthetic biology, promoters, transcription, LacI, Gal4, Light-regulation

Daniel Camsund, Department of Chemistry - Ångström, Molecular Biomimetics, Box 523, Uppsala University, SE-75120 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-8954-0
urn:nbn:se:uu:diva-223599 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-223599)

Cover illustration

Photograph: *Tobias Jakobsson*

The engineered *lac* repressor expressing strain of the cyanobacterium *Synechocystis* loaded in a 96-well plate of the type used for fluorescence measurements.

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Huang, H.H.*, Camsund, D.*, Lindblad, P. and Heidorn, T. (2010) Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Research*, 38, 2577-2593.

II Heidorn, T., Camsund, D., Huang, H.H., Lindberg, P., Oliveira, P., Stensjö, K. and Lindblad, P. (2011) Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions. *Methods in Enzymology*, 497, 539-579.

III Camsund, D., Heidorn, T. and Lindblad, P. (2014) Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *Journal of Biological Engineering,* 8, 4.

IV Camsund, D., Lindblad, P. (2014) A LacI-regulated promoter for *Synechocystis* and its use for implementing a T7 RNA polymerase-based orthogonal transcriptional system. *Manuscript*.

V Camsund, D., Lindblad, P. and Jaramillo, A. (2011) Genetically engineered light sensors for control of bacterial gene expression. *Biotechnology Journal*, 6, 826-836.

VI Camsund, D., Lindblad, P., Jaramillo, A. (2014) Development of Gal4-regulated transcriptional systems in *Escherichia coli*. *Manuscript*.

* Joint first authors.

Contents

Introduction 11
Motivation: Global warming and solar energy 11
Cyanobacteria, solar-powered cell factories 11
Synthetic biology 13
Transcription 15
The cyanobacterial RNA polymerase 16
Global effects on transcription 17
Transcription factors 17
The *lac* repressor, LacI 18
The yeast Gal4 activator 18
Orthogonal transcriptional systems 19
Orthogonal promoters 19
Orthogonal transcription factors 19
Orthogonal RNA polymerases 20
Other factors of importance for gene expression 21

Aim 22
Transcriptional tools for cyanobacterial biotechnology 22

Methods 23
Construction of DNA constructs 23
Inferring promoter activity indirectly from measurements using fluorescent protein reporters 24
A model of fluorescent protein reporter gene expression 24
Design of transcriptional test constructs 27
Sources of cross-talk that can affect apparent promoter activity 28
Design considerations and controls 29
Copy number and genomic location of expression 30

Results & Discussion 31
A broad-host-range BioBrick vector for use in cyanobacteria (Papers I & II) 31
Fluorescent protein reporters, translational and post-translational tools (Papers I & II) 33
Fluorescent proteins as reporters in cyanobacteria 33
Tools for translational regulation 35
Post-translational tools 36

Transcriptional tools (Papers I, III-VI)....37
Native *Synechocystis* promoters (Paper I)....37
Introducing common *Escherichia coli* promoters into *Synechocystis* (Paper I)....37
Characterization of a library of artificial, constitutive promoters in *Synechocystis* (Paper III)....38
Engineering LacI-regulated transcriptional systems (Papers I, III & IV)....39
Development of Gal4-regulated transcriptional systems for light-regulation (Papers V & VI)....45

Conclusions & Future Perspectives....50
Tools developed in the present study and needs for future development....50
Towards portable orthogonal gene expression systems....51

Svensk sammanfattning....52

Acknowledgments....55

References....57

Abbreviations

IPCC	Intergovernmental panel on climate change
CO_2	Carbon dioxide
Synechocystis	*Synechocystis* PCC 6803
N. punctiforme	*Nostoc punctiforme* ATCC 29133
E. coli	*Escherichia coli*
PCC	Pasteur culture collection
ATCC	American type culture collection
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
TSS	Transcription start site
mRNA	Messenger RNA
RNAP	RNA polymerase
PCR	Polymerase chain reaction
CDS	Coding DNA sequence
RBS	Ribosome binding site
EYFP	Enhanced yellow fluorescent protein
GFP	Green fluorescent protein
LacI	*lac* repressor
lacO	*lac* operator
UASG	Upstream activating sequences for galactose
CAP	Catabolite activator protein
LOV	Light-oxygen-voltage domain
SDS	Sodium dodecyl sulphate
PAGE	Polyacrylamide gel electrophoresis

Introduction

Motivation: Global warming and solar energy

From pre-historic times through the industrial revolution to our modern age and most likely further on, it is clear that human technological and societal development requires ever increasing amounts of energy. However, how we obtain that energy clearly makes a big difference to our surrounding environment and hence our own well-being. Recently, the intergovernmental panel on climate change (IPCC) released the physical science basis part of its Climate Change 2013 report [1]. It is now considered extremely likely that human activity is the dominant cause of the recent global warming, and anthropogenic release of CO_2 remains one of the main drivers of climate change.

Fossil fuels are not only unsustainable in the sense that they are in limited supply, the continued combustion of fossil fuels releases large amounts of CO_2 into our atmosphere, exacerbating climate change problems from the CO_2 already released since the industrial revolution. Therefore, there is a need to identify and make available sustainable energy sources that do not contribute to global warming.

The Sun irradiates Earth with immense amounts of energy; every hour the Sun provides our planet with the equivalent of humanity's energy consumption for one year [2]. Thus, solar energy presents a nearly inexhaustible source of energy, if only we could efficiently harness it. In essence, solar energy can be captured and converted into electricity, which is difficult to store on a global scale, or converted and stored as chemical bond energy, a fuel.

This thesis describes my contributions toward enabling the use of cyanobacteria as solar-powered cell factories, by designing and characterizing molecular tools – with a special focus on regulation of gene expression.

Cyanobacteria, solar-powered cell factories

Being the oldest known photosynthetic organisms and chiefly responsible for the transformation of our atmosphere starting ca 2.8 billion years ago into the oxygen-rich air we breathe today [3], cyanobacteria has changed the world previously and may yet do so again. As photosynthetic bacteria able to

flourish in a wide range of habitats, differing for instance greatly in salinity, pH and temperature, and being amenable to genetic engineering, cyanobacteria are well-suited for use in diverse renewable biotechnology applications [4]. The fact that many strains tolerate or thrive in high salinity water is especially important for potential global scale cyanobacterial cultivation, as sea water-based cultures will not compete for fresh water with agriculture or other human consumption. In addition to fixing CO_2 from air, eliminating the need to supplement cyanobacterial cultures with sugars or other forms of fixed carbon, some strains can also fix nitrogen, removing the need to add fixed nitrogen or fertilizer [5].

The first cyanobacterium to be fully sequenced in 1996, *Synechocystis* PCC 6803 (*Synechocystis*) [6], is a unicellular strain with moderate tolerance to salinity (Figure 1A). It is a model organism for the cyanobacterial phylum and the study of plant-like photosynthesis and as such its metabolism [7] and genetics [8] have been extensively examined. Because of this wealth of knowledge, and the ease with which *Synechocystis* can be genetically modified, it serves as the primary model organism for the research presented in this thesis.

An example of a nitrogen-fixing strain is *Nostoc punctiforme* ATCC 29133 (*N. punctiforme*), a filamentous cyanobacterium with the capacity to fix nitrogen from air using an oxygen-sensitive enzyme complex known as nitrogenase [5]. Nitrogen fixation takes place in a minority of specialized cells called heterocysts, which have evolved a low-oxygen environment to protect the activity of nitrogenase. Oxygenic photosynthesis takes place in the vegetative cells, which constitute the majority (Figure 1B).

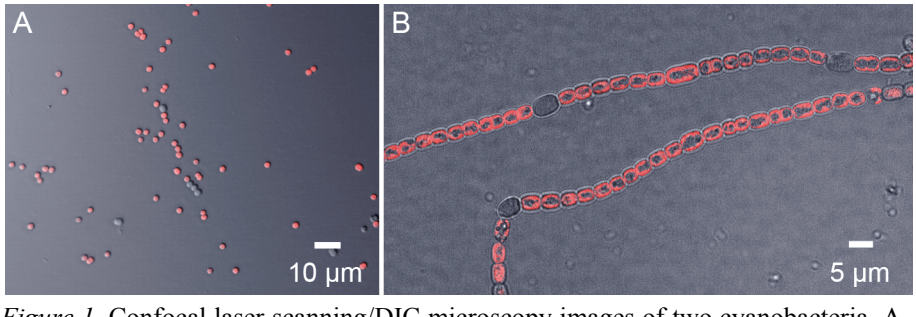

Figure 1. Confocal laser scanning/DIC microscopy images of two cyanobacteria. A. The unicellular *Synechocystis* PCC 6803. B. The filamentous *Nostoc punctiforme* ATCC 29133. The large, dispersed cells within the filaments without red autofluorescence are heterocysts, cells specialized to fix N_2 from air. The red autofluorescence comes from pigments in the phycobilisome/photosystem II complexes.

Several trends conspire to accelerate the use of cyanobacteria for renewable, solar-powered biotechnology. The recent increase in DNA-sequencing capacity allows for the sequencing of many different cyanobacterial genomes, and together with the rise of synthetic biology, this permits us to design and

engineer new traits into cyanobacteria. However, to speed up the development and enable the construction of more sophisticated cyanobacterial systems, well-characterized and robust biological parts, such as regulated promoters, must be developed and tested. This is where the burgeoning new engineering field of synthetic biology becomes important.

Synthetic biology

The focus on rationally designing and constructing new biological systems with intended properties from more basic biological parts and an understanding of how they function make synthetic biology an aspiring engineering field [9]. To help in design and to accelerate the process, engineering principles such as standardization, decoupling and modularity are at the core of synthetic biology [10]. Often, analogies are made with electronic systems, which are built up of small, standardized and modular parts to form circuits (Figure 2).

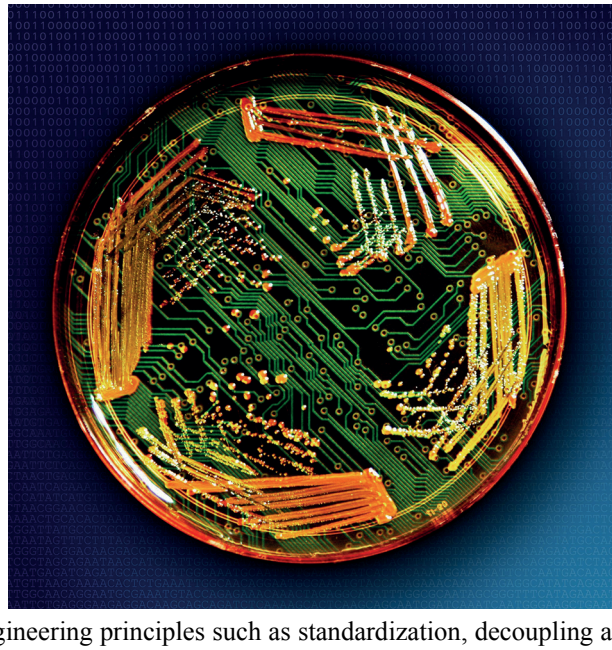

Figure 2. Engineering principles such as standardization, decoupling and modularity are at the heart of synthetic biology. Here, this is illustrated by electronic circuits made up by parts produced and combined using exactly those principles, overlaid by a Petri dish of bacteria, a living counterpart made up of genetic circuits. Image credit: Ivan Morozov (Virginia Bioinformatics Institute) / PLOS Synthetic Biology Collection.

While the engineering principles makes it different from its predecessor field of genetic engineering, synthetic biology is also broader in the sense that it

draws on multiple other fields for a comprehensive understanding of biological systems, including e.g. systems biology, biochemistry and biophysics, computational biology and design, molecular cell biology and genomics [11]. Still, the advance of synthetic biology would not be possible without the development of several technologies: First, DNA sequencing, which together with the massive increase in computational power makes high-throughput sequencing of whole genomes possible. Secondly, DNA synthesis, which enables us to design DNA sequences and implement systems that have never before existed on an increasingly large scale, as the technology and automation picks up speed.

Since living systems can produce valuable compounds and perform important services for us, there are many more applications of synthetic biology than producing biofuels. Some other examples are the production of the anti-malarial drug-precursor artimisinic acid in yeast [12] or bacteria that can detect and eliminate cancer cells [13]. Clearly, this is a field of multiple useful applications. In addition to these applications, synthetic biology also offers the possibility to learn more about and test our understanding of natural biological systems. Only when we can replicate precisely the function of a natural system, we know we truly understand it.

But why bother to build partly synthetic or fully artificial biological parts and cellular systems when we can use natural ones? Natural biological systems have evolved to perform a certain function in an organism, and it is how this function affects the organism's chances of reproduction that matters, not how the function is implemented. Further, natural biological systems are not decoupled, but have evolved to perform their function inside the cellular environment, in the myriad of interactions that occur with other biomolecules and on different levels of regulation [14]. For these reasons, the implementation of natural biological systems is often difficult to understand and consequently difficult to use or engineer. Unknown interactions or cross-talk between natural parts and other parts of the cell may cause your system to fail or perform less than optimally [15]. Therefore, decoupling is an important concept for synthetic biology. Further, together with standardization it enables modularity, which in turn enables the assembly of hierarchical systems all the way from single component genes and proteins up to pathways and whole cellular networks, analogously with the assembly of computers from the single resistor and transistor parts (Figure 3).

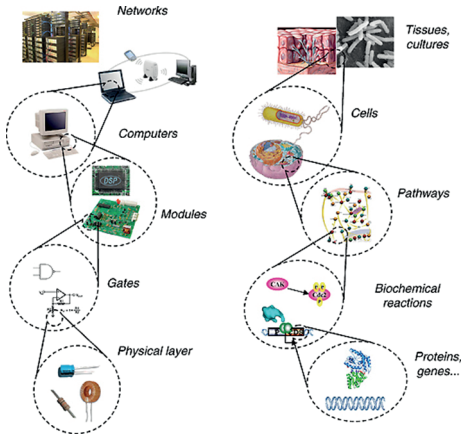

Figure 3. The assembly of hierarchical systems from decoupled parts. An analogy between the use of single electronic parts to build computers and single biological parts to build living cells. Adapted with permission from [10].

The use of orthogonal parts, i.e. parts that are not related to the implementation chassis or wholly artificial, can aid in the development of decoupled parts and systems. They do not share evolutionary history or a functional coupling, and hence confers a much lower risk for unwanted cross-talk or interactions [16]. However there is a need for the development of biological parts, especially promoters, both for commonly used biotechnological workhorses like *Escherichia coli* (*E. coli*) [17], but even more so for the photosynthetic cyanobacteria [4].

Transcription

The bacterial RNA polymerase consists of an apoenzyme made up of five subunits, $\beta\beta'\alpha_2\omega$. When it binds a sigma factor and forms the complete holoenzyme, $\beta\beta'\alpha_2\omega\sigma$, it gains the ability to bind a promoter specifically and initiate transcription. The housekeeping sigma factor, $\sigma 70$, has four different conserved domains that identify different parts of a typical $\sigma 70$ promoter. Part 1.2 binds the discriminator, a sequence situated just downstream the -10 element, which is in turn bound by part 2. Part 3 of the sigma factor recognizes the extended -10 element, just upstream the -10 element, and finally part 4 binds the -35 element (Figure 4A) [18]. After the RNAP has bound with a sigma factor at the promoter and formed a closed complex, it proceeds to melt and load the promoter DNA around the -10 element down to the transcriptional start site (TSS) at +1. These loading steps are very rapid, and finally the RNAP holoenzyme ends up in the open complex, with the melted promoter DNA loaded and the downstream double-stranded DNA

held in place by a clamp-like structure formed by the β and β' subunits (Figure 4B).

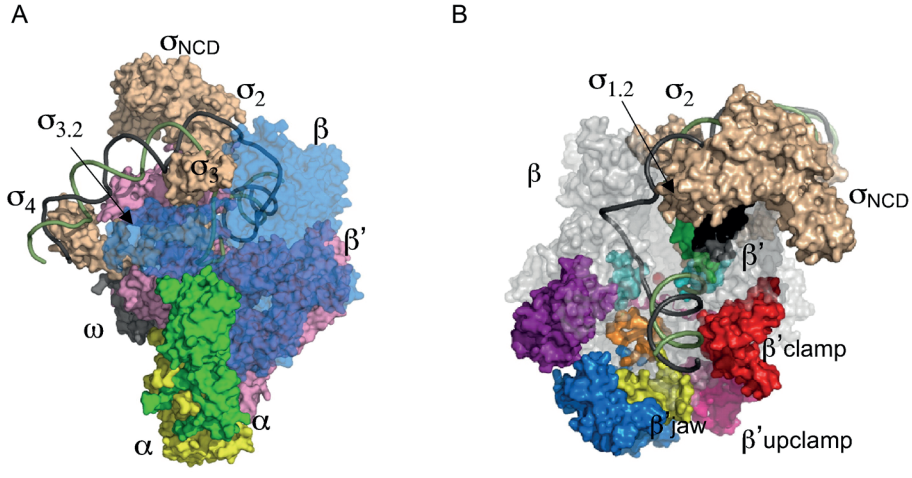

Figure 4. Structure of the *Escherichia coli* RNAP holoenzyme (ββ'α_2ωσ70) binding a promoter and forming an open complex. A. View showing the interactions between the promoter and the different parts of σ70: σ2, σ3 and σ4. B. The active site channel formed by the β, β' and σ1.2 domains, with the open transcriptional bubble, from -11 to +3 on the promoter, binding inside the cleft. The active site Mg^{2+} is situated next to the +1 transcriptional start site on the bottom of the channel and is colored as a faded red sphere. NCD, non-conserved domain of σ70. Adapted with permission from [18].

When the open complex has been reached, several steps of aborted transcription occurs while RNAP pulls in promoter DNA and tension builds up and decreases as the process is aborted. Finally, these tensions are released by the RNAP holoenzyme disengaging the promoter and it proceeds to elongation [18].

The cyanobacterial RNA polymerase

Cyanobacterial RNAP consists of the same subunits as the generic, enteric RNAP, except for the fact that the β' subunit is split into two parts. The γ subunit corresponds to the N-terminal part of the regular β' subunit, whereas the cyanobacterial β' subunit corresponds to the C-terminal part [19, 20]. It is unknown what the effect of the split β' is, if any, but differences in how enteric and cyanobacterial RNAP transcribe promoters have been observed [21] and an insertion in the cyanobacterial β' subunit has been suggested to be the cause. Later, it was suggested that the insertion is a jaw-like DNA-binding domain that interacts with the promoter [22], but this hypothesis has

yet to be tested. Further, a recent study examined the differences in Mn^{2+} tolerance between *E. coli* and cyanobacterial RNAP. While Mn^{2+} is toxic for most bacteria as it can replace the RNAP active-site Mg^{2+} ion, cyanobacteria need Mn^{2+} at higher intracellular concentrations for maintaining the photosystems. By comparing the activities of the two RNAP systems in vitro, it was concluded that the cyanobacterial RNAP transcribes its DNA slower but with higher fidelity [23]. They also suggested that the β' insertion of cyanobacterial RNAP could be responsible for the slower but more precise transcriptional elongation.

Global effects on transcription

Sigma-switching is an adaptive mechanism that allows bacteria to adapt to new environmental conditions or different types of stress. Most alternative σ-factors belong to the σ70-family, however there are examples of σ-factors belonging to the σ54-family, which generally require ATP-driven activators to unwind the promoter DNA [24].

Cyanobacteria only have sigma factors belonging to the σ70-family [25] [26] but those on the other hand can be divided into three groups. Group one consists of the primary sigma factor SigA, which corresponds to σ70 in *E. coli*, group two consists of non-essential sigma factors that provides a mechanism for environmental adaption, and group three sigma factors are involved in specific stress-survival regulons [27]. The primary sigma factor SigA binds to the same type of σ70-promoters as the *E. coli* σ70 factor does, consisting of conserved -35 and -10 elements, plus the other elements mentioned above. During stress however, the alternative group 2 sigma factors are expressed and partially replace SigA in the RNAP holoenzyme. This steers RNAP towards specific type 2 promoters, that only consist of a -10 element and distal enhancers, to initiate enhancer-stabilized stress responses.

Another global actor on gene expression is the circadian rhythm. It has been found that about half, or 30-64%, of all genes are rhythmically expressed, and initial evidence suggests that this regulation mainly occurs through modulation of DNA topology [28].

Transcription factors

Transcription factors (TFs) are important regulators of bacterial metabolism and behavior. In *E. coli*, for example, 6% of the total gene count is made up by different types of TFs. Further, TFs generally consist of a DNA-binding domain and a sensor or response type of domain, and can act on a global scale, like architectural DNA-binding proteins, or on a specific local scale, such as repression of a particular gene under a certain stimuli [24]. Except for being involved in the regulation of DNA-topology, TFs are generally

repressors or activators of transcription. Repressors mainly work by sterically hindering RNAP from binding the promoter, or destabilizing bound RNAP, meaning that they are most efficient when their operators overlap with the core promoter. Activators, on the other hand, act by stabilizing the binding of RNAP to the promoter, so their operators generally do not overlap with the core promoter [29].

The *lac* repressor, LacI

LacI is the repressor of the *E. coli lac* operon that binds its pseudo-palindromic *lac* operators (*lacO*) through the DNA-binding domains of one dimer in a tetramer that consists of a dimer of dimers [30, 31]. Because it is a homo-tetramer, the two dimers can simultaneously bind two spatially separated *lacO* while bending or looping the DNA in between [32] (Figure 5).

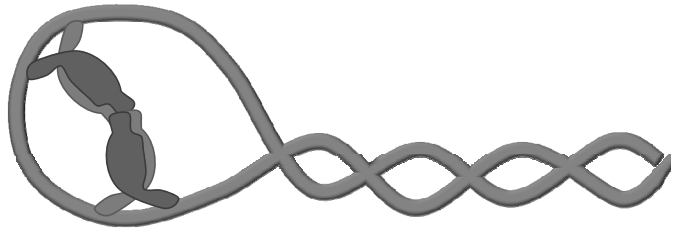

Figure 5. Conceptual image of a LacI tetramer binding two *lacO* inside an apical loop on plectonemically supercoiled DNA. (Paper III).

This dual *lacO*-binding and DNA-looping leads to cooperativity in the repression and induction behavior [33], improving repression and causing its regulation to be more switch-like upon induction with the wild-type inducer allolactose, or the artificial, stable lactose analog isopropyl β-D-1-thiogalactopyranoside (IPTG). Further, there are three native *lacO*, *lacO1*, *O2*, and *O3* that bind LacI with decreasing affinity in that order [34]. Further, the perfectly symmetric artificial operator *lacOsym* (or *lacOid* for ideal) binds LacI with an even stronger affinity [34]. Even though it constitutes a well-studied transcriptional system, which is used in many different variants both for engineered gene expression [35] or as a model-system for studying gene expression and DNA-looping [36], different aspects of its function are still debated [37, 38].

The yeast Gal4 activator

The Gal4 transcription factor mainly functions as an activator of transcription in the baker's yeast *Saccharomyces cerevisiae* [39]. It forms a homodimer that binds its upstream activating sequence (UASG), a partially palindromic operator, though a Zn^{2+}-containing DNA-binding domain [39].

Truncated versions of Gal4, where the domains important for its function as an activator in yeast have been removed, still bind its UASG operator specifically and with high affinity [40]. Therefore, these truncated versions of Gal4 are frequently used for different applications requiring a DNA-binding protein, such as two-hybrid assays [41].

Orthogonal transcriptional systems

To reduce the risk for cross-talk with the native transcriptional system of any system implementation chassis, orthogonal transcriptional parts or whole systems can be introduced. Orthogonal parts can be orthogonal in different degrees, ranging from e.g. transcription factors from related strains of bacteria that are different enough in the new host to decrease the risk for cross-talk, to TFs from very distantly related bacteria, to fully artificial, synthetic parts that have never existed before and therefore have a maximum degree of orthogonality.

Orthogonal promoters

Artificial promoters that bind the host's own RNAP can be considered orthogonal, as these DNA-sequences have not evolved in the chassis and hence are very unlikely to contain operator sequences or other target sequences that would cause unwanted interactions and cross-talk.

Orthogonal transcription factors

Transcription factors that are imported from an exogenous host, e.g. the transfer of LacI from *E.coli* to a distantly-related cyanobacterial strain, represent orthogonal transcription factors. Also fully or partially artificial TFs, like engineered zinc-finger DNA-binding proteins [42], or the recently implemented Cas9-system derived from CRISPRs [43], constitute orthogonal TFs. These are all unlikely to find specific operators to bind in the genome of the new chassis. Nonetheless, there can be unexpected cross-talk, e.g. through unspecific binding at sequences that are randomly similar to the TFs specific operator.

Another class of orthogonal TFs are light-regulated. These types of TFs are interesting for applications that require high spatial resolution, for instance in biomedicine or targeted therapeutics, but also for regulating gene expression in biotechnology. Compared to regularly used TFs like LacI, that require the addition of chemicals for induction, light-regulated TFs are preferable in situations when induction needs to be temporary, or e.g. when the induced system is solar-powered and relies on day and night cycles (Paper V). Most engineered light-regulated gene expression systems can be divided

into two categories: Two-component systems consisting of histidine kinases and a response regulator, or one-component systems that consist of the transcription factor itself or a partner, who dimerize upon light-stimulation (Figure 6).

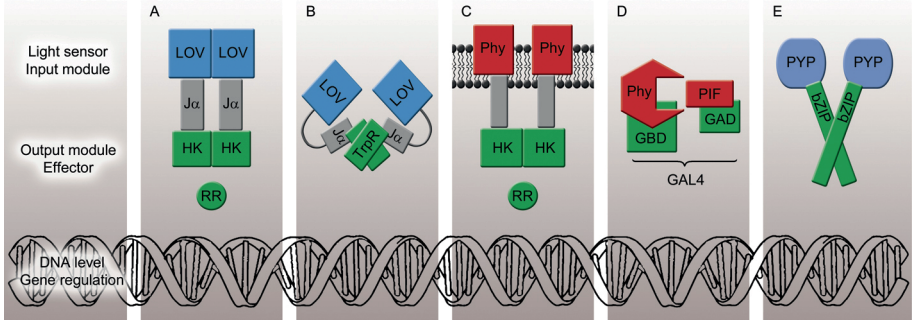

Figure 6. Light-regulated orthogonal systems for regulation of gene expression. A. The soluble, blue-light-sensitive YF1 histidine kinase chimera that activates the activity of its response regulator FixJ [44]. B. The soluble LovTAP light sensor that dimerizes upon blue-light stimulation and binds the *E. coli* TrpR binding-site [45]. C. The membrane-bound red-light sensitive chimeric Cph8 histidine kinase that activates the activity of its response regulator OmpR [46]. D. Example of a red-light sensitive yeast-two-hybrid inspired light-regulated dimerization system that binds the Gal4 UASG operator upon stimulation [47] E. A blue-light sensitive dimerization system that uses the photoactive yellow protein combined with a leucine-zipper DNA-binding protein [48] (Paper V).

Orthogonal RNA polymerases

Finally, the most orthogonal gene expression system is one that does not rely on the hosts own RNAP at all, or otherwise minimally. By using an orthogonal RNAP that does not recognize the host's own promoters, and for which the host's RNAP does not recognize the orthogonal promoters, the risk for cross-talk is strongly reduced, and combined with likewise orthogonal TFs the system is almost completely decoupled from the chassis own transcriptional systems. Obviously, even an orthogonal RNAP will still depend on the cell for substrates and energy, and at least the first orthogonal RNAP has to be produced by the cell's own machinery before it becomes self-maintaining.

One such orthogonal RNAP (O-RNAP) system is the phage T7 RNAP and its promoters. T7 RNAP does not recognize the host's promoters, and vice versa, the host's RNAP does not recognize the T7 promoters [49]. It is conceivable that similarly orthogonal RNAP can be found in other viruses or possibly other very distantly related organisms, to expand the toolbox of O-RNAPs.

Other factors of importance for gene expression

To conclude the introductory part, there are other elements and factors that are important for gene expression that are not covered in this thesis, or else mentioned very briefly. This includes for example the effect of secondary structures on mRNA and translation [50] and the engineering of these for rational design of translation initiation [51], stability of mRNAs [52, 53], the presence of small RNAs and antisense transcription in e.g. the cyanobacterium *Synechocystis* [54] plus codon choice and internal ribosome stalling [55].

Aim

Transcriptional tools for cyanobacterial biotechnology

For this thesis, there are three aims:

I The development of tools required to implement and characterize transcriptional or other systems. This includes e.g. vectors for DNA transfer and expression, ribosome binding sites and fluorescent protein reporters.

II To develop tools required to control cyanobacterial gene expression, with the ultimate aim of simplifying metabolic engineering for renewable biotechnology.

III To evaluate the utility of the developed systems and identify potential developments that could improve the systems further.

Methods

Construction of DNA constructs

After the discovery of the first type II restriction enzyme [56] that enabled recombinant DNA technology, our ability to assemble different pieces of DNA has evolved considerably.

In 2003, Tom Knight proposed the by now famous BioBrick system in the report "Idempotent vector design for standard assembly of biobricks" [57], which opened up the use of BioBricks from the at present large and growing biological parts database iGEM Registry of Standard Biological Parts [58]. This technique, which enables the continuous addition of BioBrick parts to the start or end of another BioBrick part or assembly using restriction enzymes, was used extensively throughout the present work. The continuous addition of new parts at the ends is possible because the assembly process recreates the restriction enzyme target sites upon ligation. A development of the BioBrick assembly system that enables the simultaneous BioBrick assembly of two parts into a vector, 3A-assembly [59], was also used extensively throughout this work as it speeds up the BioBrick assembly process.

The polymerase chain reaction (PCR) [60] has revolutionized all of biology in more ways than it is possible to mention, and it continues to be a practical method e.g. for amplification and modification of DNA parts. PCR in combination with mutagenic primers was used to produce almost all the different promoter reporter constructs described in the present work, and together with overlap-extension [61] it was used to produce new versions of artificial transcription factors.

Another method used in combination with PCR that was used for the simultaneous assembly of several DNA parts at once was one-step isothermal assembly [62], also known as Gibson assembly. The method is based on 20-40 bp sequence-overlaps between parts to be assembled, which are normally added to the parts by PCR. The overlapping parts are then mixed with an exonuclease, a DNA polymerase and a heat-tolerant ligase in a reaction mixture that is incubated at 50 °C for under one hour. During this time, the exonuclease will chew back the 5' ends of all parts, producing complementary sticky-ends. As the exonuclease is not heat-tolerant, it will lose activity and the polymerase will fill in the gaps of the annealed parts, where the nicks are finally filled in by the ligase to produce circular double-stranded DNA molecules. Also, these circular DNAs will accumulate, as the remaining active

exonuclease only targets 5' ends. To aid in the design of overlaps for Gibson assembly, the j5 DNA assembly design automation software was used [63].

Finally, DNA synthesis ordered from different commercial providers was used as a complement to the above methods when larger pieces of synthetic DNA were required. The GeneDesigner software [64] was used for codon-optimization of synthetic coding sequences.

Inferring promoter activity indirectly from measurements using fluorescent protein reporters

The use of reporters, especially fluorescent proteins, for estimating promoter activity is wide-spread [65]. It is a practical approach with several advantages. It is possible to measure activities or changes in activities in real time in living cells, the emitted light is easy to detect non-invasively, averages over whole populations can be measured quickly through the use of e.g. plate readers, and single cell measurements can be done with e.g. flow cytometry or microscopy. Further, as compared with enzymatic reporters, fluorescent proteins do not require a substrate except for the excitation light, meaning that promoter activity estimations are not biased by e.g. substrate limitation or the need to add substrates or other chemicals that could affect the system to be characterized.

However, there are several intermediate steps between the start of transcription and the final active, fluorescent protein that is the subject of measurements. This complicates our ability to draw conclusions about promoters based on fluorescence measurements of expressed proteins – how can we be sure that the effect we see is not on e.g. the post-transcriptional or translational level? The best, of course, would be if we could measure promoter activity directly, for instance by detecting directly the numbers of RNA polymerases that pass by the promoter per second [66]. At present, this is not practically possible though. Instead, one can make use of models of gene expression to understand the whole process, from transcription to the final fluorescent protein, to make more informed experimental designs, and to help in interpreting the data.

A model of fluorescent protein reporter gene expression

Gene expression encompasses many steps, from the binding of RNAP to the promoter, initiation and elongation of transcription, translation of the mRNA and folding of the resulting peptide chain into a mature protein, which can be modeled in many ways [67, 68]. While the nature of gene expression is stochastic, meaning that individual players in gene expression such as transcription factors and RNAP diffuse more or less randomly through the cell or

along the DNA to find their targets [69], simplified deterministic models are still useful for understanding e.g. the expression of fluorescent proteins [68].

A previously developed deterministic model of gene expression makes use of differential equations describing the separate steps of transcription, translation, and maturation of fluorescent proteins [70]. While it is not being used for simulations in this thesis, the model serves as a useful description of the cellular processes that affect the amount of final fluorescent protein that we use to infer promoter strengths. First, the change of mRNA-levels in time was described by the contribution from transcription minus degradation and dilution due to cell division:

$$\frac{d}{dt}RNA = \mathrm{P}\psi\mathrm{DNA} - (\delta_\mathrm{r} + \mu)\mathrm{RNA} \quad (1)$$

where RNA is the amount of mRNA, P the promoter activity, ψ the number of promoters per vector, DNA the copy number of the vector, δ_r the mRNA degradation rate and μ the growth rate. From Equation 1 it is apparent that even if we could measure the amount of mRNA produced by a specific promoter, for example by lysing the cells and performing quantitative reverse-transcription PCR, the obtained value would still not be a perfect measure of promoter activity, as the amount of mRNA per cell also depends on the stability of the mRNA and the cellular growth rate.

Secondly, the change of immature, non-fluorescent proteins in time was described by the contribution from translation minus the protein that is maturated, degraded, or diluted due to cell division:

$$\frac{d}{dt}PROT^{(n)} = QRNA\Omega - (m + \delta_p + \mu)PROT^{(n)} \quad (2)$$

where $\mathrm{PROT}^{(\mathrm{n})}$ is the immature protein, Q the ribosome binding site affinity or strength, Ω the number of ribosome binding sites per mRNA molecule, m the maturation rate of immature protein into mature, fluorescent protein, and δ_p the protein degradation rate. Finally, the change of the amount of mature, fluorescent protein in time, $\mathrm{PROT}^{(\mathrm{f})}$, was described by the contribution from maturation of immature protein minus degradation and cell division dilution:

$$\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{PROT}^{(\mathrm{f})} = \mathrm{mPROT}^{(\mathrm{n})} - \left(\delta_p + \mu\right)\mathrm{PROT}^{(\mathrm{f})} \quad (3)$$

From Equations 2 and 3 it is apparent that the measured fluorescence from a fluorescent protein will depend on the maturation rate of the immature protein, its degradation rate and dilution due to cell division, except for the promoter activity (Equation 1) that we desire to measure. While engineered versions of GFP and other modern fluorescent proteins have relatively high maturation rates, making their use as promoter activity reporters practically

possible, their β-barrel structure is very stable, making them resistant to both chemical denaturation and proteolytic degradation [65]. Their stability results in a low degradation rate, meaning that the production and maturation of new fluorescent proteins will mainly be balanced by dilution due to cell division. Hence, the level of fluorescent proteins per cell will depend to a great extent on the cellular growth rate. In fact, for the constitutive expression of proteins in bacteria, it was shown that while the number of mRNAs and the number of proteins per cell increase for higher growth rates, the protein concentration goes down. This decrease in concentration in spite of the increased number of molecules per cell was explained by the markedly increased cell volume at fast growth [71]. This can be illustrated by following the population average of fluorescence per cell over time for an *Escherichia coli* culture expressing a fluorescent protein constitutively (Figure 7). In this experiment, an over-night culture of stationary phase cells that had accumulated a fluorescent protein were diluted 200 times into fresh medium in the morning and the development of fluorescence and growth was followed at several time points during the day.

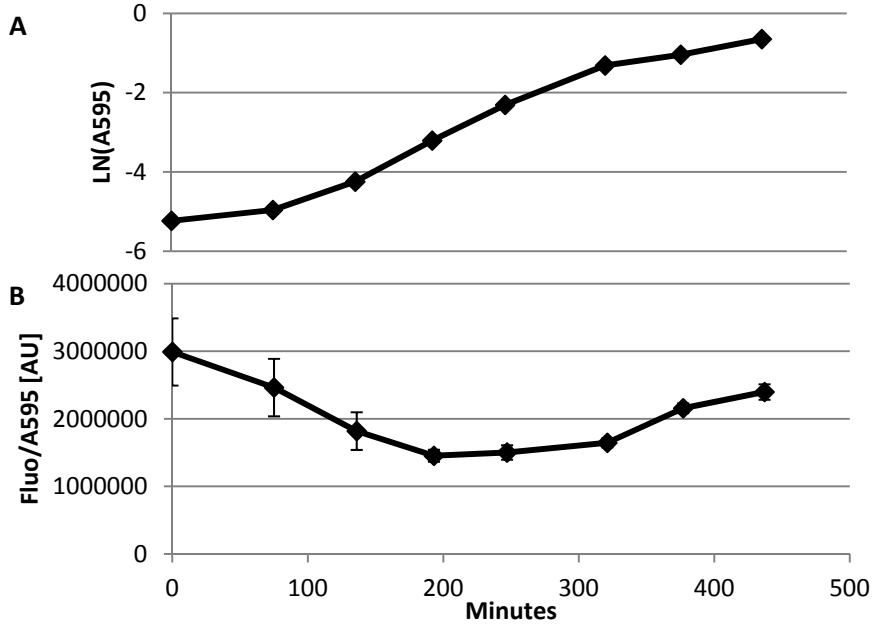

Figure 7. Growth and fluorescence per cell versus time for a growing *E. coli* culture constitutively expressing a fluorescent protein. A. Logarithmic growth curve of absorbance at 595 nm. B. Average population fluorescence per cell expressed as fluorescence normalized to absorbance at 595 nm. Error bars represent standard deviation (n=3).

The higher concentration of fluorescent proteins, which accumulated overnight as the growth rate slowed down and the cells entered the stationary

phase, can be seen in the relatively high level of fluorescence per cell in the beginning of the experiment. Then, as the cells start to divide faster after an initial lag-phase (Figure 7A), the fluorescence per cell starts to drop because of dilution due to cell division. After almost 200 minutes there is a temporary steady-state in fluorescence per cell levels, as the production of new fluorescent proteins is balanced by cell division (Figure 7B). This steady-state can be interpreted in Equation 3 as the time-derivative of $PROT^{(f)} = 0$. However, as the cells leave the first exponential growth phase and enter a second slower one, which can be seen in the decreased slope of the logarithmic growth curve after about 320 minutes (Figure 7A), the fluorescence per cell starts to increase (Figure 7B), due to decreased dilution from cell division.

The steady-state that occurs when the creation and destruction terms of the fluorescent protein in Equation 3 equal each other makes it possible to calculate the promoter activity if the other parameters are known or can be measured [68, 70]. A similar model was also used to calculate an estimate of polymerases that pass by the promoter per second, or PoPS [66], illustrating the possibility of extracting quantitative promoter activities from fluorescent protein measurements.

For qualitative comparisons of promoter activity, or relative comparisons, it is important to measure cellular fluorescence in the growth phase of interest for the comparison, or during a potential steady-state, as this value is characteristic for the system and the specific environmental conditions. Caution should be taken when comparing the activities of promoters characterized in different growth conditions, as differences in growth rate, or metabolism, could make two cultures with identical promoter-reporter constructs look different.

Design of transcriptional test constructs

To make promoter characterization results comparable, there are several design factors that need to be taken into consideration. Even if the test construct consists only of one promoter and a fluorescent reporter, and growth phases and conditions have been taken into account according to the previous section, the design of the construct will have effects on the results. Also, transcription factors or RNA polymerases that are part of the transcriptional system often need to be co-expressed from the same vector as the promoter test construct. This presents additional challenges when designing promoter test constructs.

Sources of cross-talk that can affect apparent promoter activity

There is a risk for cross-talk between several elements that constitute a typical promoter test construct that uses a reporter protein (Figure 8).

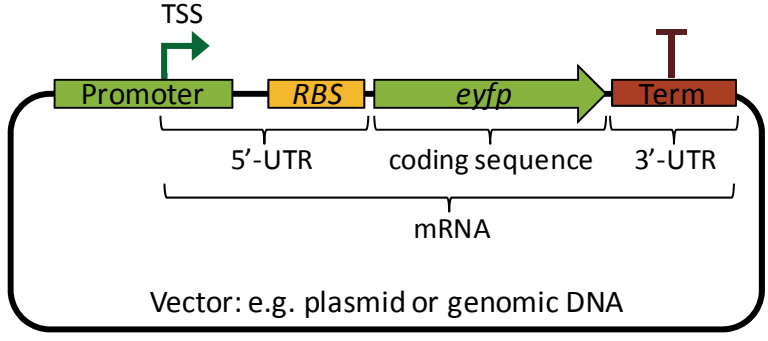

Figure 8. The anatomy of a general promoter test construct. Abbreviations: TSS (transcriptional start site), UTR (untranslated region), RBS (ribosome binding site), *eyfp* (enhanced yellow fluorescent protein gene), and Term (transcriptional terminator).

Indeed, a recent combinatorial study where many different promoters and 5'-UTRs were combined with two different fluorescent reporters found that the largest part of the variation in translation efficiency could be explained by the choice of promoter, and that mRNA abundance was mostly explained by the 5'-UTR sequence [72]. This, of course, poses a large problem for the reliable characterization of promoters, when it is not certain if the promoter sequence to be analyzed affects other stages of gene expression not involved at all in the process of transcription, and because the apparent activity of the promoter will depend both on the specific 5'-UTR and the reporter gene coding sequence.

Perhaps disturbing from a biological engineer's perspective, promoters are not always well-defined. Often, there are multiple transcriptional start sites, producing mRNAs with different 5' ends, or the promoter sequence continues downstream the TSS, contributing with excess sequence to the 5'-UTR (Figure 8). This leads to unpredictable effects on mRNA stability, as the mRNA sequence itself will affect its stability through differential association with RNase E and subsequent degradation [52]. Further, the 5'-UTR is important for ribosome binding and initiation of translation, and interactions between the part of the promoter that contributes to the 5'-UTR and the ribosome binding site (RBS), or the first part of the coding sequence, could lead to the formation of ribosome-blocking secondary structures [50].

To avoid these problems, which contribute to unnecessary cross-talk between promoters and downstream parts of an expression cassette, standardized promoters that always end with its TSS has been suggested. Going even

further, the same study also developed a bi-cistronic system for translation that prevents 5'-UTR secondary structures from blocking translation of the gene of interest, which works even for different coding sequences [73]. Other ways of solving the problem of cross-talk between promoter parts and the 5'-UTR includes adding self-cleaving ribozymes to the RBS, which will truncate the mRNA and remove any contribution to the 5'-UTR from the promoter [74]. These examples of reducing cross-talk between parts important for gene expression can be viewed as functional insulation or decoupling, which are concepts crucial to the success of the rational design of ever larger and more complex genetic circuits.

Design considerations and controls

For the more humble goal of characterizing promoters, a minimal requirement is to only compare fluorescence per cell values of promoter test constructs that share the same 5'-UTR and reporter gene coding sequences. For regulated promoters with different 5'-UTR and reporter gene coding sequences, the activity ratios of the regulated and the un-regulated promoter can also be compared between different promoters, as all post-transcriptional steps are assumed to be the same with and without regulation and hence these effects will cancel [36].

For combining several transcriptional units directionality also becomes important. Since terminators are seldom 100 % efficient at stopping elongating RNA polymerases, e.g. the commonly used double terminator BBa_B0015 has a forward termination efficiency of 0.97-0.984 and a reverse efficiency of 0.295-0.62 [58], read-through transcription from nearby promoters is a potential issue. Therefore, a design with several transcriptional units in a row is not to recommend if it is important that the second unit or later units are precisely regulated or transcribed at specific levels. In those cases, divergent designs, where two promoters are transcribing in different directions and separated by a spacer wide enough to avoid cross-talk between RNAP or TFs are suitable. Convergent designs, where two promoters lead transcription head-on are not preferable, since there is a substantial risk that elongating RNAP may collide and cause premature termination of transcription [75].

To detect disturbing cross-talk when characterizing gene expression systems, appropriate controls are necessary. For regulated promoters controlled by a transcription factor, combined designs with two or several transcriptional units of which one is the subject of measurements, or for controlling the effect of growth conditions on putative fluorescence steady-states, constitutive promoters are indispensable controls. Being unregulated and always active, comparing the fluorescence per cell levels from constitutive promoters allows the detection of cross-talk from induced systems like LacI and IPTG (Paper III) or can serve as controls for potential cross-talk between

closely spaced, divergent promoters (Paper VI). Further, minimal and orthogonal constitutive promoters can also function as excellent reference promoters, as they contain no functional sequences that transcriptional regulators can bind (Paper III).

Finally, the measurement of fluorescence itself is an important experimental design consideration. If population-wide averages of fluorescence are sufficient to answer all hypotheses or provide adequate characterization data, instruments such as plate readers are useful. If single-cell data is required, for example to detect two or more subpopulations in the experimental cultures, flow cytometry or FACS is an appropriate method. Lastly, if there is a need for both single-cell data and temporal resolution, a microfluidics platform combined with automated fluorescence microscopy is suitable.

Copy number and genomic location of expression

Gene dosage is another design criterion that merits consideration for any transcriptional system. The number of promoters per cell is not only important from a strength of expression perspective, where a higher gene dosage usually leads to higher expression levels [76], but also important for regulation. For example, the cellular concentration of repressors may be sufficient to repress a promoter under low copy number, but may be insufficient and cause a higher basal promoter activity level when the target promoter exists in too many copies.

The location of the expression construct is a factor that is connected to the copy number, as the copy numbers of plasmids and genomes normally differ. While there are plenty of plasmids available for use with a great span of different copy numbers, bacterial chromosomes may have copy numbers varying from 1-2 for *E. coli*, to 12 or up to between 40 and 200 depending on the growth phase, for *Synechocystis* [77, 78]. Further, the location of expression may affect the cellular localization of the gene product, causing for example different local concentrations of a repressor in different locations of the cell depending on where its gene is expressed [79]. Except for transcriptional regulation, this would be of importance for the expression of enzymes involved in the same pathway, which one ideally would like to be co-localized. Another less obvious but not totally unexpected factor is that the gene copy number of a gene inserted into the bacterial chromosome will depend on the distance to the origin of replication. The closer it is to the origin, the higher the gene copy number will be because of more frequent replication, and vice versa, the closer it is to the replicative terminus the lower the copy number will be [71]. Finally, global regulators of DNA topology like circadian rhythms in cyanobacteria affect the expression of many genes [28], and it is likely that this effect will be different in different locations of the genome, depending on the local state of DNA packing.

Results & Discussion

A broad-host-range BioBrick vector for use in cyanobacteria (Papers I & II)

The RSF1010-replicon of the IncQ incompatibility group has the ability to replicate in many gram-negative bacteria, making it one of the most widespread replicons known [80]. It has even been confirmed to replicate in a gram-positive bacterium [81]. To utilize this capability, enabling us to characterize the same constructs, or introduce the same metabolic circuits, in several different cyanobacterial strains, we engineered a chimeric vector from the standardized BioBrick plasmid pSB1AK3 [58] and the broad-host-range RSF1010-carrying pAWG1.1 plasmid (Figure 9)

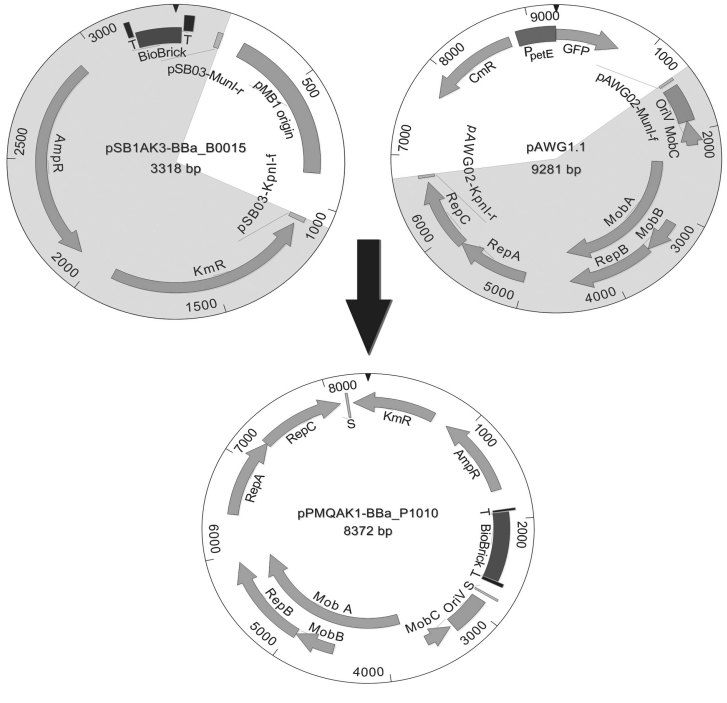

Figure 9. Assembly of the pPMQAK1 broad-host-range BioBrick vector from the pSB1AK3 BioBrick plasmid [58] and the RSF1010-carrying pAWG1.1 plasmid using PCR in combination with MunI and KpnI. (Paper I).

The new plasmid was named pPMQAK1, which stands for Photochemistry and Molecular science, the IncQ incompatibility group, Ampicillin and Kanamycin/neomycin resistance, and finally version number 1.

To confirm if it could replicate in different cyanobacterial strains such as the unicellular *Synechocystis* or the two filamentous cyanobacteria *Nostoc* sp. PCC 7120 and *N. punctiforme*, and simultaneously test a commonly used LacI-repressed promoter, P*trc*, we transferred a P*trc*-GFP reporter construct into these three different strains of cyanobacteria using conjugation. Shortly later we could confirm that pPMQAK1 indeed replicates in all three strains and that the P*trc* promoter works well for expressing high amounts of GFP per cell under unregulated conditions in all strains tested (Figure 10).

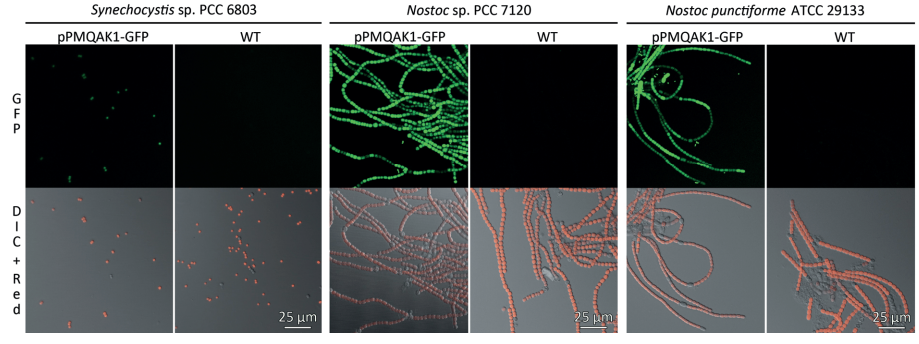

Figure 10. Confocal laser scanning/DIC microscopy images illustrating the replicative capability of pPMQAK1 by means of a constitutive GFP reporter cassette in three cyanobacterial strains. The red autofluorescence comes from pigments in the phycobilisome/photosystem II complexes. (Paper I).

Since these tests, we have been using pPMQAK1 routinely both for direct cloning of BioBrick parts or using 3A-assembly. However, its relatively low copy number in combination with large losses during purification and its apparent resistance to restriction digestions does not make it an ideal cloning vector. For expression purposes, it works well both in *E. coli* and cyanobacteria, and has an expected copy number of 10 in the former [82] and 10-30 in the cyanobacterium *Synechocystis* [83, 84]. That should be compared to the copy number of the *E. coli* chromosome, which is between 1-2 depending on growth phase, and the *Synechocystis* chromosome, which has 12 copies [77] to 40-200, depending on growth phase [78]. Finally, other labs have successfully transferred pPMQAK1-derivates to *Synechocystis* using electroporation [85], which makes the use of conjugal *E. coli* strains unnecessary and the transfer process somewhat cleaner, as *E. coli* contamination may under some circumstances be difficult to remove.

Fluorescent protein reporters, translational and post-translational tools (Papers I & II)

To develop the basic tools necessary for promoter characterization in cyanobacteria, we continued with investigating the potential interference of photosynthetic pigments with fluorescent reporters, the effect of different ribosome binding sites on fluorescent protein levels, and increased protein degradation rates through the use of degradation tags.

Fluorescent proteins as reporters in cyanobacteria

Cyanobacteria possess many pigments connected to photosynthesis that absorb visible light at wavelengths of value for exciting fluorescent proteins or for detecting their emitted light. To investigate if photosynthetic pigments would interfere with fluorescent protein (FP) spectra, we decided to characterize several FPs inside cyanobacteria and compare with values obtained for pure proteins. The resulting emission and excitation spectra for the three FPs Cerulean, GFPmut3B (GFP) and EYFP expressed inside *Synechocystis* (Figure 11) corresponded well to the previously reported values of maximum excitation and emission wavelengths.

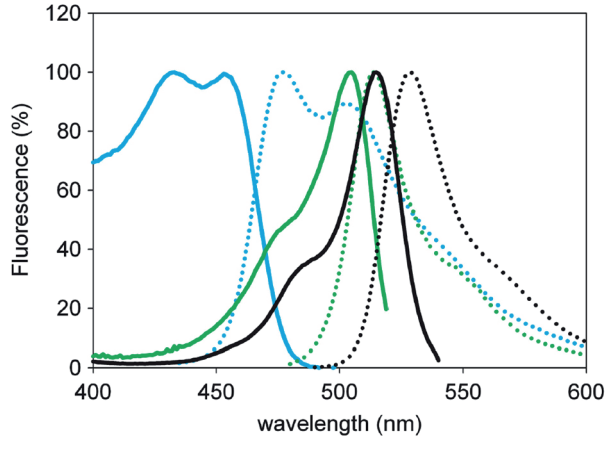

Figure 11. Excitation (solid lines) and emission (dotted lines) spectra for the fluorescent proteins Cerulean (blue), GFPmut3B (green), and EYFP (black) expressed in *Synechocystis*. All signals are normalized to the highest value as 100%. (Paper I).

Further, we performed an immunoblot study to compare the relative levels of denatured GFP per cell for several different promoters with their corresponding values of fluorescence per cell to detect any abnormalities in the pattern, implying that fluorescence does not correspond to protein levels. However, the patterns of GFP protein per cell detected by immunoblotting corresponded well to the measured fluorescence intensities per cell (Figure 12),

meaning that the FP reporters can well be used as reporters of gene expression.

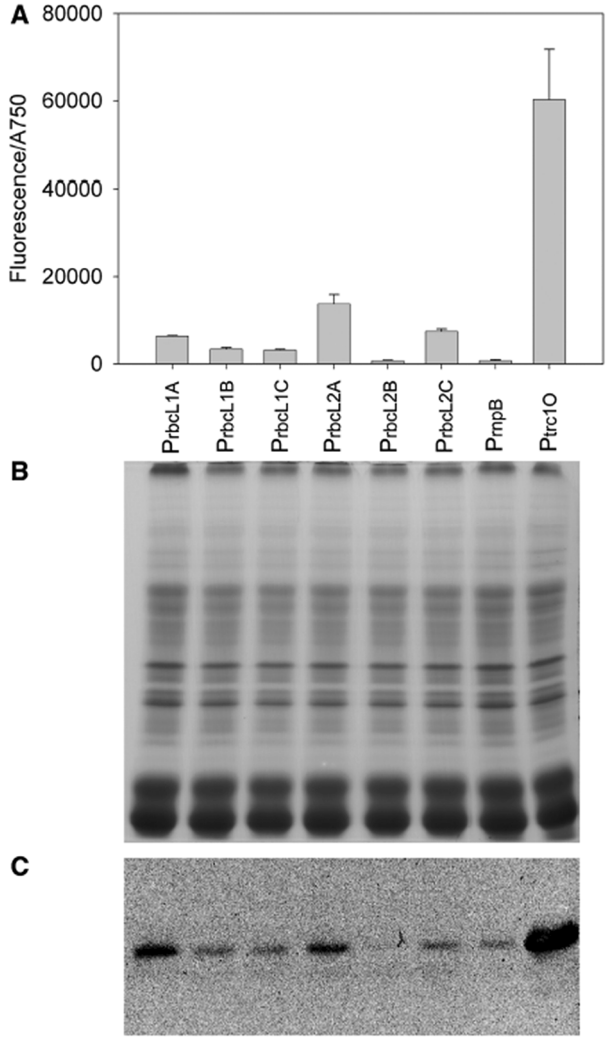

Figure 12. Comparison between measured fluorescence per cell levels and immunoblotting results measuring the relative amount of denatured protein for the same promoter reporter constructs in *Synechocystis*. A. The fluorescence per cell levels measured as fluorescence/absorbance at 750 nm for eight different promoter reporter constructs. Error bars represent standard deviations (n=3). B. Stained SDS-PAGE of the loaded amount of total protein for the immunoblot in C. C. Western blot signals from GFP-specific antibodies used to detect denatured GFP in the total protein samples. (Paper I).

In combination with the Western immunoblot study, the conserved excitation and emission spectra for the different FPs in *Synechocystis* mean that

FPs can be used to quantify promoter activities despite the presence of large levels of photosynthetic pigments in the cells.

Tools for translational regulation

As the processes of transcription and translation are tightly coupled in bacteria, ribosomes will start to assemble on the ribosome binding site of the mRNA in the order of seconds after the start of transcription [86]. Ribosome binding sites carry a core motif, the Shine-Dalgarno (SD) sequence 5'-GGAGG-3', to which the anti-SD sequence of the 16S ribosomal RNA subunit binds before initiation of translation. Several factors decide the strength of ribosome binding sites and the subsequent initiation of translation. The degree of complementarity to the SD-sequence is one factor, the distance between the SD and the start codon is another [86]. The folding status of the 5'-UTR into different secondary structures is a third factor, which can help to expose the RBS or hide it from the ribosome, preventing initiation of translation [50].

Compared to 57% of all genes in *E. coli*, a bare 26% of genes were found to have the core SD-sequence at the RBS in *Synechocystis* [87]. To test if translation initiation in *Synechocystis* can be improved by changing the core RBS to become more similar to the 16S anti-SD sequence, we designed RBS*: 5'-TAGTGGAGGT-3'. To test RBS*, we assembled it and three other artificial RBSs commonly used in *E. coli* with the P*trc* promoter and a GFP reporter gene and cloned it on pPMQAK1. When tested in *E. coli* and *Synechocystis* and compared, most of the RBS differed in the measured strength between the two species, and RBS* turned out to be the strongest RBS for *Synechocystis* (Figure 13). However, as we have seen in the previous sections, there is room for ample cross-talk between 5'-UTRs and coding sequences. Hence, it would be interesting to test these very same RBS again but with several different reporters, to see how large the variation is.

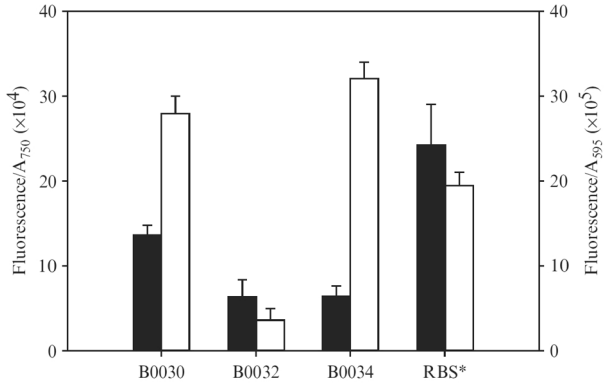

Figure 13. Test of three common RBS, BBa_B0030, BBa_B0032, BBa_B0034 and RBS*, using constitutive expression of GFP as a reporter, in both *E. coli* (white bars) and *Synechocystis* (black bars). Fluorescence per cell was measured as fluorescence divided by absorbance at 595 nm for *E. coli* and 750 nm for *Synechocystis*. Averages and error bars were calculated from three biological replicates. (Paper II).

Post-translational tools

To enable dynamic temporal studies of gene expression, when promoter activities change over time, destabilized fluorescent protein reporters are necessary. To test the activity of the in *E. coli* previously implemented *ssrA*-degradation tags that were used for targeted degradation of GFP by the ClpXP and -AP proteases [88], we designed EYFP-expression constructs with increasingly effective degradation tags. The *ssrA*-degradation system works well within *Synechocystis*, as can be seen for the progressively lower fluorescence per cell in constructs with stronger degradation tags (Figure 14).

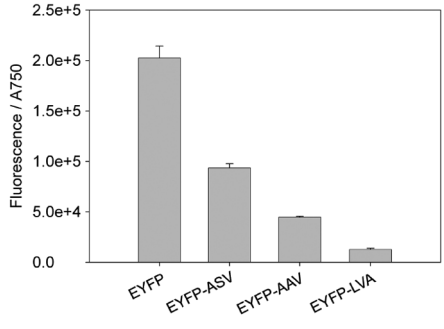

Figure 14. The use of *ssrA*-degradation tags to destabilize a constitutively expressed EYFP reporter in *Synechocystis*. Fluorescence per cell was measured as fluorescence divided by absorbance at 750 nm and error bars correspond to the standard deviation of the average (n=3). (Paper I).

Transcriptional tools (Papers I, III-VI)

Being the first step in the central dogma of molecular biology, transcription gives rise to all the species of the cell encoded by DNA. Thus, regulation of transcription is one of the most important control points in gene expression and crucial for any biotechnological application in a living system.

Native *Synechocystis* promoters (Paper I)

To find suitable promoter candidates for expressing LacI in our first study, the ribonuclease P promoter P*rnpB*, and different variants of the promoter for the large subunit of rubisco, P*rbcL*, were selected. The *rbcL* promoters were divided into two groups; group two that consists of longer promoters that contain an AT-rich sequence and a predicted binding site for the NtcA TF in the upstream part, and group one that lacks the whole upstream part. Further, the three *rbcL*-derived promoters in each group were differentiated more at the 3' end by attaching a RBS at different locations with or without an 8 bp BioBrick scar sequence in between, and the third promoters in each group had a large part of the 3' end truncated. Because the *rbcL* promoters differ in their expected 5'-UTR sequences, it is difficult to draw conclusions by comparing them. Still, it was observed that the presence of the AT-rich upstream sequence lead to an approximately two-fold increase in activity, supporting the hypothesis that this element is an enhancer, potentially for its own promoter located in the upstream region.

P*rnpB* has often been used as a "housekeeping" reference gene because of its stable expression level under different conditions of light and dark or the presence of electron transport inhibitors [89, 90]. Therefore, it was used as a reference promoter in our first study.

Introducing common *Escherichia coli* promoters into *Synechocystis* (Paper I)

Three promoters that are commonly used for different applications in *E. coli*, P*lac*, which equals the *lacZYA* operon promoter, P*tet*, which equals the PL*lac* promoter [76], and P*R*, were tested in *E. coli* and *Synechocystis* using the same GFP reporter cassette. Interestingly, even though P*lac* and P*tet* produce quite a lot of fluorescence per cell in *E. coli*, compared with P*R* which is weaker, neither P*lac* nor P*R* produced any detectable fluorescence in *Synechocystis*, and P*tet* was extremely weak (Figure 15).

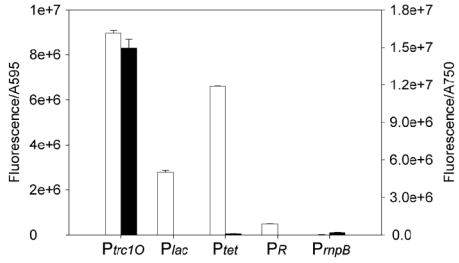

Figure 15. Promoter activities per cell for the strong P*trc* promoter, the weak reference promoter P*rnpB*, and the three test promoters P*lac*, P*tet* and P*R* in both *E. coli* (white bars) and *Synechocystis* (black bars). Fluorescence per cell was measured as GFP fluorescence divided by absorbance at 595 nm for *E. coli* and 750 nm for *Synechocystis*. Averages and standard deviation error bars correspond to three biological replicates. (Paper I).

There are likely different explanations for the lack of activity of these three promoters in *Synechocystis*. P*lac* is far from a consensus promoter, and dependent on activation by CAP in *E. coli* [30]. *E. coli* and *Synechocystis* are far from related, and therefore it is unlikely that a homologue to CAP in *Synechocystis* would bind to and activate transcription from P*lac*.

For P*tet*, recent results have shown that P*tet* is not repressed by a *Synechocystis* TetR homologue [91]. Rather, the weakness of P*tet* in *Synechocystis* is probably due an inefficient core promoter.

The inactivity of P*R* in *Synechocystis* remains unexplained. The simplest explanation is that it also has an inefficient core promoter.

These results show that promoters cannot simply be transferred from one organism to another distantly related organism and be expected to work. There are some exceptions, for example near-consensus σ70 promoters are widely conserved, just like the main housekeeping sigma factor σ70, and hence promoters like P*trc* that are close to consensus can be expected to function in many or most bacteria.

Characterization of a library of artificial, constitutive promoters in *Synechocystis* (Paper III)

Instead of focusing on native promoters, with all the inherited regulation and cross-talk, or *E. coli* promoters that are not close to consensus σ70 promoters, we selected several members from an artificial consensus σ70 promoter library. This library, referred to as the J23-library from its BioBrick part name BBa_J23### [58], consists of minimal, artificial and hence orthogonal promoters that were obtained by successively mutating an *E. coli* σ70 promoter consensus sequence, creating a library of promoters with different strengths. Selected members from this library were characterized in *Synechocystis* using an EYFP reporter cassette, and sorted according to their activities. For comparison, native *Synechocystis* promoters that have been

used previously for engineered expression, P*nirA*, P*petE* and P*rnpB*, were also included (Figure 16).

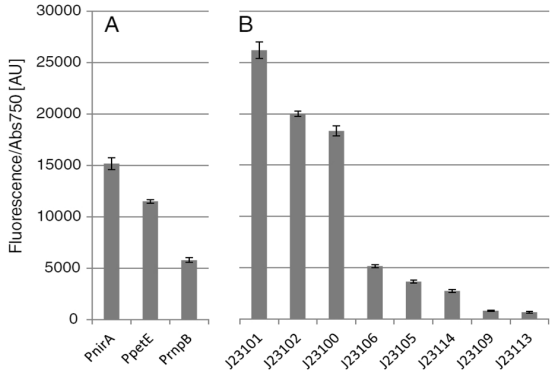

Figure 16. Promoter activities per *Synechocystis* cell. A. The native promoters P*nirA*, P*petE* and P*rnpB*. B. Eight members from a minimal, artificial and constitutive promoter library. J23### corresponds to the BioBrick part BBa_J23### where # is a number. Fluorescence per cell was measured as EYFP fluorescence divided by absorbance at 750 nm. Averages and error bar standard deviations correspond to six biological replicates. (Paper III).

As the J23-library promoters span a wide range of activity they could be used to fine-tune the expression of e.g. orthogonal transcription factors, or enzyme levels for metabolic engineering. As they are minimal they could serve as starting promoters for the engineering of new orthogonal, regulated promoters.

Engineering LacI-regulated transcriptional systems (Papers I, III & IV)

The *lac* repressor is one of the most well-characterized transcription factors and is used widely in different biotechnological applications. Furthermore, it is orthogonal to cyanobacteria, conferring a lower risk for cross-talk. Therefore, in our first study, we introduced a common and strong LacI-regulated promoter, P*trc* [35] (referred to as P*trc*1O), into *Synechocystis* together with a version of it with two *lacO* (P*trc*2O) to enable enhanced repression through DNA-looping. Comparing the repression and induction behavior of P*trc* and P*trc*2O in both *E. coli* and *Synechocystis* overexpressing LacI, however, illustrated obvious differences (Figure 17).

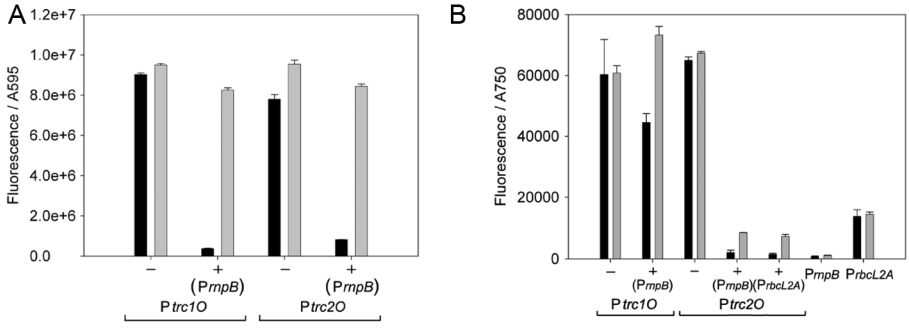

Figure 17. The P*trc* (P*trc*1O) and P*trc*2O version with two *lacO* for improved repression through DNA-looping were characterized using a GFP reporter construct. LacI was overexpressed from the same vector using either P*rnpB* or P*rbcL2A*, a version of P*rbcL* that is about 19 times stronger than P*rnpB*. The LacI-cassette was cloned divergently with the reporter cassette to avoid interference. A. Characterization in *E. coli*. Cellular fluorescence was calculated from fluorescence normalized with absorbance at 595 nm. B. Characterization in *Synechocystis*. Cellular fluorescence was calculated from fluorescence normalized with absorbance at 750 nm. '-', no LacI expressed, '+', LacI expressed from promoter in parenthesis. Black columns non-induced, grey columns induced with 1 mM IPTG for *E. coli* and 2 mM IPTG for *Synechocystis*. Averages and standard deviations are calculated from three biological replicates. (Paper I).

In *E. coli*, both P*trc*1O and 2O are very similar in both repression and induction. However, in *Synechocystis*, P*trc*1O can barely be repressed, whereas P*trc*2O is more repressed, although not completely, but cannot be induced up to unrepressed levels. Further, increased expression of LacI with the about 19 times stronger P*rbcL2A* promoter does not significantly improve repression of P*trc*2O while induction is basically the same. That additional LacI expression does not increase repression tells us that LacI is already at its maximum repression capacity in the cell for that specific promoter. If this is the case, though, it is counter-intuitive that this high level of LacI should not be enough to repress P*trc*1O. Further, that additional LacI expression does not lead to lower induction tells us that IPTG is in large excess. Hence, it seems most probable that the discrepancy between the promoters' behavior in *E. coli* and *Synechocystis* are due to effects on the promoter-DNA or transcriptional level. It also seems reasonable that the limited repression of P*trc*1O is related to another cause than the lack of induction in P*trc*2O. Perhaps the observed differences between *E. coli* and cyanobacterial RNAP causes these differences in repression and induction.

To extract information about the behavior of LacI-regulated repression through DNA-looping and potential differences between *E. coli* and *Synechocystis*, the next phase of investigation was studying LacI-mediated repression through DNA-looping.

Constructing a LacI-expressing strain of *Synechocystis*

To study the effects of LacI repression and DNA-looping, a library of different promoters will be required. To simplify the library construction process, a LacI-expressing strain of *Synechocystis* (Syn-LacI) was constructed and used throughout all LacI-regulated measurements. The Syn-LacI strain was engineered to produce similar levels of LacI as in our first study by using a promoter from the minimal constitutive J23-library, but from a neutral site on the chromosome. As a negative control for LacI, a strain with the same chloramphenicol resistance cassette that was used to make the Syn-LacI strain was also constructed (Syn-CmR).

Investigating LacI-mediated DNA-looping in *Synechocystis*

To study the observed differences of the repression and induction behavior of dual-*lacO* DNA-looping-capable promoters between *E. coli* and *Synechocystis*, a 24-member promoter library was constructed. This library consists of promoter reporter constructs with identical 5'-UTRs, including a proximal *lacOsym*, and P*trc*-derived core promoters; only the distal part of the promoter beyond the -35 element and upstream differs, making it possible to directly compare promoter activities. Further, each consecutive P*trc*2O promoter differs from the previous in that its distal *lacOsym* is one bp further away from the -35 box of the core promoter. For example, the P*trc*2O promoter with the distal *lacOsym* located next to the -35 element is named P*trc*2O-0. This makes it possible to study the effects of repression and induction when the two *lacOsym* are on the same side of the DNA-helix, on opposite sides, or somewhere in between, all around. This is interesting because DNA cannot bend or twist in any way in such short distances, making DNA-looping when the two *lacOsym* are on opposite sides of the helix impossible, leading to decreased repression due to missing cooperativity. On the other hand, when the two *lacOsym* are on the same side of the helix, the DNA bends or loops easier, and cooperative, enhanced repression is possible.

The DNA-looping library displays many similarities between *E. coli* and *Synechocystis*, but also many differences (Figure 18).

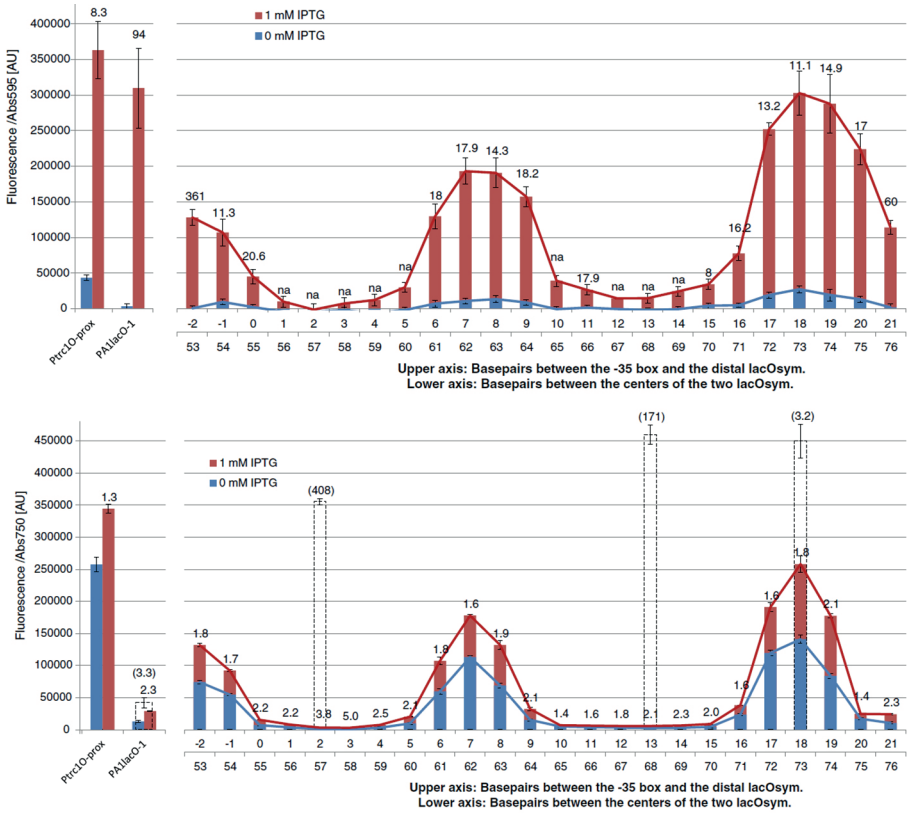

Figure 18. DNA looping library in *E. coli* (upper) and *Synechocystis* (lower) over-expressing LacI. All promoters except PA1*lacO*-1 share the same core promoter and 5'-UTR. Mean values and standard errors of the mean are calculated from six replicates for *E. coli*, four for *Synechocystis* Syn-LacI and two for Syn-CmR. Values on top of columns are induction ratios, values in parenthesis are repression ratios, broken columns represent activity in the Syn-CmR strain. Fluorescence per cell was measured as fluorescence divided by absorbance at 595 nm for *E. coli* and 750 nm for *Synechocystis*. IPTG (isopropyl β-D-1-thiogalactopyranoside). (Paper III).

When comparing the DNA-looping library between *E. coli* and *Synechocystis* it is apparent that: (i) both strains share the same periodicity between peaks and troughs of 11 bp, the in vivo number of bp/turn of chromatin DNA, at the same inter-*lacO* distances, (ii) peaks that represent unfavorable looping geometries are more repressed in *E. coli* than *Synechocystis* but can still be highly induced, whereas (iii) favorable looping geometries in the troughs are better repressed in *Synechocystis*, where repressed promoters cannot be induced like in *E. coli*. Further (iv), the repression troughs in *Synechocystis* are wider and deeper than in *E. coli*.

To conclude, the number of bp per helical repeat is the same in both strains, but there are large differences in repression and induction behavior. The inefficient repression for unfavorable geometries, the peaks, in *Synecho-*

cystis, can be explained by the lack of looping and cooperativity, causing the ineffective proximal *lacOsym* of P*trc*1O-prox to be alone in repressing the promoter. However, it is especially striking that when geometries are favorable for looping, repression is extremely efficient in *Synechocystis* and induction is not possible. If this is explained by extremely high levels of LacI/operator, it is illogical that P*trc*1O-prox or the fluorescence peaks are not more efficiently repressed. Alternative explanations are that cyanobacterial RNAP are more sensitive to DNA-looping or bending than enterobacterial RNAP, or, that different architechtural DNA-binding proteins in *Synechocystis* enhance loop stability, thereby enforcing repression through DNA looping.

Engineering destabilized LacI-regulated promoters for *Synechocystis*

The highly repressed P*trc*2O-2 promoter (Figure 18) would work as an excellent promoter if it only could be induced to higher levels. To make this possible, we developed three variants of P*trc*2O-2 where we successively reduced the strength of the distal *lacO* from *lacOsym* in P*trc*2O-2 to *lacO1* in P*trc*2O-2 dO1, *lacO2* in P*trc*2O-2 dO2, and *lacO3* in P*trc*2O-2 dO3. We then transferred the three versions into the Syn-LacI and the Syn-CmR strains for characterization. All the promoters were of similar, high strength in the absence of LacI, but were repressed down to different levels depending on the quality of the distal *lacO* (Figure 19A, left). The most efficiently repressed version was P*trc*2O-2 dO1, but it could however not be induced to practical levels (Figure 19A, right). P*trc*2O-2 dO3, on the other hand, was extremely leaky and henceforth not of use for any practical purpose. Instead, P*trc*2O-2 dO2 (dO2) could be repressed quite efficiently but still induced to significantly higher levels. To characterize dO2 in detail, we performed a time-development study of fluorescence per cell (Figure 19B, left). The initially high values of cellular fluorescence, probably from fluorescent protein accumulated during a slower previous growth phase, decreased down to levels under the detection level after 2 days under no induction. However, the induced cultures started to accumulate fluorescence immediately, leading to induction ratios from 5 to 40 and higher, as the repressed fluorescence was not detectable. This is a clear improvement as compared to previous versions of P*trc* in *Synechocystis*. Further, we characterized the induction transfer curve of dO2 and fit a Hill equation to the data to extract information about the cooperativity of the system (Figure 19B, right). The model fit the data very well, but the measured cooperativity was limited to a Hill coefficient of 1.171, which is low compared to the wild-type LacI system in *E. coli* that has a cooperativity of 2.

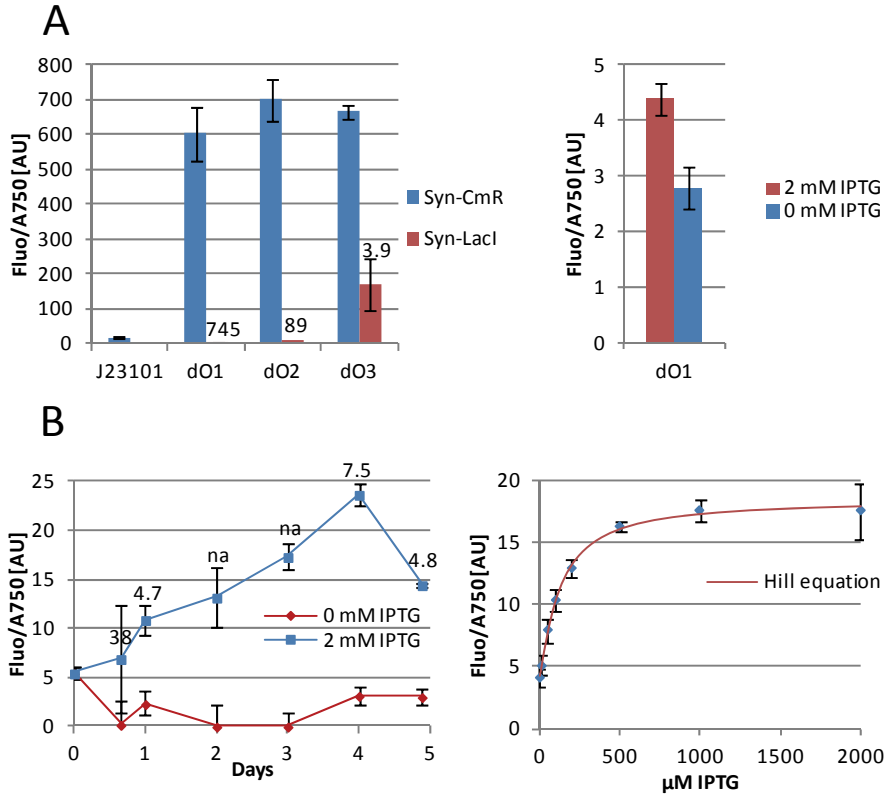

Figure 19. Characterization of the P*trc*2O-2 variants with different distal *lacO*. A. Left: Promoter activity in the absence of LacI (Syn-CmR) and in the presence of LacI (Syn-LacI). Values above columns denote repression ratios. Right: Induction of dO1 in Syn-LacI. B. Left: Time development of dO2 fluorescence per Syn-LacI cell with (2 mM) or without (0 mM) IPTG induction. Right: Induction transfer curve of dO2 in Syn-LacI. A Hill equation was fit to the data. Fluorescence per cell was measured as fluorescence divided by absorbance at 750 nm. Averages and error bar standard deviations correspond to two biological replicates. IPTG (isopropyl β-D-1-thiogalactopyranoside). dO1 (P*trc*2O-2 dO1), dO2 (P*trc*2O-2 dO2), dO3 (P*trc*2O-2 dO3). (Paper IV).

An orthogonal transcriptional system based on T7 RNA polymerase

Using P*trc*2O-2 dO1 and dO2, we engineered two genetic circuits for controlling the expression of T7 RNAP, which in turn transcribed the *eyfp* reporter. Using the dO1 construct, we could not detect any significant fluorescence, which implies that dO1 is completely silent. We could confirm fluorescence expression from the T7 promoter when dO2 was used to control the expression of T7 RNAP, but there were potentially deleterious substitutions in the T7 RNAP coding sequence, most likely limiting its transcriptional efficiency. Still, the use of orthogonal LacI-regulated promoters to control the expression and activity of an orthogonal RNAP is a great step towards

implementing decoupled transcriptional systems that will help in engineering the next generation metabolic systems for cyanobacteria.

Development of Gal4-regulated transcriptional systems for light-regulation (Papers V & VI)

Light-regulated control of gene expression opens up new possibilities for spatiotemporal control of transcription in biotechnology and biomedicine. Such orthogonal systems would also be useful tools for synchronizing the activity of genetic and metabolic circuits with day/night cycles in photosynthetic organisms such as cyanobacteria or plants.

In general, light-regulated systems for in vivo control of bacterial gene expression consist of at least two parts: a sensor domain that may be sensitive to specific wavelengths of light, and an effector domain that carries out a function when activated by the sensor domain. Most engineered light-regulated systems for gene expression belong to either one of two categories: light sensitive two-component histidine kinases or light-sensitive DNA-binding dimers (Paper V) (Figure 6). The two-component systems are large in size, both regarding their coding sequences and regarding their protein components. They consist of a sensor part that detects the light signal that is then transferred via allosteric changes to the kinase part. The kinase part then auto-phosphorylates and subsequently transfers the phosphor-group to a smaller response regulator. The activated response regulator then gains the ability to bind its specific operator sequence at a promoter, where it either acts as an activator or a repressor. Light-regulated dimers, on the other hand, can be regarded as one-component systems, since they gain activity directly when they dimerize. Hence, they are generally much smaller than histdine kinase two-component systems, which is sometimes an advantage in itself. Also, since there are fewer steps from sensing the signal to activation of the gene regulatory effect for one-component light sensors, theoretically they can be faster responders than two-component systems (Paper V). Therefore, one-component light-regulated transcription factors are the focus of this thesis.

To enable the construction and implementation of a Gal4-based light sensitive transcription factor, we first set out to optimize Gal4-repressed promoters (Figure 20). Also, such promoters would be a useful addition to the toolbox of orthogonal transcription factors and promoters for bacteria.

J23101	TTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGC
P*uasg*-paul	AACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCG GCTCGTATGTTGTGTGCTCTAGCGGAAGACTCTCCTCCGTCTAGAG**AGGA**AACAGCT
P*uasg*-c	GAGCTGTTGACACGGAGGACTGTCCTCCGTATAATGTGTGG***A***ATCACACA
P*uasg*-p1	GAGCTGTTGACAATTAATCATCCGGCTCGTATAATCGGAGG***A***CTGTCCTCCG
P*uasg*-p2	GAGCTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGG***A***ATCGGAGGACTGTCCTCCG
P*uasg*-cp1*	GAGCTGTTGACACGGAGGACTGTCCTCCGTATAATCGGAGG***A***CTGTCCTCCG
P*uasg*-cp2*	GAGCTGTTGACACGGAGGACTGTCCTCCGTATAATGTGTGG***A***ATCGGAGGACTGTCCTCCG

Figure 20. Engineered Gal4-regulated promoters for bacteria. J23101 refers to BBa_J23101 [58] and is a constitutive control promoter. P*uasg*-paul is a previously developed promoter [92] included for comparison. -35 and -10 elements are marked in red, operators are underlined, the palindromic edge of the UASG operator is marked in blue. * Unstable constructs due to recombination between UASG-repeats. (Paper VI).

We designed five promoters based on the previous observation that the core of a promoter is the most efficient location for a repressor to hinder transcription, closely followed by the proximal region [29], and consequently inserted the 17 bp UASG operator into the core or two proximal locations of the P*trc* promoter. To enable more efficient repression due to dual binding of Gal4 to the promoter, and potentially enabling cooperative effects, we also combined the core-located UASG with the first and second distal positions. However, these promoters were unstable due to recombination between the two UASG operators, resulting in loss of function.

To characterize Gal4-repression of the promoters we cloned them with a GFP cassette and measured the fluorescence per cell in the absence of Gal4 and the presence of Gal4 co-expressed from another plasmid in the same cell. The constitutive control J23101 remained unaffected by Gal4 expression, whereas all the UASG-containing promoters were affected (Figure 21).

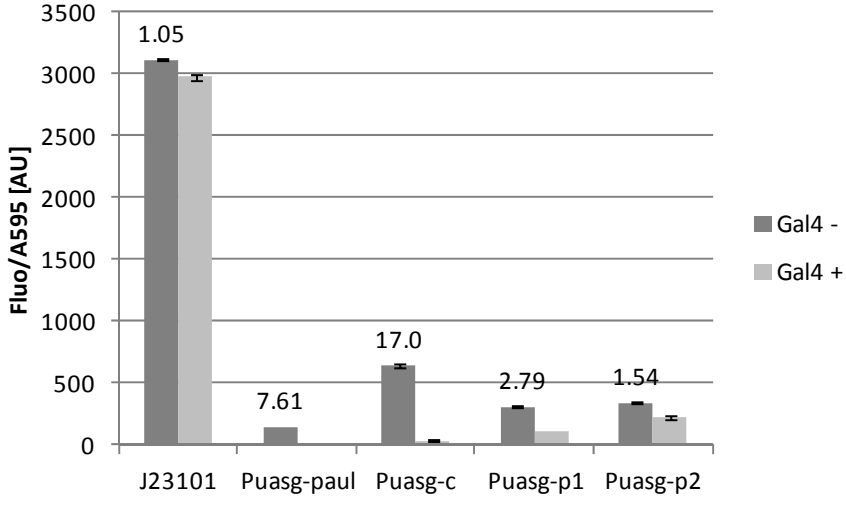

Figure 21. Promoter activity in the absence (Gal4 -) or presence (Gal4 +) of Gal4. J23101 refers to BBa_J23101 [58] and is a constitutive control promoter. P*uasg*-paul is a previously developed promoter [92] included for comparison. Fluorescence per cell was measured as fluorescence divided by absorbance at 595 nm. Averages and standard deviations are calculated from three biological replicates. (Paper VI).

P*uasg*-c, the promoter with the UASG operator in the core location, was the most efficiently repressed promoter and a clear improvement over the previously developed P*uasg*-paul promoter [92].

Recently, the chimeric GAVPO mammalian transcriptional activator was developed [93]. GAVPO consists of three domains: The Gal4 DNA-binding domain, a light-sensitive LOV-domain from the fungal protein VVD, and a mammalian activation domain. Upon blue-light stimulation, it forms a dimer through LOV-domain dimerization and binds to the UASG site close to a mammalian promoter that it subsequently activates.

Using an analogous design of fusing the Gal4 DNA-binding domain and the VVD LOV-domain we created GAV. Further, we hypothesize that GAV that binds the core UASG of P*uasg*-c will function as a potentially light-regulated repressor of transcription in bacteria. To test this, we made four versions of GAV with different mutations to the blue-light sensitive LOV-domain: One mutation causing it to become a permanent dimer (GAV-D), one that makes it a permanent monomer (GAV-M), one with the same mutations as GAVPO (GAV-O), and finally one version with the wild-type VVD sequence (GAV-W). To characterize the activity of the GAV-variants, we cloned it divergently under a LacI-repressed promoter together with the P*uasg*-c test promoter driving expression of a GFP reporter (Figure 22).

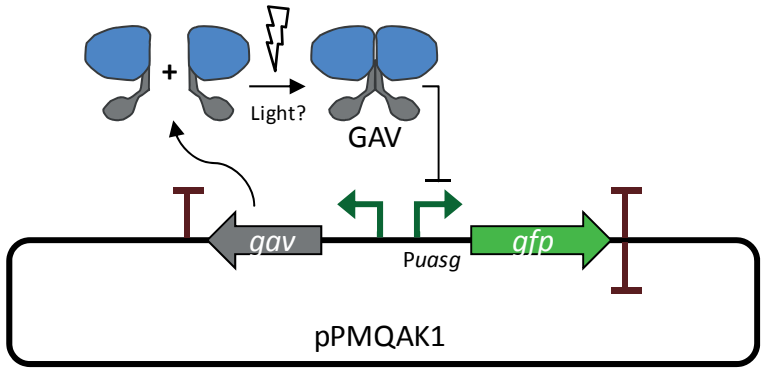

Figure 22. Design of GAV-variant expression and P*uasg* test promoter constructs. The GAV cassette and the test promoter reporter are cloned divergently to avoid promoter interference. (Paper VI).

We then characterized the effect of the different GAV-variants, with or without extra IPTG induction of GAV expression, by measuring the effect on the fluorescence per cell from the test promoter constructs while using J23101 as a constitutive control (Figure 23).

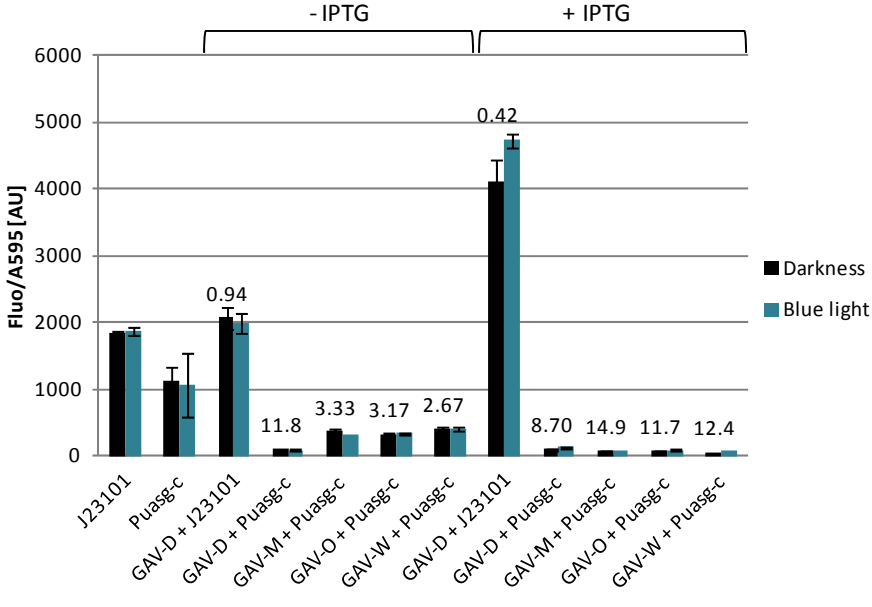

Figure 23. Characterization of GAV-repression under blue light and darkness. J23101 refers to BBa_J23101 [58] and is a constitutive control promoter. J23101 or P*uasg*-c refers to only the test-promoter reporter construct, GAV-D + P*uasg*-c refers to the combination of GAV-D and the test promoter. Values above columns refer to the repression ratios. Fluorescence per cell was measured as fluorescence divided by absorbance at 595 nm. Averages and standard deviations are calculated from three biological replicates. (Paper VI).

First, an increase in the constitutive control promoter J23101 can be observed. As J23101 is not a target for Gal4 repression (Figure 21) or even Gal4-binding, it is unlikely that GAV would function as an activator for J23101. Instead, the increase in fluorescence from J23101 while co-expressed with GAV-D, and the further increase upon induction of GAV-D, can be explained by a severely decreased growth rate, which leads to an accumulation of the stable GFP. The decrease in growth rates for the P*uasg*-c + GAV-variant test cultures were less severe than for J23101 + GAV-D, but there is a small decrease in growth upon induction that likely leads to an underestimation of GAV-repression.

Looking at repression of P*uasg*-c without induction, the permanent dimer version of GAV, GAV-D, is the most efficient. This was also expected, as GAV needs to dimerize for the Gal4 DNA-binding domain to bind its operator. Still, there is a significant level of repression from the other versions of GAV, including the monomer version where the contribution to dimer-stability from the light-regulated LOV-domain should be zero. This means that the Gal4 DNA binding domain by itself has the capacity to form dimers and bind its operator, repressing the promoter independently on the presence of blue light or darkness. This may also be why no light-regulated effects on repression can be seen, as GAV forms dimers even in darkness. If this is the cause for the lack of light-regulated repression, it could be solved by (i) truncating the Gal4 domain to destablilize the dimer-interface that forms even in darkness, and/or (ii) by lowering the expression levels of GAV, lowering the effective concentration and hence the driving force to form dimers in darkness.

Conclusions & Future Perspectives

Tools developed in the present study and needs for future development

Herein was presented the necessary theory for developing and testing new transcriptional tools. The developed tools for engineering cyanobacteria included the following non-transcriptional parts:

- The broad-host-range BioBrick vector pPMQAK1.
- The characterization of fluorescent proteins in a strong background of photosynthetic pigments and the confirmation of their use as reporters.
- Ribosome binding sites for more efficient initiation of translation in *Synechocystis*.
- Protein degradation tags that destabilize proteins to enable dynamic temporal studies, and that decreases the dependence of protein accumulation on growth rate.

A range of transcriptional tools and knowledge were developed:

- The *Synechocystis rnpB* promoter that serves as a stable, native reference promoter.
- A library of minimal and orthogonal constitutive promoters was characterized and it could be used for direct applications, as orthogonal reference promoters, or as basis for the development of novel regulated, orthogonal promoters.
- One completely and one strongly LacI-repressed promoter that could be used for applications requiring stringent repression but not requiring induction such as genetic switches.
- An inducible and repressed LacI-regulated promoter that could be used for direct applications or improved further according to my suggestions.
- An orthogonal transcriptional system based on T7 RNAP was shown to function and additional work to make it more practically useful was suggested.
- The optimization of Gal4-repressed promoters for use with Gal4 as an orthogonal repressor in bacteria.
- Work towards implementing a light-regulated one-component repressor based on Gal4 and a blue-light sensitive LOV-domain.

Towards portable orthogonal gene expression systems

From the cloning and sequencing of single genes to the synthesis of genomes and the sequencing of populations, our ability to read and write the language of life develops further every month. The first synthetic bacterial genome has already been booted up to life in a cellular chassis [94] and the first designer yeast chromosome was recently synthesized and inserted functionally into yeast [95]. As our ability to synthesize and design larger pieces of DNA and genomes develops, there is going to be a need to decouple genetic systems from each other. For instance, for running a parallel metabolism only activated at certain environmental cues. These kind of parallel transcriptional systems are already in use, they are orthogonal RNA polymerases like T7 RNAP [49]. Instead of transcribing only one or a few genes involved in forming a certain product or metabolic flux, orthogonal polymerases like T7 could be used to run hundreds or a whole genome's worth of genes simultaneously and decoupled from the host's native system. Further, the unification of orthogonal polymerases with orthogonal ribosomes, or o-ribosomes [96], enables a whole orthogonal gene expression system. Such systems are not only interesting because they allow us to decouple our system from the host's native system, but also because they allow us to incorporate for instance artificial amino acids [97] that open up new biochemistry. Still, though, having robust orthogonal systems would yield possible benefits already now. Biological circuits could be constructed using orthogonal systems in one strain, and then ported to other strains for e.g. functional screening without having to switch promoters and other host-specific parts to make the system work.

Svensk sammanfattning

Energi är något som vi människor alltid använt, och med utvecklingen av vår civilisation har tendensen varit att energiåtgången per människa ökat. Det vore inget problem om energin vi använder kom från en källa som inte kommer att sina, eller vars restprodukter inte ger negativa effekter på vår omgivning. Men de fossila bränslen som vi idag använder är en ändlig resurs. Det innebär inte att de plötsligt kommer att ta slut, men utvinningen av dem kommer troligen att bli dyrare och tillgången kommer långsamt att minska. Dessutom leder förbränningen av fossila bränslen till stora utsläpp, bland annat av växthusgasen koldioxid. Alldeles nyligen kom FN:s klimatpanel (IPCC) ut med en delrapport till den stora klimatförändringsrapporten 2013 (Climate Change 2013) som beskriver den naturvetenskapliga grunden till klimatförändringarna. I den fastslår de forskare som undersökt de vetenskapliga bevisen att det är extremt sannolikt att människan är den dominerande orsaken till de pågående klimatförändringarna. De fastslår även att utsläppen av koldioxid är en av de största fysiska drivkrafterna bakom klimatförändringarna. Jämfört med den föregående rapporten, som gavs ut 2007, har bevisen för människans påverkan stärkts, och det råder nu alltså i princip inget tvivel om att människan är huvudorsaken till förändringarna. Med andra ord har vi ett stort behov av att identifiera alternativ till användningen av fossila bränslen som energikälla.

Solen belyser konstant vår planet med enorma mängder energi. Varje timme ger solen oss ungefär samma mängd energi som mänskligheten använder under ett år. Därför är solen en i stort sett outsinlig energikälla, och skulle kunna lösa våra energiproblem om vi bara kunde fånga och lagra solenergin på ett effektivt sätt. I princip kan solenergin fångas och konverteras till elektricitet, som är svårt att lagra på en global skala, eller konverteras till kemisk energi i ett bränsle, som är lätt att lagra.

Cyanobakterier, även kända som blågröna alger, är fotosyntetiska bakterier som med början för ca 2.8 miljarder år sedan var med och bildade den syrerika atmosfär vi har idag. Cyanobakterier finns utspridda över hela jorden, i världshaven, i sjöar, varma källor och frusna vidder. Idag finns det ett stort intresse av att använda cyanobakterier bioteknologiskt, framförallt för produktion av biobränslen. Anledningarna är flera: först och främst fångar de solenergidrivna cyanobakterierna koldioxid genom fotosyntesen och kräver därmed inte att man tillsätter fixerat kol, som exempelvis socker vid tillverkningen av bioetanol, utan de tar aktivt bort koldioxid från luften. Dessutom

trivs många stammar i saltvatten, vilket innebär att de kan odlas utan konkurrens med jordbruk eller annan mänsklig aktivitet som kräver sötvatten. Vidare har vi DNA-sekvenserna till ett flertal olika cyanobakteriers arvsmassa vilket underlättar möjligheten att biotekniskt förändra dem. Med hjälp av det nya fältet syntetisk biologi har vi ett tillvägagångssätt för att programmera om cyanobakterierna till bl.a. gröna biobränslefabriker.

Syntetisk biologi är en biologisk ingenjörskonst där man ser levande system som maskiner vars egenskaper och beteenden är programmerbara. Målet med syntetisk biologi är att kunna designa nya, biologiska enheter med utgångspunkt från små funktionella delar, och sedan kunna kombinera dessa enheter med andra delar för att skapa nya biologiska system. Tillämpningarna av syntetisk biologi är många förutom bioteknik och biobränsleproduktion. Till exempel kan det användas för att programmera organismer till att detektera och neutralisera miljöföroreningar, eller för att göra om små bakterier till målsökande förgörare av cancerceller. Men för att kunna bygga robusta, användbara biologiska system krävs byggstenar, till exempel pålitliga och karaktäriserade promotorer som driver uttrycket av gener. En annan viktig egenskap hos biologiska byggstenar är att de ska kunna användas i kombination med andra delar i levande organismer utan att påverkas i sin funktion. En kategori av delar som uppfyller detta kriterium är ortogonala delar. Ortogonala delar kommer antingen från en helt annan organism än den cell där de ska tillämpas, eller är helt eller delvis artificiella. Detta gör att de inte har nedärvda interaktioner med andra delar i systemet eller organismen som kan göra det slutliga systemet instabilt eller dysfunktionellt.

Denna avhandling fokuserar på att ta fram och karaktärisera nya byggstenar, speciellt promotorer och andra delar som är viktiga för att kontrollera genuttryck, som kan användas till att programmera om cyanobakterier till att bli effektiva, gröna biobränslefabriker.

Den första byggstenen som producerades är en plasmid, ett stycke cirkulärt DNA som används för att föra över gener, vid namn pPMQAK1. Denna plasmid har egenskapen att kunna föras över till en mängd olika cyanobakterier vilket möjliggör tester av delar och hela biobränslesystem i flera intressanta stammar. Dessutom togs ett syntetiskt ribosomalt inbindningsställe (RBS) fram, som testades tillsammans med andra syntetiska RBS i sin förmåga att driva translationen av mRNA till färdiga proteiner i cyanobakterien *Synechocystis* PCC 6803 (*Synechocystis*). Även en sorts markörer för att aktivera proteindegradering implementerades och karaktäriserades i *Synechocystis*.

För att kontrollera genuttryck designades en mängd olika promotorer och andra transkriptionella element. Ett bibliotek av konstant aktiva, artificiella, minimala och därmed ortogonala promotorer karaktäriserades i *Synechocystis*. Dessa promotorer visade sig producera ett brett spektrum av genuttrycksnivåer och kan användas för att styra uttrycket av olika gener för olika applikationer eller användas som grund för designen av nya reglerade pro-

motorer. Ett transkriptionellt system som bygger på transkriptionsfaktorn LacI, ett protein som blockerar genuttryck och därför benämns repressor, konstruerades. Två promotorer som kontrolleras av LacI konstruerades med antingen ett (P*trc*1O) eller två (P*trc*2O) inbindningsställen för LacI. De visade sig ha olika regleringsegenskaper när de kontrolleras av LacI i bakterien *Escherichia coli* (*E. coli*) eller cyanobakterien *Synechocystis*. P*trc*1O kunde inte blockeras i samma utsträckning i *Synechocystis* som i *E. coli*, medan P*trc*2O kunde blockeras i *Synechocystis* men inte induceras (avblockeras). För att undersöka orsakerna till dessa skillnader utvecklades ett bibliotek av P*trc*2O- promotorer där avståndet på DNA:t mellan inbindningsställena för LacI varierades. Karaktäriseringen av biblioteket i både *E. coli* och *Synechocystis* visade på flera likheter, men även flera olikheter. Resultaten tyder på att LacI fungerar som en ortogonal repressor i *Synechocystis*, men att blockeringen av transkription i promotorer som har två LacI-inbindningsställen på samma sida av DNA-spiralen inte alls kan avblockeras i *Synechocystis.* Promotorer som har två LacI-inbindningsställen på motsatta sidor kan, å andra sidan, endast blockeras mycket bristfälligt. Detta skiljer sig från situationen i *E. coli* där promotorerna med inbindningsställen på samma sida både kan blockeras men även induceras något, och där promotorerna med inbindningsställen på motsatta sidor både kan delvis blockeras och induceras. Anledningarna till dessa skillnader är inte klargjorda, men orsakerna har förmodligen att göra med skillnader på strukturell nivå i promotor-DNA:t eller skillnader i RNA-polymerasen mellan *E. coli* och *Synechocystis*. En promotor från biblioteket som är starkt blockerad men inte kan induceras i *Synechocystis* kan användas för att tillverka en genetisk switch. En variant av denna promotor som både kan blockeras och induceras konstruerades och karaktäriserades i *Synechocystis*. Dessutom användes denna förbättrade promotor för att driva uttrycket av det ortogonala T7 RNA polymeraset som karaktäriserades i sin förmåga att transkribera från den ortogonala T7-promotorn i *Synechocystis*.

För att öka antalet ortogonala transkriptionsfaktorer som finns tillgängliga för syntetisk biologi i bakterier konstruerades och karaktäriserades även ett genetiskt system som bygger på aktivatorn Gal4 från jäst. Därtill togs flera steg för att implementera en ortogonal ljusreglerad repressor som bygger på Gal4.

Slutligen bidrar denna avhandling även med en teoretisk och praktisk bakgrund för design av transkriptionella system och ger förslag på hur dessa kan utvecklas ytterligare. Min förhoppning är att denna kunskap och dessa biologiska byggstenar kan bidra till designen av effektivare produktionssystem för biobränslen i cyanobakterier, som ett förnyelsebart alternativ till bruket av fossila bränslen.

Acknowledgments

First, I would like to express my heartfelt gratitude to Peter Lindblad, my main supervisor, for the guidance and support over all these very interesting years as a PhD student in your group. You have given me the freedom to develop and test my research ideas while providing much appreciated guidance, support and humor. For this I offer you my warmest thanks.

I would also like to extend my warmest thanks to my co-supervisor Thorsten Heidorn for all the great advice, guidance and good times. Both in the lab but also while on conferences or winning sailing competitions in lake Mälaren. It should be added for the benefit of other readers that Thorsten is an excellent sailor while I am merely a good swimmer.

My second co-supervisor Alfonso Jaramillo has been a great inspiration, full of visionary ideas and interesting projects. It has been a great pleasure to work together with you and for this I offer you my sincere gratitude.

Karin Stensjö, it has been a true pleasure to work together with you. Also, you have always been a great source of advice and support, in good times and in testing times, and for this you have my deepest gratitude.

Stenbjörn Styring, you have been a great source of inspiration and support, and for this I give you my most sincere thanks. Your lectures about the state of global energy and solar fuels have always given motivation and a strong sense of purpose that I consider very important.

Leif Hammarström, I truly appreciate your warm but direct leadership style, always clear and with a smile. Thank you for good times and taking so good care of the department.

To my colleagues and friends within the group and the department: Working with you has been truly great. You make a good day even better and take the thorns out of more challenging days. To my closest colleagues without any order of significance: Susanne, you have been a great office-mate during my final time as a PhD student, brightening things up with your amazing humor. Fikret, you are likewise always there with a smile and a joke. We should develop our plans for putting a giant laser on the moon. Fredrik, Johannes, Felix, Patricia, Anders, Ping and Ann for being such impressive experts on different aspects of photosynthesis and biochemistry, always ready to answer questions or a have good laugh. Nigar, for lots of good times and sharing some of your amazing Azeri food. Pia, Gustaf, Åsa and Helder for always being open to a good discussion or just having fun. Namita for being an awesome collaborator, Bagmi for having such a great and positive

philosophy of life, and both of you for adding so much joy and color by sharing your culture. To Elias, Claudia, Rui, Xin, Feiyan, Olegs, Mihir, Kostas, Christoph, Giovanny, Starla, Sonja, Edgar, Adam and Ken, you all are such awesome people. Some special things worth mentioning: Elias, thanks for a great version of Sinatra's "My way" and for enlightening perspectives. Claudia, for having such a fun and relaxed attitude to life. Rui for being such a kind person. Xin for Tai Chi assistance. Feiyan for all the kung fu moves in the corridor. Olegs for showing another perspective in foreign policy. Mihir for your happy and forward ways. Kostas for good ideas and humor. Christoph for your open, supportive and optimistic personality.

To previous colleagues and friends for many good times in the past and lots of great times to come. Tanai, you have been a true inspiration and great friend. I believe C&C is on its way and will surpass our ambitions. Hsin-Ho for being a great long-time colleague and friend, I hope we can work together again in the future. Amr for interesting perspectives and being such a positive person with a great drive. Ievgen and Katha for good times and good laughs. Åsa, Marie, Ellenor, Paulo and Fernando, for making my first time in the group so good, and thank you for all good advice. For my old time "Ludvika" people: Tobias, for all the good times and always being there as a great friend and brother. Stefan, Tias, Magnus and Peter for the friendship and adventures, in this world and other ones. To Gabriel, Björn and Silver: For merry times, craziness and great friendship - Glory to the Brave! To Adrian and Hamid: Even though it has been a risky business knowing you, with shoot-outs and drive-by's, thank you for a great time. Femman femman, see you at Condeco! To Erik, Christian, Niclas and Mikael, a.k.a. the engineers that meets and eats sushi, thank you for still being there for a great time. To Marlen for all the fun and risky adventures by canoe. To Marcus for great times teaching how to purify the most amazing protein in the world, "GFP – it's dynamite!" To Markus and Holger for introducing me to the best of German culture, brause pulver, and great times. To Martin and Linda for being good old friends and for all the Ludvika memories. To the T-shirts, Rudbeckers and supporters of the team: Angie, Brinkster, Ammar, Anna, Daniel, Jelena, Jess, Martin, Ben, Todd, Travis, Ola, Claire, Caro, Johannes, Aniñha, Kasia, Sahar, Iwa, Sara K, Sara DO, Mattias, Nicky, Mi, Anna G, Christina, Anna K, David, Yuliya and Lucie. We have gone through too much to mention, from adventures in beautiful Colombia to enjoying life in Uppsala. Thank you all for being great friends!

Finally I would like to thank my family. My brother Johan for always being a great brother and all the good times. My brothers Anders and Patrik, my aunt Gärd and my father Hans for inspiration and support. Marianne, Lars-Göran, Simon and Erik for many happy occasions. LG and my late grandparents Ragnhild and Bror for a good life growing up. Special thanks to my mother Birgitta for all the memories and raising me to be curious, making this thesis possible. My warmest thanks to you all!

References

1. Stocker TF, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.): **Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.** Cambridge, United Kingdom and New York, USA: Cambridge University Press; 2013.
2. Crabtree GW, Lewis NS: **Solar energy conversion.** *Physics Today* 2007, **60:**37-42.
3. Olson JM: **Photosynthesis in the Archean era.** *Photosynth Res* 2006, **88:**109-117.
4. Heidorn T, Camsund D, Huang HH, Lindberg P, Oliveira P, Stensjö K, Lindblad P: **Synthetic biology in cyanobacteria engineering and analyzing novel functions.** *Methods Enzymol* 2011, **497:**539-579.
5. Bothe H, Schmitz O, Yates MG, Newton WE: **Nitrogen fixation and hydrogen metabolism in cyanobacteria.** *Microbiol Mol Biol Rev* 2010, **74:**529-551.
6. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, et al: **Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.** *DNA Res* 1996, **3:**109-136.
7. Knoop H, Grundel M, Zilliges Y, Lehmann R, Hoffmann S, Lockau W, Steuer R: **Flux balance analysis of cyanobacterial metabolism: the metabolic network of *Synechocystis* sp. PCC 6803.** *PLoS Comput Biol* 2013, **9:**e1003081.
8. Kopf M, Klahn S, Pade N, Weingartner C, Hagemann M, Voss B, Hess WR: **Comparative Genome Analysis of the Closely Related *Synechocystis* Strains PCC 6714 and PCC 6803.** *DNA Res* 2014.
9. Endy D: **Foundations for engineering biology.** *Nature* 2005, **438:**449-453.
10. Andrianantoandro E, Basu S, Karig DK, Weiss R: **Synthetic biology: new engineering rules for an emerging discipline.** *Mol Syst Biol* 2006, **2:**2006 0028.
11. Kitney R, Freemont P: **Synthetic biology - the state of play.** *FEBS Lett* 2012, **586:**2029-2036.

12. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al: **Production of the antimalarial drug precursor artemisinic acid in engineered yeast.** *Nature* 2006, **440:**940-943.
13. Anderson JC, Clarke EJ, Arkin AP, Voigt CA: **Environmentally controlled invasion of cancer cells by engineered bacteria.** *J Mol Biol* 2006, **355:**619-627.
14. Young E, Alper H: **Synthetic biology: tools to design, build, and optimize cellular processes.** *J Biomed Biotechnol* 2010, **2010:**130781.
15. Cardinale S, Arkin AP: **Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems.** *Biotechnol J* 2012, **7:**856-866.
16. Rao CV: **Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression.** *Curr Opin Biotechnol* 2012, **23:**689-694.
17. Keasling JD: **Synthetic biology and the development of tools for metabolic engineering.** *Metab Eng* 2012, **14:**189-195.
18. Saecker RM, Record MT, Jr., Dehaseth PL: **Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis.** *J Mol Biol* 2011, **412:**754-771.
19. Schneider GJ, Haselkorn R: **RNA-Polymerase Subunit Homology among Cyanobacteria, Other Eubacteria, and Archaebacteria.** *J Bacteriol* 1988, **170:**4136-4140.
20. Xie WQ, Jäger K, Potts M: **Cyanobacterial RNA-Polymerase Genes *rpoc1* and *rpoc2* Correspond to *rpoc* of *Escherichia coli.*** *J Bacteriol* 1989, **171:**1967-1973.
21. Schyns G, Jia L, Coursin T, Tandeau de Marsac N, Houmard J: **Promoter recognition by a cyanobacterial RNA polymerase: in vitro studies with the *Calothrix* sp. PCC 7601 transcriptional factors RcaA and RcaD.** *Plant Mol Biol* 1998, **36:**649-659.
22. Imashimizu M, Fujiwara S, Tanigawa R, Tanaka K, Hirokawa T, Nakajima Y, Higo J, Tsuzuki M: **Thymine at -5 is crucial for *cpc* promoter activity of *Synechocystis* sp. strain PCC 6714.** *J Bacteriol* 2003, **185:**6477-6480.
23. Imashimizu M, Tanaka K, Shimamoto N: **Comparative Study of Cyanobacterial and *E. coli* RNA Polymerases: Misincorporation, Abortive Transcription, and Dependence on Divalent Cations.** *Genet Res Int* 2011, **2011:**572689.
24. Seshasayee AS, Sivaraman K, Luscombe NM: **An overview of prokaryotic transcription factors : a summary of function and occurrence in bacterial genomes.** *Subcell Biochem* 2011, **52:**7-23.
25. Khudyakov IY, Golden JW: **Identification and inactivation of three group 2 sigma factor genes in *Anabaena* sp strain PCC 7120.** *J Bacteriol* 2001, **183:**6667-6675.

26. Fujisawa T, Narikawa R, Okamoto S, Ehira S, Yoshimura H, Suzuki I, Masuda T, Mochimaru M, Takaichi S, Awai K, et al: **Genomic Structure of an Economically Important Cyanobacterium, *Arthrospira* (*Spirulina*) *platensis* NIES-39.** *DNA Res* 2010, **17:**85-103.
27. Imamura S, Asayama M: **Sigma factors for cyanobacterial transcription.** *Gene Regul Syst Bio* 2009, **3:**65-87.
28. Dong GG, Kim YI, Golden SS: **Simplicity and complexity in the cyanobacterial circadian clock mechanism.** *Curr Opin Genet Dev* 2010, **20:**619-625.
29. Cox RS, 3rd, Surette MG, Elowitz MB: **Programming gene expression with combinatorial promoters.** *Mol Syst Biol* 2007, **3:**145.
30. Lewis M: **The *lac* repressor.** *C R Biol* 2005, **328:**521-548.
31. Gilbert W, Müller-Hill B: **Isolation of the *lac* repressor.** *Proc Natl Acad Sci U S A* 1966, **56:**1891-1898.
32. Krämer H, Niemoller M, Amouyal M, Revet B, von Wilcken-Bergmann B, Muller-Hill B: ***lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators.** *EMBO J* 1987, **6:**1481-1491.
33. Oehler S, Eismann ER, Kramer H, Müller-Hill B: **The three operators of the *lac* operon cooperate in repression.** *EMBO J* 1990, **9:**973-979.
34. Oehler S, Amouyal M, Kolkhof P, von Wilcken-Bergmann B, Müller-Hill B: **Quality and position of the three *lac* operators of *E. coli* define efficiency of repression.** *EMBO J* 1994, **13:**3348-3355.
35. Brosius J, Erfle M, Storella J: **Spacing of the -10 and -35 Regions in the Tac Promoter - Effect on Its In vivo Activity.** *J Biol Chem* 1985, **260:**3539-3541.
36. Bond LM, Peters JP, Becker NA, Kahn JD, Maher LJ: **Gene repression by minimal *lac* loops in vivo.** *Nucleic Acids Res* 2010, **38:**8072-8082.
37. Daber R, Sharp K, Lewis M: **One Is Not Enough.** *J Mol Biol* 2009, **392:**1133-1144.
38. Oehler S, Alberti S, Müller-Hill B: **Induction of the *lac* promoter in the absence of DNA loops and the stoichiometry of induction.** *Nucleic Acids Res* 2006, **34:**606-612.
39. Traven A, Jelicic B, Sopta M: **Yeast Gal4: a transcriptional paradigm revisited.** *EMBO Rep* 2006, **7:**496-499.
40. Hong M, Fitzgerald MX, Harper S, Luo C, Speicher DW, Marmorstein R: **Structural basis for dimerization in DNA recognition by Gal4.** *Structure* 2008, **16:**1019-1026.
41. Miller J, Stagljar I: **Using the yeast two-hybrid system to identify interacting proteins.** *Methods Mol Biol* 2004, **261:**247-262.

42. Dhanasekaran M, Negi S, Sugiura Y: **Designer zinc finger proteins: tools for creating artificial DNA-binding functional proteins.** *Acc Chem Res* 2006, **39:**45-52.
43. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression.** *Cell* 2013, **152:**1173-1183.
44. Möglich A, Ayers RA, Moffat K: **Design and Signaling Mechanism of Light-Regulated Histidine Kinases.** *J Mol Biol* 2009, **385:**1433-1444.
45. Strickland D, Moffat K, Sosnick TR: **Light-activated DNA binding in a designed allosteric protein.** *Proc Natl Acad Sci U S A* 2008, **105:**10709-10714.
46. Levskaya A, Chevalier AA, Tabor JJ, Simpson ZB, Lavery LA, Levy M, Davidson EA, Scouras A, Ellington AD, Marcotte EM, Voigt CA: **Synthetic biology: engineering *Escherichia coli* to see light.** *Nature* 2005, **438:**441-442.
47. Shimizu-Sato S, Huq E, Tepperman JM, Quail PH: **A light-switchable gene promoter system.** *Nat Biotechnol* 2002, **20:**1041-1044.
48. Morgan SA, Al-Abdul-Wahid S, Woolley GA: **Structure-Based Design of a Photocontrolled DNA Binding Protein.** *J Mol Biol* 2010, **399:**94-112.
49. Temme K, Hill R, Segall-Shapiro TH, Moser F, Voigt CA: **Modular control of multiple pathways using engineered orthogonal T7 polymerases.** *Nucleic Acids Res* 2012, **40:**8773-8781.
50. De Smit MH, Vanduin J: **Secondary Structure of the Ribosome Binding-Site Determines Translational Efficiency - a Quantitative-Analysis.** *Proc Natl Acad Sci U S A* 1990, **87:**7668-7672.
51. Salis HM, Mirsky EA, Voigt CA: **Automated design of synthetic ribosome binding sites to control protein expression.** *Nat Biotechnol* 2009, **27:**946-950.
52. Deutscher MP: **Degradation of RNA in bacteria: comparison of mRNA and stable RNA.** *Nucleic Acids Res* 2006, **34:**659-666.
53. Lalaouna D, Simoneau-Roy M, Lafontaine D, Masse E: **Regulatory RNAS and target mRNA decay in prokaryotes.** *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms* 2013, **1829:**742-747.
54. Mitschke J, Georg J, Scholz I, Sharma CM, Dienst D, Bantscheff J, Voss B, Steglich C, Wilde A, Vogel J, Hess WR: **An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp PCC6803.** *Proc Natl Acad Sci U S A* 2011, **108:**2124-2129.
55. Quax TEF, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran WQ, Blombach F, Makarova KS, Brouns SJJ, Forster AC, et al: **Differential Translation Tunes Uneven Production of Operon-Encoded Proteins.** *Cell Reports* 2013, **4:**938-944.

56. Smith HO, Wilcox KW: **A Restriction Enzyme from *Hemophilus-Influenzae* .1. Purification and General Properties.** *J Mol Biol* 1970, **51:**379-391.
57. Knight T: **Idempotent vector design for standard assembly of biobricks.** In *MIT Synthetic Biology Working Group Technical Reports*. Cambridge: MIT; 2003.
58. **The iGEM Registry of Standard Biological Parts** [http://parts.igem.org/]
59. Shetty RP, Endy D, Knight TF, Jr.: **Engineering BioBrick vectors from BioBrick parts.** *J Biol Eng* 2008, **2:**5.
60. Mullis KB, Faloona FA: **Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction.** *Methods Enzymol* 1987, **155:**335-350.
61. Dong B, Mao R, Li B, Liu Q, Xu P, Li G: **An improved method of gene synthesis based on DNA works software and overlap extension PCR.** *Mol Biotechnol* 2007, **37:**195-200.
62. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO: **Enzymatic assembly of DNA molecules up to several hundred kilobases.** *Nat Methods* 2009, **6:**343-345.
63. Hillson NJ, Rosengarten RD, Keasling JD: **j5 DNA Assembly Design Automation Software.** *ACS Synthetic Biology* 2012, **1:**14-21.
64. Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S: **Gene Designer: a synthetic biology tool for constructing artificial DNA segments.** *BMC Bioinformatics* 2006, **7:**285.
65. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA: **Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues.** *Physiol Rev* 2010, **90:**1103-1163.
66. Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, de Mora K, Glieberman AL, Monie DD, Endy D: **Measuring the activity of BioBrick promoters using an in vivo reference standard.** *J Biol Eng* 2009, **3:**4.
67. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R: **Transcriptional regulation by the numbers: models.** *Curr Opin Genet Dev* 2005, **15:**116-124.
68. Leveau JH, Lindow SE: **Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria.** *J Bacteriol* 2001, **183:**6752-6762.
69. Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG, Elf J: **The *lac* Repressor Displays Facilitated Diffusion in Living Cells.** *Science* 2012, **336:**1595-1598.
70. Carrera J, Rodrigo G, Singh V, Kirov B, Jaramillo A: **Empirical model and in vivo characterization of the bacterial response to synthetic gene expression show that ribosome allocation limits growth rate.** *Biotechnol J* 2011, **6:**773-783.
71. Klumpp S, Zhang ZG, Hwa T: **Growth Rate-Dependent Global Effects on Gene Expression in Bacteria.** *Cell* 2009, **139:**1366-1375.

72. Mutalik VK, Guimaraes JC, Cambray G, Mai QA, Christoffersen MJ, Martin L, Yu A, Lam C, Rodriguez C, Bennett G, et al: **Quantitative estimation of activity and quality for collections of functional genetic elements.** *Nat Methods* 2013, **10:**347-353.
73. Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai QA, Tran AB, Paull M, Keasling JD, Arkin AP, Endy D: **Precise and reliable gene expression via standard transcription and translation initiation elements.** *Nat Methods* 2013, **10:**354-360.
74. Lou CB, Stanton B, Chen YJ, Munsky B, Voigt CA: **Ribozyme-based insulator parts buffer synthetic circuits from genetic context.** *Nat Biotechnol* 2012, **30:**1137-1142.
75. Callen BP, Shearwin KE, Egan JB: **Transcriptional interference between convergent promoters caused by elongation over the promoter.** *Mol Cell* 2004, **14:**647-656.
76. Lutz R, Bujard H: **Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements.** *Nucleic Acids Res* 1997, **25:**1203-1210.
77. Labarre J, Chauvat F, Thuriaux P: **Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803.** *J Bacteriol* 1989, **171:**3449-3457.
78. Griese M, Lange C, Soppa J: **Ploidy in cyanobacteria.** *FEMS Microbiol Lett* 2011, **323:**124-131.
79. Kuhlman TE, Cox EC: **Gene location and DNA density determine transcription factor distributions in *Escherichia coli*.** *Mol Syst Biol* 2012, **8:**610.
80. Meyer R: **Replication and conjugative mobilization of broad host-range IncQ plasmids.** *Plasmid* 2009, **62:**57-70.
81. Davison J: **Genetic tools for pseudomonads, rhizobia, and other gram-negative bacteria.** *Biotechniques* 2002, **32:**386-388, 390, 392-384, passim.
82. Frey J, Bagdasarian M (Eds.): **The molecular biology of IncQ plasmids**. London: Academic Press; 1989.
83. Marraccini P, Bulteau S, Cassierchauvat C, Mermetbouvier P, Chauvat F: **A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera *Synechococcus* and *Synechocystis*.** *Plant Mol Biol* 1993, **23:**905-909.
84. Ng WO, Zentella R, Wang Y, Taylor JS, Pakrasi HB: **PhrA, the major photoreactivating factor in the cyanobacterium *Synechocystis* sp. strain PCC 6803 codes for a cyclobutane-pyrimidine-dimer-specific DNA photolyase.** *Arch Microbiol* 2000, **173:**412-417.
85. Anfelt J, Hallström B, Nielsen J, Uhlen M, Hudson EP: **Using Transcriptomics To Improve Butanol Tolerance of *Synechocystis* sp Strain PCC 6803.** *Appl Environ Microbiol* 2013, **79:**7419-7427.

86. Laursen BS, Sorensen HP, Mortensen KK, Sperling-Petersen HU: **Initiation of protein synthesis in bacteria.** *Microbiol Mol Biol Rev* 2005, **69:**101-123.
87. Ma J, Campbell A, Karlin S: **Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures.** *J Bacteriol* 2002, **184:**5733-5745.
88. Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, Molin S: **New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria.** *Appl Environ Microbiol* 1998, **64:**2240-2246.
89. Alfonso M, Perewoska I, Kirilovsky D: **Redox control of *psbA* gene expression in the cyanobacterium *Synechocystis* PCC 6803. Involvement of the cytochrome b(6)/f complex.** *Plant Physiol* 2000, **122:**505-516.
90. Alfonso M, Perewoska I, Kirilovsky D: **Redox control of *ntcA* gene expression in *Synechocystis* sp. PCC 6803. Nitrogen availability and electron transport regulate the levels of the NtcA protein.** *Plant Physiol* 2001, **125:**969-981.
91. Huang HH, Lindblad P: **Wide-dynamic-range promoters engineered for cyanobacteria.** *J Biol Eng* 2013, **7:**10.
92. Paulmier N, Yaniv M, Von Wilcken-Bergmann B, Müller-Hill B: **Gal4 Transcription Activator Protein of Yeast Can Function as a Repressor in *Escherichia coli.*** *EMBO J* 1987, **6:**3539-3542.
93. Wang X, Chen XJ, Yang Y: **Spatiotemporal control of gene expression by a light-switchable transgene system.** *Nat Methods* 2012, **9:**266-269.
94. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, et al: **Creation of a bacterial cell controlled by a chemically synthesized genome.** *Science* 2010, **329:**52-56.
95. Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM, et al: **Total synthesis of a functional designer eukaryotic chromosome.** *Science* 2014, **344:**55-58.
96. Wang K, Neumann H, Peak-Chew SY, Chin JW: **Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion.** *Nat Biotechnol* 2007, **25:**770-777.
97. Davis L, Chin JW: **Designer proteins: applications of genetic code expansion in cell biology.** *Nat Rev Mol Cell Biol* 2012, **13:**168-182.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1149

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-223599

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014