NTNU – Trondheim
Norwegian University of
Science and Technology

Fluorescent Reporters for investigating the Bacterial Stringent Response

Jarle Magnus Ribe Pahr

Chemical Engineering and Biotechnology
Submission date: July 2013
Supervisor: Eivind Almaas, IBT
Co-supervisor: Rahmi Lale, IBT
Martin F. Hohmann-Marriott, IBT

Norwegian University of Science and Technology
Department of Biotechnology

Preface

This thesis is submitted in fulfilment of the requirements for the degree Master of Science/Siv.ing. in chemical engineering and biotechnology at the Department of Biotechnology at the Norwegian University of Science and Technology (NTNU), July 2013. It is the result of a one-semester, full-time (30 ECTS credits) thesis project. The project has been carried out under guidance by Professor Eivind Almaas, Dr. Rahmi Lale and Associate Professor Martin Hohmann-Marriott. The project has been carried out in connection with a collaboration with the laboratory of Dr. Vasili Hauryliuk at Umeå University (Sweden), who has consulted on the project.

Acknowledgements

This project sprung out of the participation of NTNU in the International Genetically Engineered Machine (iGEM) competition in two ways: Firstly, the project is the result of a collaboration between the iGEM 2011 Team NTNU project. I therefore give my thanks to the NTNU iGEM 2011 team. Without their work, this project would not have happened. Secondly, through my participation with the NTNU Team in the iGEM 2012 competition, I acquired much of the lab experience which enabled me to embark on this project. I therefore thank the other members of the NTNU iGEM 2012 team: Gunvor Røkke, Ove Øyås, Eirin Korvald, Rolf Hellemyren and Nina Hole. Thanks also go to Jacob Lamb and Danja Hinke for help and company in the lab. At the Department of Biotechnology, my thanks go to Merete Christensen and the technical staff for helping me get everything I needed, and keeping the lab running every day.

All my advisors deserve thanks for their helpfulness, their understanding and calm response to my inquiries and problems. I feel lucky to have come into contact with and given the opportunity to work with them. Rahmi Lale has been my main advisor for both practical matters and the general direction of the project.

Many big thanks go to Vasili Hauryliuk at Umeå University, Sweden, and Vallo Varik at the University of Tartu, Estonia, for supplying crucial materials and much helpful advice.

Finally, I want to express my gratitude to my parents for their support which I have never appreciated properly.

Abstract

The "stringent response" is a bacterial physiological adaptation to nutritional stress most well studied in the model organism *Escherichia coli*. It is a research topic of interest both for increasing our understanding of bacterial physiology in general, and because of the potential for informing development of treatments for bacterial diseases. The small regulatory molecules penta- and tetra-guanosine phosphate, collectively referred to as (p)ppGpp, are known to be central in the regulation of the stringent response, but convenient methods for assessing intra-cellular concentrations of (p)ppGpp are currently lacking. One possible way to assess intracellular levels of a regulatory molecule is to place a "reporter gene" encoding a fluorescent protein, under transcriptional control of a promoter regulated by the molecule. Here, several plasmids encoding a variant of the green fluorescent protein (GFP) under different promoter sequences for this purpose have been constructed. The promoters differ with respect to whether and how they are effected by (p)ppGpp level, previous evidence showing either postive (up-regulation), negative (down-regulation) or indifferent (no *direct* regulation) transcriptional response to increasing (p)ppGpp levels. Preliminary characterisation of GFP production directed by the different promoters, as measured by fluorescence, have been carried out with cells growing under both non-perturbed conditions and in the face of elevated (p)ppGpp levels. To elevate (p)ppGpp levels, production of a fragment of the ppGpp synthase RelSeq from *Streptococcus equisimilis* was induced by use of a second plasmid.

Sammendrag

Den fysiologiske mekanismen i bakterien *Escherichia coli*, og visse andre arter, kjent som "stringent response" er et forskningstema av interesse både for å øke vår kunnskap om bakteriell fysiologi generelt, og på grunn av potensialet for å dirigere utviklingen av behandlinger for bakterielle sykdommer. Det er kjent at molekylene penta- og tetra-guanosinfosfat, kollektivt betegnet (p)ppGpp, er sentrale i reguleringen av "stringent response", men det er i øyeblikket en mangel på praktiske metoder for å estimere den intra-cellulære mengden ppGpp i levende celler. I dette prosjektet har seks plasmider blitt konstruert der produksjonen av grønt fluoriscerende protein (GFP) blir regulert av ulike promoter-sekvenser som kan være nyttige for dette formålet. Plasmider for uttrykk av en ustabile variant av GFP er også konstruert for flere av promotorene. Disse plasmidene er potensielt mer nyttige for

å følge dynamiske endringer i genuttrykk. Promotorsekvensene er utvalgt på grunnlag av tidligere studier der genuttrykk fra promotorene har vist enten positiv (oppregulering av genuttrykk), negative (nedregulering) eller nøytral (ingen *direkte* regulering) response med hensyn til økte nivåer av (p)ppGpp i cellen. Promotoraktivitetene har blitt målt ved hjelp av fluorescens fra GFP som proxy under rike vekstbetingelser og under betingelser av overproduksjon av (p)ppGpp. De innledende resultatene indikerer en positiv effekt av (p)ppGpp på aktiviteten til to promotorer, i tråd med litteraturen. Disse promotorene kan dermed tenkes å være kandidater for bruk i videre utvikling.

Contents

1 Introduction

The bacterium *Escherichia coli* is among the most well studied of all organisms. For many decades, it has served as a model organism for studies of microbial genetics and physiology. One extensively studied physiological mechanism of *E. coli* is the *stringent response*, which can be activated by several types of physiological stress including nutrient starvation. *Alarmones* are intracellular messenger molecules which are produced as a response to environmental stress. The alarmones guanosine tetraphosphate and guanosine pentaphosphate, collectively referred to as (p)ppGpp, play a central role in the stringent response. The stringent response is characterised by growth arrest and a major decrease in rRNA synthesis. However, (p)ppGpp affects all major biosynthetic pathways, and (p)ppGpp signalling can be considered a global control system in *E. coli*.[1] Affected pathways include cell cycle regulation, virulence induction, mutation frequency, programmed cell death, phage development and quorum sensing.[1]

The significance of the stringent response as a research area is increased by potential medical applications dependent on understanding microbial responses to stress. The stringent response has been implicated as a factor in the formation of bacterial persisters, phenotypic variants that show antibiotic tolerance due to metabolic control rather than specific antibiotic resistance genes.[2] As such, understanding the formation of persisters and maximizing the effect of antibiotics would benefit from increased understanding of the stringent response. Regulators involved in the stringent response has also been found to be important for the virulence of several human pathogens, including *Yersinia pestis*, the bacterium responsible for the Black Death.[3] With a view to this, a synthetic inhibitor of ppGpp synthesis has recently been described and suggested to hold potential as an antibacterial agent.[4]

Because of the major impact of (p)ppGpp on bacterial physiology, it is desirable to be able to assess the intra-cellular levels of these alarmones in real time. However, convenient methods for this are currently lacking. One common method for measuring (p)ppGpp levels requires the use of radioactive labelling compounds and a complex extraction and analytic procedure.[5] This method is time-consuming, and has obvious drawbacks with respect to cost and safety. Chromatographic analysis may also be employed, but is similarly labor-intensive.[6] Recently, several fluorescence-based chemosensors for (p)ppGpp have been developed.[7, 8, 9] Although these systems may offer some advantages, they also require the synthesis of specialized reagents and like the previous methods they cannot be used for real-time, continous measurements. These drawbacks of conventional analytic methods suggest

that an approach to monitoring (p)ppGpp levels using a "reporter gene" system may be worth developing.

Fluorescent proteins are widely used as reporters to assess gene expression and promoter activities, with the Green Fluorescent Protein (GFP) from the jellyfish *Aequorea victoria* being the first widely used example. In place of direct measurements, the expression of genes regulated by these alarmones may be a useful proxy of their concentrations. In this regard GFP has several desirable attributes as a reporter. GFP concentrations may be assessed by measuring fluorescence by the protein at a specific wavelength, offering the possibility for monitoring GFP production in real-time and at the level of single cells.

The goal of this project has been to construct plasmids useful for the purpose of assessing (p)ppGpp levels *in vivo*. For this purpose, several promoter sequences whose activity levels are known to be regulated by (p)ppGpp have been placed in control of GFP expression on separate plasmids. For increased usefulness in monitoring gene expression dynamics, reporter plasmids encoding an unstable version of GFP have also been constructed. Initial experiments have been carried out to demonstrate the feasability of the system and investigate the effect of cellular stress and/or (p)ppGpp levels on GFP levels as measured by fluorescence. To this end, automated fluorescence measurements using a microplate reader has been employed. This approach yields measurements indicating average fluorescence of the cell population rather than at the single cell level. However, the procedure also allows higher-throughput experiments to be performed.[10]

Firstly in this report, the basic features of gene expression regulation and the stringent response in *E. coli* are described, along with known regulatory effects of (p)ppGpp on transcription. A project by the NTNU iGEM 2011 team to construct a genetic circuit to assess production of (p)ppGpp is then reviewed. Under Results, the construction of fluorescent reporter plasmids are described along with results from preliminary characterization experiments showing fluorescence levels as a proxy for promoter activities under both elevated and non-elevated levels of (p)ppGpp.

2 Background

2.1 DNA

All known forms of life utilize deoxyribonucleic acid (DNA) as its hereditary material. The double-helix structure of DNA, consisting of two separate

linear polymers (DNA strands) was first described by Watson and Crick in 1953.[11] A single DNA strand consists of a sequence of nucleobases, organic chemical groups containing nitrogen, linked together by a phosphate-deoxyribose backbone. The two strands run in opposite directions, with the two directions of each strand being referred to as the $5'$ and $3'$ directions or ends. This notation refers to the position of the carbon atom involved in binding between the deoxyribose and phosphate groups when moving in each direction.

Although DNA had already been implicated as a carrier of genetic information [12], the discovery of the DNA structure paved the way for understanding how the biological function and genetic content of DNA depends on the sequence of its nucleotides. The four nucleotides in DNA are Adenine, Guanine, Cytosine and Thymine - abbreviated as A, T, G and C, respectively. The two DNA strands are held together by hydrogen bonding between the nucleotides on each strand, which occur together in complementary pairs. Adenine is paired with Thymine, and Guanine is paired with Cytosine. This complementary pairing of nucleotides implies that if the nucleotide sequence of one strand is known, the sequence on the other strand is also defined. As such, DNA sequences are customarily written by describing only the nucleotides on the strand of interest, proceeding from the $5'$ to the $3'$ end with respect to that strand.

The process of determining the partial or complete nucleotide sequence in a DNA molecule is called DNA sequencing. For small-scope sequencing applications, a method based on that described by Sanger et al [13] and referred to as "Sanger sequencing" is generally used.[14] The raw output data from Sanger sequencing is a chromatogram with colored traces. Each nucleobase is represented by a different color, and a peak of one color at a point in the chromatogram indicates the presence of the corresponding base. Ideally, the chromatogram peaks are neatly ordered and separated, allowing the sequence to be automatically determined ("called") by a computer program. However, the signal quality in chromatograms very and manual inspection of the may sometimes be necessary.

Most bacteria contain the vast majority of their genetic information in a single circular DNA molecule, the bacterial chromosome, with an average size of 4 million basepairs.[15] Many bacteria also contain plasmids, smaller DNA molecules which are generally not essential for life of the organism and that replicate independently from the chromosome.[16] In some cases, a bacterium may acquire plasmid DNA by direct transfer from another bacterium by conjugation.[17] To varying degrees, cells may also take up DNA from the environment, a process which may be enabled (or made greatly more

efficient) by specific manipulations.[18] The stable uptake of plasmid DNA is called *transformation*, and methods for inducing transformation have been developed for a number of species. In those organisms for which transformation protocols are available, plasmids are convenient genetic engineering tools as new genetic elements can be readily introduced. Plasmids are also of great medical importance, as they often carry genes conferring resistance to various antibiotics.[19] Antibiotic resistance genes are also used as selective markers to isolate transformed bacteria and ensure plasmid retention. By treating a cell population with an antibiotic to which a given plasmid confers resistance, those cells which lack the plasmid are selectively killed or prevented from growing.

2.2 Mechanisms of gene expression

The first, and only universal, step in gene expression is the production of a ribonucleic acid (RNA) molecule mirroring the DNA gene sequence. The process producing RNA is called *transcription*, and the RNA molecules produced are referred to as *transcripts*. During transcription, the two strands of DNA separate to allow one of the strands to act as template for the synthesis of a single strand of RNA. RNA is structurally similar to DNA, but contains a ribose group in place of deoxyribose in the backbone, is usually single-stranded, and the nucleotide Thymine has been replaced by the close structural analogue Uracil. The nucleotide position in the DNA molecule which corresponds to the first nucleotide in the RNA strand is the *transcription start site*, and the DNA strand to which the RNA sequence corresponds is called the *coding strand*.

Transcription of DNA to produce RNA is accomplished by the enzyme RNA Polymerase (RNAP) and a region of DNA which stimulates binding by RNAP is called a *promoter*. Depending on the sequence following a promoter, one or several genes may be expressed from a single promoter. The level of transcription due to binding of RNAP to a promoter may be referred to as the *activity* of the promoter and promoter activities may be affected by diverse regulatory factors.[20] While specific transcription factors may regulate a single or a limited number of genes, promoter activities on the global scale are influenced by the overall physiological state of the cell and by the abundance of different RNAP variants.

E. coli RNAP consists of a multi-subunit core enzyme bound to a transcription initiation factor known as the sigma factor.[21] *E. coli* produces several different sigma factors and the core enzyme can bind to each of these vari-

ants to produce a different RNAP holoenzyme preferentially transcribing a different set of genes. Under favorable conditions, the "house keeping" sigma factor σ^{70} dominates, while specialized sigma factors may be produced in response to varying environmental conditions. Thus, major changes in gene expression can occur as the result of production of different sigma factors. In particular, the sigma factor σ^S (also known as RpoS) is considered a "master regulator", is associated with the general stress response of *E. coli*, and alters the gene expression profile majorly when binding to the RNAP core. In addition to regulation by the sigma factors, other transcription factors may also bind to RNAP and affect transcription.

The RNA transcript produced by RNAP may be functional by itself (as in the case of rRNA, or in tRNA, discussed below), or direct protein synthesis through the process known as *translation*.

The directed synthesis of a protein by processing of an mRNA transcript is called translation. While transcription and translation in eukaryotic cells happens in separate cell compartments, the two processes are closely coupled in eubacteria and translation generally starts before transcription is finished. Translation of mRNA is accomplished by *ribosomes*, enzyme complexes containing both protein and RNA components. The RNA component of ribosomes is called ribosomal RNA (rRNA). Another class of RNA molecules called transfer RNA (tRNA) bind amino acids to give a *charged* tRNA and interact with the ribosome to either initiate protein synthesis or continuing it by transferring the amino acid to the growing polypeptide chain. tRNA defines and implements the genetic code (the relationship between DNA sequence and protein sequence) by selectively delivering amino acids to the protein undergoing synthesis. Each type of tRNA binds a specific amino acid and recognizes one or several *codons*, a sequential triplet of nucleotides, in the mRNA sequence. In addition to the codons specifying amino acids, two additional signals are recognized: A *start codon* initiates translation, while a *stop codon* halts translation.

A nucleotide sequence may be divided into different series of triplets depending on where in the sequence the series begins. In a given region of DNA, different codons may thus be translated depending on where translation starts. Each possible way of "reading" a nucleotide sequence as codons is called a *reading frame*. With three reading frames on each DNA strand (transcription always proceeds in the $5'$ to $3'$ direction with respect to the coding strand) and two possible coding strands, there are six possible reading frames for any given DNA segment. A sequence of DNA beginning with a start codon and ending with a stop codon in the same reading frame is referred to as an Open Reading Frame (ORF) and may or may not correspond

to a gene. An *operon* consists of several genes which are co-transcribed to yield a single mRNA. The gene products are then produced by translation of separate ORFs on the mRNA.

To discern between a gene and its gene product when both have the same name, nomenclature dictates that gene names are written in cursive with the first letter in small caps, while protein names are written in non-cursive with the first letter capitalized.

Regulation of gene expression may happen at both the transcriptional and translational level, and several mechanisms can affect the number of proteins produced per transcript for a given gene. *Transcription factors* are proteins or other molecules which bind to DNA or RNAP to affect the level of transcription. Some transcription factors are specific to a single gene, while others may affect the regulation of many genes. By regulating the expression of genes encoding other transcription factor, a single transcription factor may simultaneously affect the expression of many genes. In *E. coli*, a hierarchy of transcription factors mediates gene regulation with many "local" and a few "global" regulators.[22] Regulation of promoters by transcription factors may be positive, negative, or both. Furthermore, promoters may be simultaneously regulated by transcription factors with opposing effects. As such, regulatory effects can be very complex. The action of transcription factors may be altered directly or indirectly by numerous types of molecules. A molecule which stimulates transcription from a given promoter when added to a cell culture is called an *inducer* of that promoter.

Constitutive promoters are those promoters which are not affected by specific transcription factors. However, expression is still dependent on the general physiological state of the cell, such as the availability of RNAP, ribosomes and free amino acids. Even for nominally unregulated promoters, predicting the gene expression level at any given time thus requires knowledge of the overall physiological state of the cell.

For introductory texts covering gene expression and regulation, see references [23] and [24].

2.3 The stringent response in *E. coli*

Adaptation to changing environmental conditions is a crucial aspect of microbial physiology. The stringent response is an adaptational response to nutritional stress that is most well studied in the bacterium *E. coli*. The meaning of the term has varied over time and between different writers. "Stringent response" has been used to refer both to the general physiological

response to nutrient starvation, but also more specifically to all responses caused by elevation of (p)ppGpp.[25] Most recently Dalebroux and Swenson [26] describe stringent response as "a stress response coordinated by guanosine tetraphosphate and guanosine pentaphosphate, in which cells rapidly inhibit synthesis of stable RNA, ribosomes and proteins, leading to growth arrest."

In addition to being present in DNA and RNA, nucleotides are also found in important energy- and signal-carrying molecules. The nucleotides guanosine pentaphosphate (pppGpp) and guanosine tetraphosphate (ppGpp) are central regulators of the stringent response. They can be considered derivates of the more abundant nucleotides guanosine tetraphosphate (GTP) and guanosine diphosphate (GDP), differing by the addition of a pyroposphate group at the $3'$ carbon of the ribose ring. The two nucleotides have similiar regulatory properties and are thus often referred to collectively as (p)ppGpp. They were originally named "magic spots" I and II (MSI, MSII) due to circumstances of their discovery. The discovery was made in the form of two distinct spots appearing on an autoradiogram prepared during a study of nucleotide abundance in *E. coli*, and was reported by Michael Cashel and Jonathan Gallant in 1969 ([27], cited in [28]). In a review article co-authored by Cashel, the authors declared their opinion that much of the "magic" surrounding the molecules has not been lost, in the sense that much is still unknown about their role and modes of action in bacteriological physiology. [25] Nevertheless, knowledge in the field has increased greatly, both with respect to the global cellular effects of (p)ppGpp [29] and the mechanisms of (p)ppGpp synthesis and regulation. Evidence indicates that most bacteria encode ppGpp synthase genes.[30] As such, understanding the stringent response in *E. coli* may also serve to increase our understanding of bacterial physiology in general.

For review articles covering the stringent response, see references [31], [32], [33] and [34].

2.3.1 Synthesis and degradation of (p)ppGpp

Synthesis and degradation of (p)ppGpp is catalyzed in various organisms by proteins belonging to the RelA/SpoT Homologue (RSH) family with (p)ppGpp synthase and/or hydrolase activities.[35] RSH proteins are found in gamma- and beta-proteobacteria and are thought to have evolved through gene duplication from an ancestral Rel protein more widely found in bacteria.[30] The family is named after the (p)ppGpp synthesis/degradation proteins in *E. coli*, which were the first to be discovered.

In *E. coli*, (p)ppGpp is primarily synthesized by the RelA protein.[25] RelA contains domains for both synthase and hydrolase activity, but the hydrolase domain is inactive.[25] The production of ppGpp by the RelA protein is dependent on binding of the enzyme to a ribosome to which an uncharged tRNA is bound, termed a "blocked" ribosome.[36] Such blocked ribosomes are likely to be higher in number under nutrient starvation conditions in which protein synthesis stalls due to lack of amino acids, and gives one potential trigger for the stringent response. Wendrich et al. found that (p)ppGpp inhibited the binding of RelA to blocked ribosomes, while the (p)ppGpp synthase activity has been found to be subject to positive allosteric feedback regulation by (p)ppGpp.[37] Overexpression of RelA in *E. coli* causes accumulation of ppGpp.[38] This decreases growth rate and may inhibit growth completely, but mutations in the gene *rplK* may nullify the effect, as RelA depends on the RplK protein (also known as ribosomal subunit protein L11) for its ppGpp synthase activity.[36]

The SpoT protein is bifunctional, having both weak ppGpp synthesis activity and strong degradation activity.[39] The degradation activity is dependent on manganese ions.[40] Levels of ppGpp have been shown to vary between different laboratory strains, which can be an effect of mutations in the *spoT* gene.[41] Differing basal levels of ppGpp are related to differing levels of stress resistance by affecting the concentration of the alternative RNAP subunit RpoS.[42]

In *E. coli*, deletion of the *spoT* gene alone is generally lethal, as ppGpp accumulates and inhibits growth. Deletion of both the *relA* and *spoT* gene results in ppGpp-deficient cells exhibiting auxotrophies.[43] Different mutation variants produce various complex amino acid requirements, also varying between strains.[43] These requirements may be due to ppGpp-dependent expression of amino acid biosynthesis genes, as (p)ppGpp is known to positively regulate several such genes.[44]

2.3.2 Interaction with general stress adaptation system

The alternative sigma factor RpoS (σ^{38}) is a major regulator of the general stress response of *E. coli*, its function mediated by binding to the RNAP holoenzyme.[42] Elevated levels of RpoS are associated with a number of stress conditions including temperature shock, pH downshift and transition into stationary phase.[45] It has been observed that (p)ppGpp-free mutant strains (due to deletions of the *relA* and *spoT* genes) contain decreased amounts of RpoS. A correlation between increased levels of (p)ppGpp and increased RpoS led to the suggestion that (p)ppGpp is a positive transcriptional

regulator of the *rpoS* gene.[46] However, later evidence has not supported this view.[47] It has been established that RpoS is stabilized by multiple proteins acting to counteract degradation of the RpoS protein.[48] The genes encoding these proteins have been found to be positively regulated by (p)ppGpp, and this may explain in part the observed correlations between (p)ppGpp and RpoS levels.[49][50] However, it has also been reported that ppGpp induces production of RpoS at the level of mRNA translation.[51]

2.4 Transcriptional regulation by (p)ppGpp

2.4.1 Growth rate dependent control of RNA synthesis

Adaptation to varying environmental conditions is an important aspect of microbial physiology.[32] The rate of growth, a central physiological parameter, is dependent on protein synthesis, while proteins are produced by ribosomes. Regulation of growth is thus closely tied to control of the rate of ribosome synthesis.[52] The term growth rate-dependent control refers to the phenomenon when the abundance of a gene transcript or product increases with increasing growth rates.[52] In *E. coli*, the number of ribosomes relative to the protein content of the cell is proportional to the growth rate. Ribosomes consists of rRNA and ribosomal proteins (r-protein), but the rate of r-protein synthesis is adjusted to the rate of rRNA synthesis. Therefore, the study of ribosome synthesis regulation is focused on rRNA synthesis.

Reduction of rRNA synthesis is a major effect of the stringent response caused by down-regulation of promoters in the rRNA operons. The genome of *E. coli* contains 7 rRNA (rrn) operons with the same general structure.[53] Each operon contains two promoters, called in each case P1 and P2. The P1 promoters have been subject to most study, as most transcription originates from this promoter in medium- to fast-growing cells.[53] Upstream of the -35 sequence of each P1 promoter, an activator region contains three to five binding sites for the transcription factor known as Factor for Inversion Stimulation (Fis).[52] Binding of Fis stimulates expression from P1, while (p)ppGpp reduces the strength of the P1 promoter both directly by interactions with RNAP and indirectly by reducing synthesis of the stimulating factor Fis.[52] In addition to the Fis binding sites, an Adenine-Thymine (AT) rich sequence called the UP element is found between the Fis binding sites and the -35 sequence of each P1 promoter.[53] Both the UP elements and binding of Fis increases transcription through interaction with RNAP. It is believed that Fis acts by stabilizing the interaction between RNAP and the rrnBp1 promoter at the open complex step.[54] In addition to regulation by Fis, P1 promoters

may be inhibited by the protein H-NS.[55] The kinetic properties of *E. coli* *rrn* promoters have been investigated in detail by Zhang et al.[56]

Anti-termination: On average, rRNA transcripts elongate at a higher rate than mRNA transcripts.[57] The higher elongation rates of rRNA are caused by anti-termination sequence features in the *rrn* operons which prevent pausing of RNAP during transcription. Each operon contains two anti-termination sequences with three RNA sequence motifs named BoxA, BoxB and BoxC.[58] The BoxA sequence has been found to be necessary and sufficient to confer the anti-termination effect.[59]

Klumpp and Hwa used a stochastic model to simulate the effect of transcription pausing, termination and anti-termination during rRNA transcription in *E. coli*.[60] Their results suggested that suppression of transcriptional pausing by anti-termination mechanisms are essential for sustaining fast growth. Heinrich et al. showed that point mutations in the BoxA anti-termination sequence of a plasmid-encoded *rrn* operon reduced the amount of plasmid-derived rRNA from about 75% to 50% in cells transformed with the plasmid.[61] The mutations did not affect the number of transcription initiations.

During rapid growth in rich media, rRNA and tRNA accounts for 90% of all transcripts, mRNA making up only about 10%.[62] The elongation rate of *rrn* transcripts is nearly unaffected by the growth rate, while the elongation rate of mRNA increases with increasing growth rate.[63] However, when the anti-termination sequence near a *rrn* promoter is removed the *rrn* transcript elongation rate is reduced, and adding an anti-termination sequence to an mRNA-producing gene increased the mRNA elongation rate.[58] Furthermore, for the mRNA genes studied, addition of the anti-termination sequence suppressed the decrease in mRNA elongation rate associated with the stringent response.[58] As the effect of anti-termination on transcription from ribosomal promoters has implications for the behaviour of the promoters when removed from their native context, it is an aspect of rRNA promoter regulation worth paying some attention to.

In addition to more direct regulation of transcription, (p)ppGpp also appears to modulate RNAP sigma factor binding. Gummeson et al. showed that elevating the concentration of *E. coli* RNAP holoenzyme with bound σ^{70} subunit caused increased expression of ribosomal genes and decreased expression of biosynthetic and maintenance/repair genes.[64] The effect was reminiscent of cells with diminished or no production of (p)ppGpp. Cells lacking (p)ppGpp showed elevated levels of free $E\sigma_{70}$ compared with the wild-type cells, and that repression of rRNA synthesis in mutants with elevated

levels of ppGpp could be suppressed by overproduction of σ^{70}. Gummeson et al. concluded that (p)ppGpp modulates the levels of free σ^{70}, and that this is an integral part of the mediation of the stringent response by (p)ppGpp.

As could be expected from the influence on RNAP sigma factor binding, the stringent response is associated with a global restructuring of gene expression patterns involving differential expression of several hundred genes.[65] Traxler et al. used DNA microarrays to profile the transcriptional response to amino acid starvation in wildtype *E. coli* and a (p)ppGpp null mutant. A number of studies have also been carried out where the effect on fewer genes have been investigated more closely and also the mechanisms of regulation been explored. Among the genes positively regulated by (p)ppGpp are several necessary for amino acid biosynthesis genes and transport.[44] For the activation of at least some of these genes the transcription factor DksA plays an important role.[66]

2.4.2 DksA

The transcription factor DksA was discovered in the early 1990s, but its role in the stringent response was unrecognized for some time.[67] Paul et al. discovered that DksA binds to RNAP and is required for the down-regulation of rRNA transcription associated with the stringent response.[68] (p)ppGpp and DksA may act both alone or together to regulate a given gene positively or negatively and in some cases the two regulators have opposite effects.[69] Expression of the *dksA* gene is subject to negative feedback regulation by the DksA protein in conjunction with ppGpp.[70] Recently, Gummesson et al. identified a promoter element facilitating positive transcriptional control by (p)ppGpp/DksA.[71] Most importantly, a mechanistic explanation for how and when (p)ppGpp/DksA inhibit or promote transcription was proposed. Understanding the mechanism of regulation and the sequences which confer different regulatory properties may allow tweaking or designing promoter elements to have a desired transcriptional response to ppGpp and/or DksA.

2.4.3 Differential regulation of ppGpp versus pppGpp

Although both ppGpp and pppGpp have been implicated as regulators in the stringent response, their roles have until recently been investigated largely collectively, and little information has been available on how, if at all, their effects differed.[72] Mechold et al. recently investigated the differential regulation of ppGpp versus pppGpp, and found that pppGpp was a less potent regulator for growth rate and with respect to several promoter activities.[72]

Preferential accumulation of ppGpp or pppGpp was achieved by inducing production of various fragments of a (p)ppGpp synthethase, RelSeq, from *Streptococcus equisimilis*. The RelSeq fragment coding sequences were expressed from plasmids previously used by the same group in a study of the synthetic and degradational activities of the RelSeq protein in which the protein was found to contain two major domains with differing activities.[5]

The native (p)ppGpp synthase gene *relA* had been deleted in all the strains used by Mechold et al. Preferential accumulation of ppGpp was achieved by the production of an N-terminal fragment of RelSeq, RelSeq79-385, expressed from the plasmid pUM9. Predominant accumulation of pppGpp resulted from production of a different fragment, RelSeq1-385, expressed from the plasmid pUM66. In addition to production of (p)ppGpp synthethase, the levels of the pppGpp phosphate hydrolase GppA were manipulated by either chromosomal deletion or consitutive plasmid-based expression of the *gppA* gene. GppA is native to *E. coli* and catalyzes the conversion of pppGpp to ppGpp. One plasmid, named pUM76, was used to achieve preferential accumulation of ppGpp by simultaneous expression of RelSeq79-385 and GppA. In that scenario, ppGpp increased from 9 to 13% of the total guanine nucleotide pool while reducing the pppGpp fraction from 1 to 0.2%.

2.4.4 Regulatory complexity

Although several key regulatory molecules connected with the stringent response, and genes affected by them, have been identified and characterized, it is still the case that other factors will also affect expression of any gene. Even for a hypothetical gene for which (p)ppGpp was the sole specific transcription factor, the translation rate would also depend on the availability of ribosomes, which is itself affected by (p)ppGpp. Thus, (p)ppGpp intrinsically affects gene expression on at least two levels. Furthermore, additional transcription factors and regulatory molecules may interact to determine the final result.

As an example to illustrate the potentially highly complex regulatory networks affecting expression of a gene whose expression is presumed to be directly affected by (p)ppGpp, the gene regulation map from the EcoCyc entry for the argI gene is shown in Figure 1. The database entries for other genes show even more complex patterns and it is clear that predicting or tracking all the concentrations and effects of all the molecules shown in a typical map would be a daunting task. Thus, the goal of properly understanding the interplay of regulatory factors affecting the expression of even a single gene may call for a systems biology approach, taking into account as

much of the molecular context as feasible and using network analysis tools when appropriate. In this vein, Dennis et al. published the results of an analysis of the interactions of RNAP and the rRNA promoters in *E. coli*, in which (p)ppGpp plays an important role.[52]

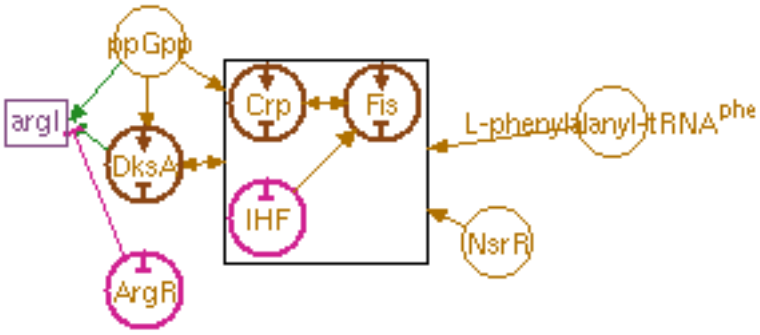


Figure 1: Genetic regulation schematic for gene ArgI in the EcoCyc database. The regulators in the box inter- regulate each other closely, so that interactions between them and other regulators are treated as a collective action or effect. (p)ppGpp and DksA along with the ArgR are shown as the direct-acting regulators. The schematic is very simple compared to many other such schematics found in the EcoCyc database for genes for which evidence of regulation by (p)ppGpp exists.

Below, several promoters of possible use for constructing (p)ppGpp-regulated reporter systems are presented. Both positively regulated, negatively regulated and promoters presumed not to be (directly) regulated by (p)ppGpp are included. All nucleotide positions refer to the transcription start site, with the first transcribed base defined as +1.

2.4.5 Negatively regulated promoters

rrnBp1: There are seven rRNA operons in *E. coli*, named *rrnA*, *rrnB*, *rrnC*, *rrnD*, *rrnE*, *rrnG* and *rrnH*, each regulated by two promoters named P1 and P2. Most work on *rrn* promoters have been focused on the *rrnB* P1 promoter (herafter referred to as rrnBp1).[73] Elevated levels of (p)ppGpp have been shown to repress the activity of the rrnBp1 promoter.[74], while the two promoters differ in their response to amino acid starvation. Zhang, Liang and Bremer found that adding any one of eight amino acids (alanine, glutamine, glutamic acid, isoleucine, leucine, methionine, serine or valine) to a minimal growth medium caused up to a two-fold change in transcription from the rrnBp1 promoter without affecting transcription from the P2 promoter.[75]

Both an upstream (UP) element and binding sites for the transcription factor Fis are known to be important for the level of transcription from rrnBp1.[73] In addition, an anti-terminator sequence, BoxA, immediately downstream of

the P2 promoter, may contribute to the high level of transcription from the *rrnB* operon in its native condition.[58]

Hirvonen et al. studied the relative contributions from Fis binding and UP elements to the transcription from *rrn* P1 promoters. They concluded that all *rrn* P1 promoters have UP elements and are activated by Fis binding, but that the relative contribution from UP elements and Fis binding varied significantly between the operons.[73] Hirvonen et al. defined a "full-length" promoter sequence for rrnBp1 as extending to 154 bp upstream of the transcription start site. Hirvonen et al. further write that the core promoter (defined as nucleotides 41 to +1 with respect to the transcription start site) accounts for less than 1% of the activity of the full-length promoter. The full-length promoter as such defined begins before the first Fis binding site.

In their study of the effect of ppGpp on transcription from several promoters, Barker et al. used an rrnBp1 promoter fragment extending from -88 to +1. [44] This includes the UP element and one Fis binding site, but excludes the first two Fis binding sites.[73] The same promoter region was also used by Paul et al.[66] Barker et al. reported that inhibition *in vitro* depended on the salt concentration: At 30 mM NaCl, the ratio of transcription in the presence of ppGpp to transcription in the absence of ppGpp was 0.99 (Table 1, [44]); At 200 mM NaCl, the ratio was 0.32.

The rrnBp1 promoter can be made insensitive to ppGpp by a three-base mutation (CGC→ATA) between the -10 sequence and transcription start.[76] The mutated promoter is named rrnBp1(dis). In an *in vitro* experiment, the ratio of transcription in the presence of ppGpp to transcription in the absence of ppGpp was 0.81 for the rrnBp1(dis) promoter, compared with 0.32 for the regular promoter (Table 1 in [44]).

2.4.6 ppGpp-insensitive promoters

lacUV5: The *lac* operon in *E. coli* has been extensively studied and has served as model system for the study of bacterial gene regulation.[77] The *lac* operon is both negatively regulated by the repressor LacI and positively regulated by the catabolite activating protein (CAP).[78] The *lac* promoter sequence is shown in Figure 2. LacUV5 is a mutated version of the *lac* promoter which drives strong, constitutive gene expression independent of the regulatory molecules involved in regulation of the *lac* operon.[77]. In their study of transcriptional regulation by ppGpp, Barker et al. used LacUV5 as a reference promoter.[44] They used a segment of the lacUV5 promoter region extending from -48 to + 40.[44] The ratio of *in vitro* transcription from the

promoter in the presence of 200μM ppGpp to transcription in absence of ppGpp was 0.94.

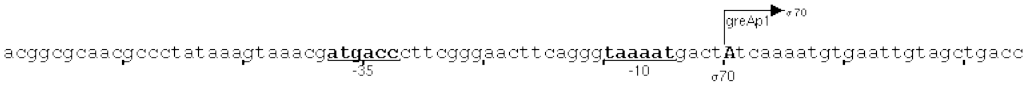


Figure 2: Nucleotide sequence at the wildtype LacZp1 promoter site. Figure from the EcoCyc database.[79]

greA: The transcription elongation facor GreA is an RNAP-binding protein and a structural homolog of DksA.[80] It is involved in re-starting transcription in the case of RNAP halting, by a mechanism involving cleaving of the incomplete RNA transcript.[81] Potrykus et al. found that *greA* expression is driven by two strong, overlapping P1 and P2 promoters - the P1 promoter is σ^{70}-dependent and P2 is σ^{E}-dependent.[82]. Potrykus et al. found in one experiment that two thirds of all *greA* transcripts were terminated early, with only a third becoming mature *greA* mRNA.[82] This was shown to be due to an intrinsic terminator beginning at the first transcribed nucleotide of the P1 promoter.

Potrykus et al. tested the promoter activities of different fragments of the promoter region. The promoter fragment with highest activity was P1, consisting of bases -214 to -146 with respect to the start codon AAT (see supplementary table S1,.[25]). Potrykus et al. reported that the intrinsic terminator accounts for the differences in promoter activities between constructs with the complete promoter region, P1 alone and P1+P2. The sequence at the GreAp1 site is shown in Figure 3

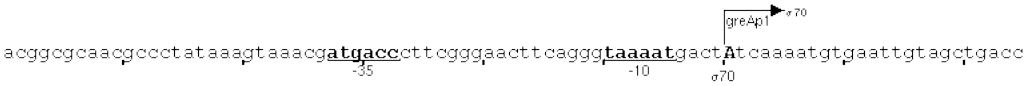


Figure 3: Sequence at the GreAp1 promoter site. Figure from the EcoCyc database.[79]

2.4.7 Positively regulated promoters

iraP: IraP is a small protein which inhibits degradation of the stress response regulator RpoS by inhibiting targeting of RpoS to the protease complex ClpXP.[49] Bougdour and Gottesman studied the change in expression of the *iraP* gene in response to phosphate starvation, which is known to increase ppGpp levels.[49] *iraP* mRNA was barely detectable in growing cells, while transient gene expression (<1 h) was observed upon phosphate starvation. A peak of ∼19-fold induction 15 min after removal of phosphate from

the growth medium was reported. In a ppGpp null mutant strain, *iraP* transcript levels remained low. *iraP* expression was also observed in non-starved cells in stationary phase. Bougdor and Gottesman suggested that inhibition of the SpoT hydrolase activity during phosphate starvation allowed ppGpp accumulation and induction of *iraP* transcription.[49] Glucose starvation also increased the detected amount of *iraP* transcripts 4-fold.

Upon induction of amino acid starvation by addition of serine-hydroxamate to exponentially growing cells, 50-fold induction of *iraP* was observed after 15 minutes. In a *relA* deletion mutant, only two-fold induction was observed under the same conditions. Increase in *iraP* transcription was also observed when ppGpp levels were increased by overexpression of *relA*.

The nucleotide sequence at the *iraP* transcription start site is shown in Figure 4. Bougdour and Gottesman produced several mutated version of the *iraP* promoter. A mutant promoter PiraP(-10-2) increased basal *iraP* transcription 28-fold, and induction upon phosphate starvation was retained. In a strain carrying the PiraP(-10-2) mutation, 8- to 12-fold induction of transcription was observed, compared with 3-fold induction observed in a ppGpp null mutant strain. Additionally, promoters with mutations in the discriminator region between the -10 sequence and transcription start site were produced. The sequence of one promoter with a discriminator mutation, PiraP(dis2) was combined with the sequence of the previously mentioned promoter-up mutation PiraP(-10-2) to give the double-mutation promoter PiraP(-10-2dis2). For this promoter, phosphate starvation gave only a 3.5 fold induction. That is, the relative induction of the up-mutation/discriminator-mutation promoter was comparable to induction of the up-mutant in the absence of ppGpp. For the discriminator-mutant alone, mRNA did not acculumate after phosphate starvation, again similar to results seen with the wildtype promoter in the absence of ppGpp.

iraP has also been shown to be positively regulated by CsgD, a protein associated with biofilm formation.[83]

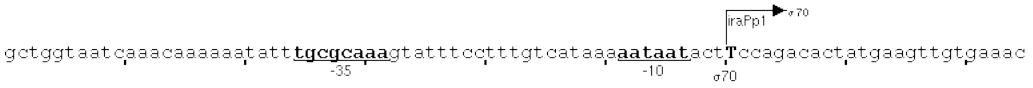


Figure 4: Nucleotide sequence at the iraP promoter site. Figure from the EcoCyc database.[79]

argIp: The *argI* gene encodes ornithine carbamoyltransferase, which catalyzes the sixth step of arginine biosynthesis.[84] The nucleotide sequence at the argIp transcription start site is shown in Figure 5. The *argI* promoter was among the promoters used by Barker et al. in their study of the mechanism

of regulation of transcription initiation by ppGpp.[44] They reported that ppGpp did not stimulate transcription of *argI* or other promoters related to amino acid metabolism significantly *in vitro*, but suggested that slow association of RNAP with the promoter may be responsible for indirect regulation by ppGpp of the promoter *in vivo*. However, the transcription factor DksA, which has later been found to potentiate the effects of ppGpp [68, 66], was not added during the *in vitro* transcription experiments. When testing the effect of ppGpp on *argI* transcription, Barker et al. used a DNA fragment covering nucleotides -45 to +32 with respect to the transcription start site. The same sequence was also used in the study by Paul et al.[66] This sequence contains a binding sites for the ArgR repressor protein[44], and Barker et al. therefore performed experiments in an ArgR mutant strain. With addition of 100 μM ppGpp plus 2 μM DksA, *in vitro* transcription from the *argI* promoter showed 3.5 and 4-fold activation relative to transcription with no addition in *in vitro* transcription experiments using supercoiled and linear DNA templates, respectively. ([66], Figure 1 A).

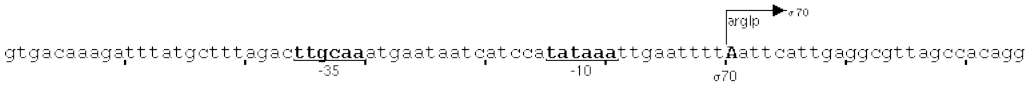


Figure 5: Nucleotide sequence at the argI promoter. Figure from the EcoCyc database.[79]

livJ: The *livJ* gene is part of the LivFGHMJ amino acid transport operon and encodes a periplasmic animo acid-binding protein.[85] A binding site for the tanscription factor Lrp is present between -133 and -121 nt with respect to the transcription start site. The livJp promoter was among the promoters used by Barker et al. in their study of the mechanism of regulation of transcription initiation by ppGpp.[44] The nucleotide sequence at the transcription start site is shown in Figure 6. In their study, Barker et al. used a DNA fragment covering nucleotides -60 to +13 with respect to the transcription start site. The same sequence was also used by Paul et al in a later study.[66] With addition of 100 μM ppGpp plus 2 μM DksA, *in vitro* transcription from the *livJ* promoter showed 8-fold activation relative to transcription with no addition (Figure 1 A in [66]).

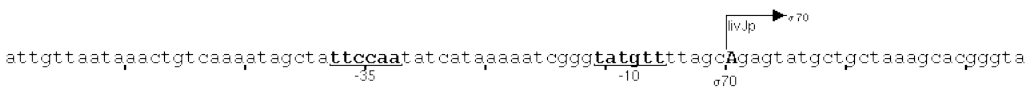


Figure 6: Nucleotide sequence at the *livJ* promoter site. Figure from the EcoCyc database.[79]

hisLp: Genes for biosynthesis of the amino acid histidine are encoded in the hisLGDCBHAFI operon. Paul et al. found that positive regulation of

promoters for amino acid biosynthesis by ppGpp requires the transcription factor DksA. With addition of 100 μM ppGpp plus 2 μM DksA, *in vitro* transcription from the *hisG* promoter showed 4.7 and 10.6-fold activation relative to transcription with no addition, in *in vitro* transcription experiments using supercoiled and linear DNA templates, respectively. In comparison, addition of ppGpp only resulted in 1.3 and 1.7-fold activation, respectively (Figure 1 A in [66]). The nucleotide sequence at the transcription start site of *hisL* is shown in Figure 7.

Figure 7: Nucleotide sequence at the *hisL* transcription start site. Figure from the EcoCyc database.[79]

thrLp: The *thrLABC* operon encodes four enzymes involved in biosynthesis of the amino acids threonine and homoserine. Paul et al. found that addition of 100 μM ppGpp plus 2 μM DksA increased *in vitro* transcription from the thrABC promoter from 7.5 to 7.7-fold.(Figure 1 A in [66]). The nucleotide sequence at the thrLp promoter is shown in Figure 8.

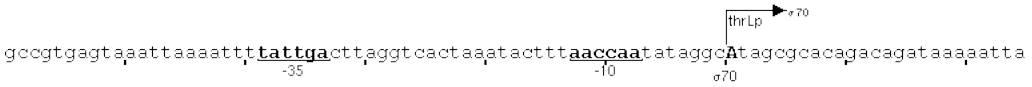


Figure 8: Nucleotide sequence at the Thr promoter.

2.5 Molecular cloning

General methods: The basic methods of molecular cloning, including restriction-ligation based cloning and the polymerase chain reaction (PCR) have been well described by Green and Sambrook.[86] These methods include the use of restriction enzymes for cutting DNA strands at predetermined, sequence-dependent positions and DNA ligase for joining together the ends of DNA strands which have been cut with the same restriction enzymes.

Sequence- and Ligation-Independent Cloning: In conventional restriction- and ligation-based cloning, double-stranded DNA molecules which are to be combined are typically treated with restriction enzymes which leave a short, single-stranded overhang at the ends of the cut molecules.[86] This allows two molecules with complementary overhangs to become associated due to hydrogen bonding between the overhanging nucleotides. Because the overhangs are only a few nucleotides in length, the hydrogen bonding is

not strong enough to permanently join the molecules. Therefore, the enzyme DNA ligase is used to covalently link the sugar-phosphate backbones.

In a paper published by Li and Elledge in 2007, a method for Sequence- and Ligation-Independent Cloning (SLIC) was described which does not require the use of DNA ligase and does not rely on the use of restriction enzymes.[87] By exploiting the exonuclease activity of T4 DNA polymerase, longer single-stranded overhangs can be generated than those resulting from digestion with restriction enzymes. The DNA molecules to be joined must have homologous regions at their ends - sequences resulting in complementary overhangs of about 20 nucleotides have been found to be sufficient.[87]

Altough potentially much faster, the SLIC method has some limitations not presented by conventional cloning. For short DNA molecules, excessive exonuclease activity may pose a problem. It has been observed that the method works best for fragments longer than about 250 basepairs.[88] However, the method has been used sucessfully at least once for a fragment as short as 150 bp.[89] Additionally, the longer overhangs needed compared to restriction-ligation cloning increases the cost of DNA synthesis if this is required.

2.6 Fluorescent reporter proteins

Fluorescence is a phenomenon where a molecule absorbs a photon at one wavelength, followed by the emission of a new photon shortly after.[90] The second photon will be less energetic, and the emitted light thus has a longer wavelength. The entire process may happen at a timescale of nanoseconds. The part of a molecule responsible for fluorescence is called a fluorophore or fluorochrome. Fluorophores contain delocalized electrons able to absorb energy, and may typically consist of aromatic ring structures.[90] Several kinds of biological molecules contain fluorophores and thus exhibit fluorescence.

Green Fluorescent Protein (GFP) is a widely used reporter molecule useful for studying gene expression and other aspects of cellular biology. It is a relatively small protein consisting of 238 amino acid residues, and can be fused to other proteins while retaining its fluorescent activity.[90] GFP was initially isolated from the jellyfish *Aequiora victoria* in 1962 by Shimomura et al. ([91], cited in [90]), a discovery for which in part the 2008 Nobel Prize in Chemistry was awarded.

The amount of fluorescent protein present in a sample can be quantified by using a laser light source and a light detector in tandem. The laser is typically tuned to a wavelength matching the excitation peak of the fluorophore,

the wavelength for which the fluorophore absorbs light maximally. Wild-type GFP has one major and one minor excitation peak at 395 and 475 nm respectively, and an emission peak at 509 nm.[90] GFP fluorescence measurements may be implemented in any of a number of instruments, such as microscopes, flow cytometers or microplate readers. For the first two instruments, it is possible to measure the fluorescence orginating from single cells, while the latter will give bulk data on the entire cell population.

According to March et al., GFP can be applied to report information in three domains: time, location and intensity.[92] In experiments where gene expression is studied or used as a proxy for something else, it is data about the time and intensity domains of fluorescence we wish to obtain. Three amino acid residues, Ser65, Tyr66 and Gly67, forms the fluorophore in GFP, becoming fluorescent after a autocalytic oxidation reaction.[90] This post-translational maturation process causes a lag between the expression of GFP and the onset or increase of fluorescence. It is pH-dependent and requires oxygen - the minimum oxygen requirement has been estimated to be between 0.025 and 0.1 ppm.[93] In the native *Aequiora victoria* GFP, this maturation process takes several hours.[94] This makes the protein unsuitable for observing rapid induction of gene expression. However, many new variants of fluorescent proteins have been derived from the original GFP with shorter maturation times.

The variant of GFP used in this project, GFPmut3b, was originally described by Cormack, Valdivia and Falkow.[95] With respect to the wildtype GFP sequence, GFPmut3b contains the mutations Ser-65 to Gly and Ser-72 to Ala and has several advantages over the wildtype variant. GFPmut3b fluoresces with an intensity approximately 20 times that of wildtype GFP, and with a significantly decreased maturation time - fluorescence may be detected as soon as 8 minutes after induction.[95]

The native GFP protein is very stable, having a lifetime of at least several days. The crystal structure of GFP showing its characteristic barrel motif is shown in Figure 9. The fluorophore is on the inside of the barrel, protecting it from the sorrounding environment and contributing to the long lifetime of the fluorophore.[90] When using GFP as a transcriptional reporter, this stability makes it easy to assess the total transcriptional activity from the start of induction. However, it also makes monitoring transient changes in gene expression difficult. One option is to calculate the rate of change of GFP fluorescence as a measure of gene expression.[96] However, it has been shown that at high concentrations GFP may aggregate to form inclusion bodies, which may reduce the amount of fluorescence.[97]

To facilitate tracking of transient changes in gene expression, Andersen et al. created several unstable variants of GFP by appending short amino acid sequences to the C-terminal end of the GFP variant GFPmut3*.[98] The peptide sequences were associated with the ssrA-mediated peptide-tagging system for protein degradation. The *ssrA* gene of *E. coli* encodes a 362-nucleotide stable RNA molecule which functions to abort translation of damaged or incomplete mRNA transcripts, in the process appending the 11 residue amino acid sequence AANDENYALAA to the unfinished protein.[99] The peptide is then degraded by one of several proteases, among which the ClpXP and ClpAP appear especially active.[100] In the case of degradation by the ClpXP complex, the degradation rate may be increased by the linker protein SspB, which binds both the ssrA tag and the degradation complex, thus bringing them into contact.[101] It has been observed that the degradation rate of ssrA-tagged proteins may be temperature-dependent.[102]

The amino acid sequences used by Andersen et al. were variations of the wild-type ssrA-tag sequence shown above, with different residues in the final three positions.[98] The resulting protein variants were named after the final three residues, giving the protein variants GFP(LAA), GFP(LVA), GFP(AAV) and GFP(ASV). For the mature fluorescent proteins, *in vivo* half-lives equaling approximately 40 minutes for GFP(LAA) and GFP(LVA) were reported, while the half-lives for GFP(AAV) and GFP(ASV) were found to be 60 and 110 minutes, respectively. Sternberg et al. used one of the degradation-tagged GFP variants in a study of bacterial biofilm growth, and reported that the signal to background ratio for stable GFP was three-fold at the slowest investigated growth rate (0.1 $h^{-}1$), while signal from AAV-tagged GFP was detectable only at growth rates above 0.4.[103]

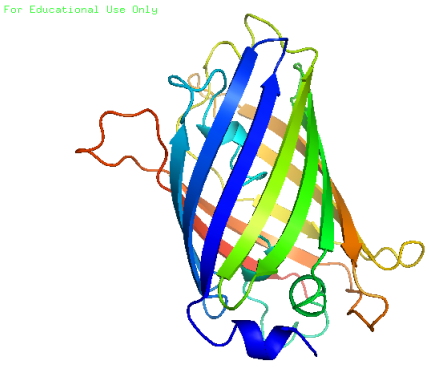


Figure 9: Ribbon diagam showing the β barrel structure of the Green Fluorescent Protein (GFP) from the jellyfish *Aequiora victoria*. The GFP fluorophore is on the inside and protected by the barrel structure. The protection of the wall contributes to the long life-time of GFP fluorescence which makes the protein useful for many imaging applications. Rendered in PyMol from PDB Entry 1EMA.[104]

2.7 Previous work

The International Genetically Engineered Machine (iGEM) competition (`http://igem.org`) is a yearly competition where undergraduate students (and high school students, in a separate division) apply genetic engineering and synthetic biology methods to construct biological parts and devices. The biological "parts" consists of DNA sequences, which may be submitted to the Registry of Standard Biological Parts (`http://partsregistry.org/`), maintained by the BioBricks Foundation (`https://biobricks.org/`). Sequences submitted to the Registry are known as BioBrick parts, and are identified by a number preceded by the prefix `BBa_`. Parts should have a defined biological function, and may classified as *basic* or *composite* parts. Whereas a basic part is defined by submitting a nucleotide sequence directly, composite parts may be defined as a combination of several BioBrick parts in a specified order. The motivation behind the registry lies with the idea that a variety of useful systems can be constructed by combining parts. A combination of sequence elements or parts designed to act or otherwise fulfill a function together is sometimes referred to as a "genetic circuit", by analogy to electric circuits. To give an example of a popular basic part, the BioBrick part `BBa_B0034` consists of a 12-nucleotide sequence comprising an RBS.

Students from the Norwegian University of Science and Technology (NTNU) have participated in the iGEM competition since 2011. The project carried out by the NTNU iGEM 2011 team aimed to produce a sensor circuit in which cells undergoing the stringent response (producing high levels of ppGpp, or in laymans terms, "stressed" cells) would produce a red fluorescent protein. The project led to the contact between the Hauryliuk lab at the Umeå University (Sweden) and the iGEM group at NTNU, which has led to the present project. As such, when starting the present project it was natural to first review the effort by the NTNU iGEM 2011 team. As some of the challenges the NTNU iGEM 2011 team faced may be of relevance also for systems with a different design, a discussion of some of the potential issues affecting the operation of their synthetic gene circuit is given here.

A sketch showing the structure and desired operation of the genetic circuits is shown in Figure 10. The repressor molecule lacI from the *lac* operon of *E. coli* is placed under control of the ribosomal promoter rrnBp1, transcription from which is known to be decreased by (p)ppGpp. The expression of a fluorescent protein, mCherry, is placed under the control of the lac promoter, which is inhibited by LacI. Theoretically, in the event of high ppGpp levels, the levels of LacI should then decrease, repression of the *lac* promoter be lifted and the result be (increased) production of mCherry protein causing cells to fluoresce

red.

In practice, although a difference in fluorescence levels between stressed and non-stressed cells was observed, significant "leakage" was also observed, in which non-stressed cells also expressed enough mCherry protein to become red and fluoresce at significant strength. Some possible reasons for this are discussed below.

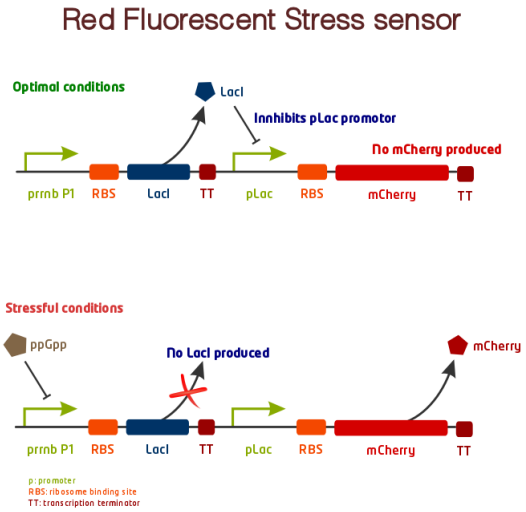


Figure 10: Genetic circuit designed by NTNU iGEM 2011 team. Production of the LacI repressor is placed under control of the rrnBp1 promoter, which is inhibited by ppGpp. When ppGpp levels increase, production of LacI decreases, lifting repression of the lacZ promoter, which controls production of the fluorescent protein mCherry. Figure by NTNU iGEM 2011 team.

Components of the *E. coli lac* operon have been widely employed for genetic engineering purposes. The LacI protein acts as a repressor to inhibit transcription from the *lac* promoter. In the absence of LacI, or in the presence of an inducer counteracting the inhibition, the gene(s) controlled by the *lac* promoter is expressed. The BioBrick part used for LacI production in the NTNU iGEM 2011 circuit is BBa_J24679. In this part, an LVA degradation tag (AANDENYALVA peptide sequence at the C-terminus) has been appended to the LacI protein sequence. If working as intended, the LVA tag causes accelerated degradation of the protein, which would decrease the efficiency of repression. However, the effect of degradation tags on the protein half-life vary between proteins making the exact effect magnitude of the effect hard to predict. Additionally, the LVA tag could hypothetically interfere with the aggregation of the LacI protein into a functional tetramer, as the tetramerization region is also at the C-terminal. Clearly, for any design similar the NTNU iGEM 2011 team were to be used in a reporter for

ppGpp, the stability, production - and degradation-kinetics of the repressing molecule would be of great importance.

It should be noted that the *lac* promoter used was not the native *E. coli lacZ* promoter, but a hybrid lambda Pl LacI regulated promoter, available as the BioBrick partBBa_R0011.

From the lab journal of the NTNU iGEM 2011 team, it appears that the testing of the ppGpp sensor was performed with the BioBrick part in the vector pSB1A2, which is a high copy number (100-300 copies/cell) plasmid. Although the down-regulation of the rrnBp1 by ppGpp is well established, it is not clear if all major regulatory factors are taken into account. Paul et al. write that the protein DksA is absolutely required for regulation of ribosomal promoters by ppGpp.[68] It is conceivable if not likely that the levels of DksA may be insufficient to achieve down-regulation of all copies of rrnBp1 on a high-copy number plasmid. The transcription factor Fis may also contribute to the normally high level transcription from rrnBp1.[105]

The experiment by Tedin et al. which inspired the use of rrnBp1 in the circuit, was performed in the *E. coli* strain VH271.[74] VH271 was constructed from strain MC4100 by phage transduction of a rrnbP1-lacZ construct.[106] The MC4100 strain carries the relA1 mutation, which eliminates the regular RelA-dependent synthesis of ppGpp. The Lac operon is also deleted in this strain. Residual synthesis by the SpoT protein can still be expected.

Hernandez and Bremer found that rrnBp1 promoter activity decreased exponentially with cytosolic ppGpp concentration, observing more than 90% repression at ppGpp saturation in their experiment. Meanwhile, it has been reported that ppGpp strongly up-regulates transcription of the lac operon *in vitro*.[107]

In the experiment by Paul et al. the rrnBp1 promoter with *lacZ* reporter gene was present as a single copy in the bacterial chromosome, while the *relA* ppGpp synthase gene was present on plasmid pKT31.[68] pKT31, sconstructed in that study, was based on the plasmid pUHE21-2, derived from the pDS-plasmid family, in which expression is controlled by a phage T7 A1 promoter with two *lac* operators. As such, protein production can be induced with IPTG, eliminating repression by LacI. Replication of the plasmid was controlled by the ColE1 replicon, for which a medium number of plasmid copies ($\sim$15-20) is expected. However, under some circumstances, up to 9-fold copy number variation can be observed in ColE1-replicon plasmids, depending on the growth conditions and specific promoters/terminators in the plasmid.[108] The rrnBp1 promoter was shut off by inducing expression of a RelA fragment from the plasmid. The issue of plasmid copy number is

not directly addressed by Tedin et al., but it may be worth noting that the ppGpp-synthesis gene was present in more copies than the *rrnBp1* promoter.

Taking into account the above, although the negative effect of ppGpp on transcription from the rrnBp1 promoter is not in question, all the available information pertains to artificially created strain-plasmid systems in which the individual effects of the different factors potentially affecting expression are difficult to ascertain.

3 Materials and methods

3.1 Identification of promoter sequences

A literature search was performed to identify possibly suitable promoter sequences for a ppGpp-regulated fluorescent protein reporting system. Relevant research articles were reviewed to obtain sequences which have been previously used and for which experimental data was available indicating their response to ppGpp. The EcoCyc database (`http://www.ecocyc.org/`, see also [79]) was used by reviewing the list of transcription units associated with ppGpp or the ppGpp-DksA complex.

3.2 Growth media

Unless otherwise noted, all culturing was carried out using low-salt Lysogeny Broth (LB Lennox: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl in H_2O) with appropriate antibiotics as needed (kanamycin: 50 μg/mL, ampicillin: 100 μg/mL). For growth on solid media, LA medium (LB + 20 g agarose/L) was used.

M9 minimal medium with 4% glucose was prepared by mixing 175 mL 10X M9 salts (64 g Na_2HPO_4 ·7 H2O, 15 g KH_2PO_4, 2.5 g NaCl, 5 g NH_4Cl. Dissolve in deionized water to 1000 mL) with 175 mL H_2O before autoclaving, followed by addition of 200 uL each of filter-sterilized (0.2μm filter) solutions of μL 0.1 M CaCl2 and 2M $MgSO_4$, plus 4 mL 20% glucose, also filter-sterilized.

3.3 Strains and plasmids

A list of all plasmids used or constructed in this project is shown in Table 1. Plasmid pSB-M1g [109] was obtained from transformed *E. coli* Dh5α cells

Table 1: Plasmids and strains used in this project

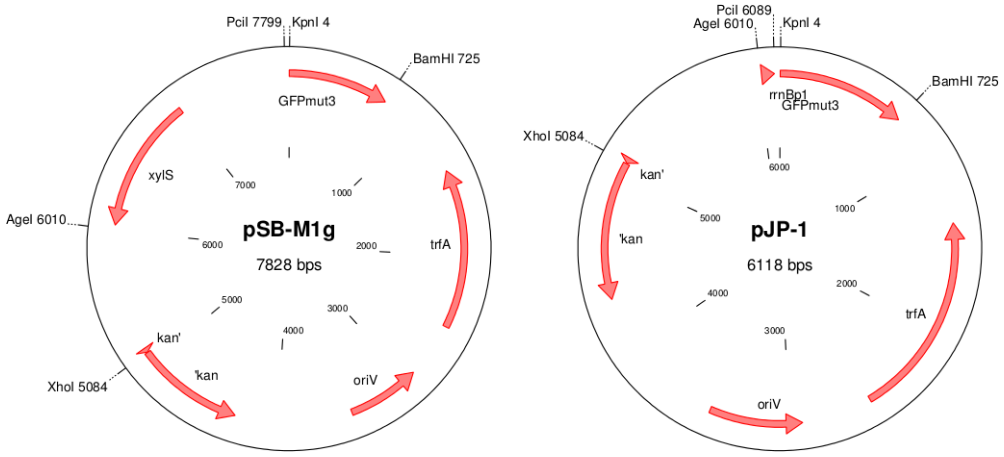
Plasmids	Description	Source
pSB-M1g	XylS/Pm GFPmut3b	[109]
pJP-1	rrnBp1 GFPmut3b	This project
pJP2	GreAp GFPmut3b	This project
pJP-3	LacUV5 GFPmut3b	This project
pJP-4	ArgIp GFP mut3b	This project
pJP-5	IraPp GFP mut3b	This project
pJP-6	LivJp GFP mut3b	This project
pJP-11	XylS/pM GFPmut3b-LVA	This project
pJP-13	GreAp GFPmut3b-LVA	This project
pJP-14	LacUV5 GFPmut3b-LVA	This project
pJP-15	ArgIp GFP mut3b-LVA	This project
pJP-16	IraPp GFP mut3b-LVA	This project
pJP-17	LivJp GFP mut3b-LVA	This project
pUM9	RelSeq(aa 79-385); C-terminal His tag	[5]
Dh5α	Cloning strain	[111]
BW27784	*E. coli* K-12 derivate strain constitutively expressing arabinose transporter.	[112]

as frozen glycerol stock at the Department of Biotechnology. pSB-M1g is a 7828 bp low copy number (5-7 copies) ([110], cited in [109]) circular plasmid containing an origin of replication (oriV) from the plasmid RK2, the *kan* gene conferring kanamycin resistance (Kan^{R}) and the XylS/*Pm* regulator/promoter system regulating the expression of the Green Fluorescent Protein variant GFPmut3b.[109] The low copy number property may is desirable for the present use of the vector in a fluorescent reporter system, as variations in output caused by variations in copy number are limited. An RBS is present immediately after the PciI restriction site to which the downstream end of the promoters are ligated. Thus it was not necessary to include an RBS or any part of the untranslated region when cloning the promoters.

Plasmid maps of pSB-M1g and of plasmid pJP-1 are shown in Figure 11. The nucleotide sequence of pSB-M1g is shown in Appendix A.

E. coli strain DH5α was used for cloning and in the experiment comparing GFP with GFP-LV, while *E. coli* strain BW27784 was used for the majority of the experiments measuring GFP production. This strain was chosen for its usefulness in allowing homogeneous induction of the pBAD promoter. pBAD is the promoter regulating the araBAD operon in*E. coli*, and has been exploited to regulate protein expression.[113] The araBAD operon encodes

genes facilitating uptake and metabolism of the simple sugar arabinose, and expression of genes under control of the pBAD promoter can be induced by addition of arabinose to the growth medium. In wildtype *E. coli*, the uptake of arabinose is facilitated by an high-affinity transporter encoded by the *araFGH* operon and a low-affinity transporter encoded by the *araE* gene.[112]. Expression of both transporters is activated by arabinose, mediated by the activator protein AraC and initial induction in presence of the inducer is assumed to be a stochastic process.[112]. As a result, arabinose-inducible promoters, such as the pBAD promoter of the araBAD operon, are auto-inducible: Cell populations to which the inducer is added show a mix fully induced and non-induced cells.[114] *E. coli* strain BW27784 has suffered a chromosomal deletion of the araFGH operon, while the araE transporter is expressed constitutively. [112]. This allows for homogenous induction of the pBAD promoter.



(a) pSB-M1g plasmid map.

(b) pJP-1 plasmid map.

Figure 11: Plasmid maps of plasmids pSB-M1g (starting material) and pJP-1 (representative constructed plasmid). Genes are shown as arrows pointing in the direction of translation. pSB-M1g is a low-copy number plasmid encoding the fluorescent protein GFPmut3b under transcriptional control of the XylS/*Pm* promoter system. Plasmid replication is initiated from the *oriV* region, with copy number regulation being dependent on the *trfA* gene product. Plasmids pJP-1 to pJP-6, containing different promoter elements, were constructed by digestion of pSB-M1g with restriction enzymes AgeI and PciI, excising the XylS/*Pm* region, followed by ligation of the promoter element to the resulting major fragment. In plasmid pJP-1, expression of GFP from the rRNA promoter rrnBp1. Transcription from rrnBp1 is inhibited by (p)ppGpp.

3.4 General procedures

Instruments: When necessary, bacterial growth was monitored by measuring optical density (OD) at 600 nm (OD600) using a PerkinElmer Lambda

Table 2: DNA oligomers used in this project.

Name	Sequence (5′-3′)	Length (bp)
rrnB p1_74bp_FWD	agccgggcgatgccaaccgggttgcgcggtcagaaaatta	40
rrnB p1_74bp_REV	ctccattattattgtacatgagtggtggcgcattatagg	39
GreA_60bp_FWD	agccgggcgatgccaaccggggcgcaacgccctataaagt	40
GreA_60bp_REV	ctccattattattgtacatgatagtcattttaccctgaagttccc	45
rrnB p1_74bp_FWD_R	caaccggtgttgcgcggtcagaaaatta	28
rrnB p1_74bp_REV_R	gtacatgtagtggtggcgcattatagg	27
GreA_60bp_FWD_R	taaccggtggcgcaacgccctataaagt	28
GreA_60bp_REV_R	gtacatgtatagtcattttaccctgaagttccc	33
LacUV5_49bp_R_FWD	caaccggtgcaccccaggctttacactttatgcttccggctcg	43
LacUV5_49bp_R_REV	gtacatgttccacacattatacgagccggaagcataaagtgta	43
ArgI_46bp_R_FWD	caaccggtgctttagacttgcaaatgaataatcatccatat	41
ArgI_46bp_R_REV	gtacatgttaaaattcaatttatatggatgattattcattt	41
iraP_61bp_R_FWD	caaccggtgctggtaatcaaacaaaaaatatttgcgcaaagtatttcc	49
iraP_61_bp_R_REV	gtacatgtaagtattatttttatgacaaaggaaatactttgcgcaaat	48
LivJ_61_bp_R_FWD	caaccggtattgttaataaactgtcaaaatagctattccaatatcata	48
LivJ_61_bp_R_REV	gtacatgttgctaaaacatacccgatttttatgatattggaatagcta	48
His_61bp_R_FWD	caaccggtgccataaaatatataaaaaagcccttgctttctaacgtgaa	49
His_61bp_R_REV	gtacatgtgtcttttaacctaaaccactttcacgttagaaagcaagggc	49
Thr_73bp_R_FWD	caaccggtaactggttacctgccgtgagtaaattaaaattttattgacttaggtc	55
Thr_73bp_R_REV	gtacatgttgcctatattggttaaagtatttagtgacctaagtcaataaaatttt	55
GFP-END-LVA-REV	agaggatcccttaagttaagctactaaagcgtagttttcgtcgtttgctgctttgtatagttcatccatgcc	72
pJP-1_seq5	cagcgtgcgagtgattat	18
pJP-1_seq6	agaccacatggtccttct	18
pSB-SeqA	tgcaagaagcggatacag	18

35 UV/Vis Spectrophotometer. Fluorescence was measured using a Tecan Inifinite M200 Plate Reader, with excitation = 485 nm and emission = 520 nm.

***In silico* cloning and sequence annotation:** Theoretical plasmid sequences were edited and managed using the CloneManager V.6.0 software (Sci-Ed Software, USA). The *E. coli* MG1655 genome sequence was viewed using the Artemis genome browser (Wellcome Trust Sanger Institute, USA).

Oligonucleotides: Oligonucleotides for use in PCR and sequencing reactions were purchased from Sigma Aldrich. Depending on length, the purification process was specified as either desalted or cartridge-purified for each oligomer. A list of oligomers used in this project is shown in Table 2.

DNA sequence analysis: Plasmid DNA was sequenced using the LightRun Sanger sequencing service (GATC Biotech AG, Germany). Sequence files were obtained as .abi (chromatogram) and .fasta (high-quality bases trimmed sequence) files. When necessary, chromatograms were inspected using the CodonCode Aligner software (CodonCode Corporation, USA). The desired and determined sequences were compared using the Basic Local Alignment Search Tool (BLAST) at the NCBI website. (http://blast.ncbi.nlm.nih.gov/Blast.cgi, [115])

Preparation of competent cells: Depending on the required level of com-

petence, competent were prepared using either a rubidium chloride method or the single-step method published by Chung et al.[116]

Following the rubidium chloride method, cells were prepared as follows: Overnight culture of *E. coli* DH5α was diluted 1:50 by adding 4 mL culture into 200 mL pre-warmed ψB medium (2.5 g yeast extract, 10 g bactotryptone, 0.38 g KCl, H2O to 500 mL, adjust to pH 7.6 with KOH then autoclave and add 17 mL sterile 1 M MgSO4) in a 1L ErlenMeyer bottle and allowed to grow (incubating at 37 ℃ with shaking) until an OD600 of 0.3-0.4. The culture was then transferred to four pre-chilled 50 mL falcon tubes, incubated on ice for 10 min and centrifuged for 10 min at 2760 *g* and 4 . In each tube, cells were then resuspended with 15 mL TfBI buffer (30 mM potassium salt, 50 mM MnCl2, 100 mM RbCl, 1 mM CaCl2, 14% glycerol, pH5.8) and centrifuged at 4 ℃. In each tube, cells were then resuspended with 2 mL TfBII buffer (10 mM MOPS, 75 mM CaCl2, 1 mM RbCl, 15% glycerol, pH 7.0) and snap-frozen in a dry ice and ethanol bath before storage at <-70 ℃.

Following the method of Chung et al., cells were prepared as follows [116]: Cells from frozen glycerol stock were inoculated into 5 mL antibiotic-free LB medium and incubated at 37 ℃ with shaking overnight. The overnight culture was diluted 1 : 100 into 30 mL LB and incubated until OD600 reached 0.4-0.5. The culture was then transferred in equal parts to two 50 mL tubes pre-chilled on ice and centrifuged at 6000g at 4 ℃ for 10 minutes. After centrifugation the supernatant was discarded and the cells placed on ice before adding 3 mL Transformation and Storage Buffer (TSS: 10% PEG8000, 5% DMSO, 50mM MgCl2 pH 6,5 and LB medium to 100 mL, filter-sterilized through 0.22 μm filter) to each tube and vortexing. Resuspended cells were aliquoted in volumes of 100 μL into 1.5 mL tubes and used directly for transformation or stored at <-70 ℃.

Bacterial transformation: Supercompetent *E. coli* cells were either obtained from the stocks at the NTNU Department of Biotechnology or prepared using either of the procedures above. Except for cells prepared by the one-step method, the cells were transformed using the heat shock transformation method. For cells prepared by the one-step method, the transformation procedure was identical, except that the heat shock step is omitted. For the heat shock transformation procedure, aliquots of 100 μL frozen cells in 1.5 mL tubes were thawed on ice, 1-5 μL DNA added and the cells incubated for 30 minutes on ice. The cells were then heat shocked in a water bath at 42 ℃ for 45 s, placed back on ice and incubated for 2 min before 200 μL sterile, antibiotic-free LB medium was added to each tube. Cells were then incubated at 37 ℃ with shaking for from 40 min to 2 h before spreading cells on antibiotic-containing agar plates.

3.5 Molecular cloning

Sequence- and Ligation-Independent Cloning (SLIC): In attempts at cloning by SLIC, the one-step protocol presented by Jeong et al. [117] was followed. A reaction mixture containing ∼100 ng plasmid backbone, 1 μL purified PCR product, 0.2 μL 10X BSA, 0.4 μL T4 DNA polymerase in a total volume of 20 μL was incubated for 5 minutes in room temperature, placed on ice and 2 μL of the reaction mix used in transformation.

Cloning of promoter sequences: *rrnB* and *GreA*: The sequences containing the rrnBp1 and greAp promoters were amplified from *E. coli* chromosomal DNA. For use as template, colony material of E. coli Dh5α growing on an LA plate was dissolved in 20 μL H_2O and heated to 98 ℃for 20 min. 1 μL of the resulting suspension was added to the each PCR reactions as template. For promoter sequences other than rrnBp1 and greAp, primers were designed to produce double-stranded DNA by sequence overlap extension (SOE). In this method, the two primers both act as templates by binding to each other and no chromosomal or other DNA template is used. The obtained DNA fragments were purified, digested with enzymes AgeI + PciI, and ligated to pSB-M1g digested with the same enzymes and treated with Calf Intestine Phosphatese to prevent re-ligation of the original vector.

Isolation of DNA: Plasmid DNA was isolated from liquid cultures using the Promega SV Wizard Miniprep kit and the MacheryNagel NucleoBond Xtra Midi kit for small scale and medium-scale preparations, respectively.

DNA amplification: DNA was amplified by the Polymerase Chain Reaction (PCR) using either an EppenDorf Mastercycler or an OpenPCR thermocycler. For all amplification reactions, the Phusion High Fidelity polymerase (New England Biolabs) was used, following the protocol provided by the supplier.

Purification of PCR products: PCR products were purified using the QiaQuick PCR purification kit (Qiagen corporation, the Netherlands).

Extraction of DNA fragments from agarose gel: DNA fragments isolated by electrophoresis were purified using the QIAEX II Gel Extraction Kit (Qiagen, Netherlands).

Gel electrophoresis: Visualization and separation of DNA fragments was performed by gel electrophoresis in TAE buffer, 0.8 - 1.2 % agarose and GelGreen dye. Imaging was performed using the BioRad GelDoc XR+ system.

DNA fragment ligation: DNA fragments were ligated using a reaction mixture 0.5 μL T4 DNA Ligase (NEB), 2 μL T4 DNA ligase Buffer, appropriate amounts of each molecular fragment (to give molar ratios of 1:1 to

5:1 insert to backbone) and H_2O to 20 μL. Samples were incubated from 3h to overnight at 16 ℃with shaking (∼350 RPM) using an EppenDorf Thermomixer.

Colony PCR: For direct PCR amplification from plasmid DNA in colonies on agar plates, the following procedure was used. Bacterial material from distinct colonies was dissolved in 20 μL H_2O each and suspended by vortex mixing. 1 μL of each suspension was pipetted out in labelled, separate spots on a new agar plate, and 1 μL of the suspension used as template in the PCR reaction.

LVA degradation tag:

The GFP ORF in plasmid pSB-M1g was edited to incorporate the LVA degradation tag by the following procedure: The DNA sequence extending from 101 bp upstream of the GFP ORF in pSB-M1g, ending at the last base before the stop codon TAA in the GFP ORF and appending the LVA tag sequence, stop codon and BamHI restriction site, was amplified by PCR. The primer pair pSB-SeqA/GFP-END-LVA-REV was used, in which the inserted sequence is appended to the 5′ end of the latter primer. The PCR product was purified, digested with KpnI and BamHI, and ligated together with pSB-M1g cut KpnI BamHI and treated with Calf Intestinal Phosphatase (CIP) to prevent re-ligation of the original plasmid.

3.6 GFP fluorescence measurements

GFP production in overnight cultures: Colonies of *E. coli* BW27784 transformed with plasmids pJP-1 to pJP-6 were inoculated into LB or M9 medium supplemented with 50 μg/mL kanamycin and incubated at 37 ℃with shaking overnight. A control culture of *E. coli* BW27784 in antibiotic-free LB or M9 medium was also inoculated. The next day, 100 μL culture was transferred in triplicate to separate microwells on a transparent-bottom microwell plate. Gain was set to 50 in manual. Fluorescence and OD600 values For both growth media the experiment was repeated with three biological replicates (independently inoculated and sampled cultures).

GFP production profiles:

Colonies of *E. coli* BW27784 transformed with plasmids pJP-1 to pJP-6 were inoculated in LB with appropriate antibiotics and incubated as for the overnight culture experiments. The following day, the cultures were diluted $1:100$ into fresh medium with antibiotics and 100 μL of the diluted cultures transferred in triplicate to separate microwells. Using the microplate reader

running in kinetic mode and maintaining an incubation temperature of 30 ℃, Fluorescence and OD600 values was measured every 15 minutes for 6 hours. Gain was manually adjusted to 50, but automatic gain-correction and signal correlation was allowed in case of signal saturation.

Effect of (p)ppGpp on promoter activities: *E. coli* BW27784 was transformed with plasmid pUM9, and a single transformant used to inoculate an overnight culture. Competent pUM9-bearing cells were prepared from this culture using the method by Chung et al. Separate co-transformants carrying pUM9 and pJP-series plasmids were then produced by following the same transformation protocol as previously. For maintenance of pUM9 and pJP-plasmids together, co-transformed cells were grown using medium supplemented with 50 μg/mL kanamycin and 100 μg/mL ampicillin. Cultures of co-transformants and a pUM9-only transformant were inoculated from agar plates and incubated overnight in LB with the appropriate antibiotics. The following day, cultures were diluted in duplicate 1:100 into fresh medium with or without inducer (0.5% w/v arabinose). From each of the induced and non-induced samples, 100 μL was transferred in triplicate to separate wells in a microplate. The microplate was immediately transferred to the plate reader and fluorescence/OD600 monitored for 6 hours as in the previous experiment.

Investigating effect of LVA-tag on GFP stability:

The effect stability of GFP and GFP-LVA was investigated by the following experiment. Cultures of *E. coli* DH5α transformed with plasmids pJP-3 (LacUV5-GFP) and pJP-14 (lacUV5-GFP-LVA) inoculated from single colonies on LA plates, along with a culture of Dh5α inoculated from frozen glycerol stock, were incubated overnight at 37 Cwith shaking, and diluted 1:50 or 1:100 (depending on measured OD600) into 10 mL LB + Kan or LB, respectively. The freshly inoculated cultures were allowed to grow until OD600 0.5-0.6, when 2 x 1 mL were sampled from each culture into separate 13 mL tubes. For each culture, one sample was termed "control" and the other "inhibited". To each "inhibited" sample, 5 μL 15 mg/mL chloramphenicol in 100% ethanol was immediately added to a final concentration of 75 μg/mL, and the sample mixed by rapidly inverting the tube 10 times. From all samples, 3 x 100 μL were then transferred to separate wells on a Nunclon 96 Flat Bottom Black Polystyrol microtiter plate. Fluorescence was measured every 2 minutes for 30 minutes starting immediately after completion of sample transfer. The experiment was repeated twice for a total of three biological replicates.

4 Results

4.1 Promoter sequence selection

A list of the promoter sequences used in this project is shown in Table 3. Where possible, the promoter fragments used were based on sequences already used in studies of transcriptional regulation by (p)ppGpp. For the rrnBp1 promoter, the sequence region was based on that used by Sternberg et al.[103] For the LacUV5, ArgI, HisG, ThrABC and LivJ promoters, regions were based on those used in the study by Barker et al.[44].

Table 3: Promoter sequences used in this project. rrnBp1 is an rRNA promoter whose activity is inhibited by ppGpp; GreAp1 and LacUV5 promoters are not directly regulated by (p)ppGpp; the IraPp1, ArgIp, LivJp and HisGp promoters are all subject to positive regulation by (p)ppGpp. DNA fragments carrying the promoter sequences were produced using either PCR amplification from *E. coli* genomic DNA, or by joining pairs of oligonucleotides by sequence-overlap extension (SOE).

Promoter	Sequence (5′-3′, plus strand)	Length (bp)
rrnBp1	gttgcgcggtcagaaaattattttaaatttcctcttgtcaggccggaataactccctataatgcgccaccact	73
GreAp1	ggcgcaacgccctataaagtaaacgatgacccttcgggaacttcagggtaaaatgactat	60
LacUV5	ggcaccccaggctttacactttatgcttccggctcgtataatgtgtggA	49
IraPp1	gctggtaatcaaacaaaaaatatttgcgcaaagtatttcctttgtcataaaaataatactt	61
ArgIp	gctttagacttgcaaatgaataatcatccatataaattgaatttta	46
LivJp	attgttaataaactgtcaaaatagctattccaatatcataaaaatcgggtatgttttagca	61
HisGp	gccataaaatatataaaaaagcccttgctttctaacgtgaaagtggtttaggttaaaagac	61
ThrLp	aactggttacctgccgtgagtaaattaaaattttattgacttaggtcactaaatactttaaccaatataggca	73

For the choice of greA promoter sequence, two options might seem reasonable: To either clone the entire promoter region, or clone only the P1 fragment. In the first case, the native RBS would be included, but from what Potrykus et al. reported this might result in many prematurely terminated transcripts.[82] Alternatively, the P1 fragment only could be cloned. This might give higher activity, but require the separate inclusion of an RBS. As the plasmid pSB-M1g already contains an RBS immediately after the promoter insertion site, this option was viable. For this project, the sequence extending from -58 to +2 with respect to the P1 promoter native transcription start site was selected.

Regarding the iraP promoter, little information about the sequence upstream from the core promoter is available, making any sequence cutoff somewhat arbitrary. A 61bp "core" sequence extending from nucleotide -60 to +1 with respect to the transcriptional start site as described by Bougdour and Gottesman was chosen for use in this project.[49]

4.2 Molecular cloning

Cloning of promoter sequences: Promoter sequences for rrnBp1 and GreA with adapter sequences for SLIC were amplified from *E. coli* chromosomal DNA using primers `rrnB_FWD` and `rrnb_REV`. The PCR products were visualized by gel electrophoresis as shown in Figure 12, which indicated succesful amplification. However, no transformants were obtained after several independent PCR amplifications and transformation attempts. After attempts at cloning using SLIC failed, new primers were ordered with the SLIC adapters replaced by restriction sites for PciI and AgeI. Due to the pattern of the overhangs created when digesting the pSB-M1g vector with AgeI and PciI, the PciI site would not have been preserved in the plasmid constructed by SLIC, and the primers/PCR products could therefore not be reused for conventional restriction- and ligation-based (RL) cloning.

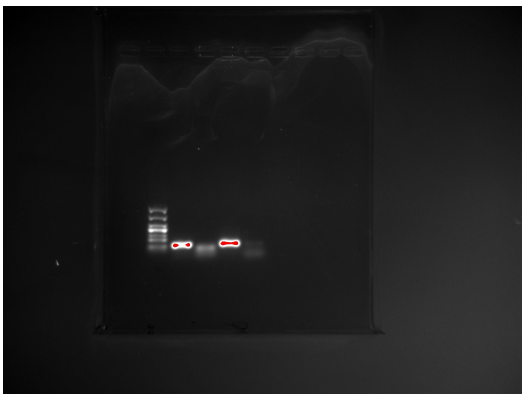


Figure 12: Gel electrophoresis of PCR products containing promoter sequences with adapters for SLIC cloning. From left to rights: DNA reference ladder (DongSheng Bio 50bp), rrnBp1, rrnBp1 negative control (no template), GreA, GreA negative control.

PCR was repeated using the new primers `rrnB_FWD_R` and `rrnB_REV_R`. The result of gel electrophoresis is shown in figure 13. PCR products were purified, digested with restriction enzymes AgeI+PciI and ligated with the major fragment of pSB-M1g digested with the same enzymes. Transformants were obtained for both rrnB and GreA and single clones picked for culturing and sequencing of plasmid DNA.

PCR products for promoters LacUV5 (pJP-3), ArgI (pJP-4), IraP (pJP-5), LivJ (pJP-6), His and ThrABC were obtained by SOE, using the primer pairs `LacUV5_49bp_R_FWD/REV`, `ArgI_46bp_R_FWD/REV`, `iraP_61bp_R_FWD/REV`; `LivJ_61_bp_R_FWD/REV`, `His_61bp_R_FWD/REV` and `Thr_73bp_R_FWD/REV`, respectively. The result of gel electrophoresis of the PCR products for LacUV5 and ArgI promoter sequences is shown in Figure 14, while the re-

sult of gel electrophoresis of IraP, LivJ, His and Thr amplicons are shown in Figure 15.

Plasmids with correct sequences were obtained for promoters rrnBp1, GreA, LacUV5, ArgI, IraP and LivJ. The sequence data is shown in Appendix A.

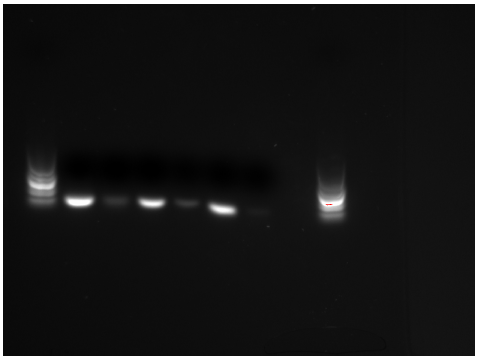


Figure 13: PCR products for SLIC (LacUV5) and RL (rrnB, GreA) cloning. From left to right, amplified/negative control samples (no polymerase) in pairs of two for LacUV5, rrnB and GreA. Far left and right: DNA reference ladder (DongSheng Bio 50bp).

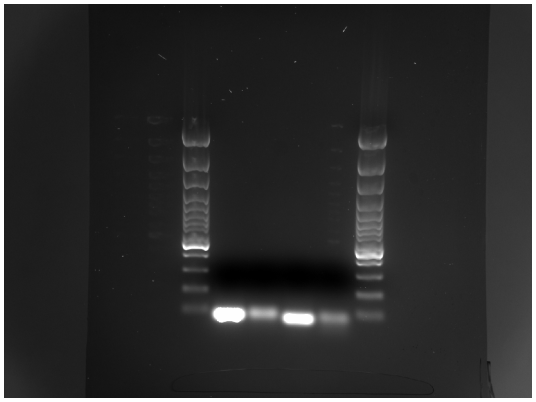


Figure 14: SOE PCR products containing LacUV5 and ArgI promoter sequences for RL cloning. From left to right: LacUV5, LacUV5 negative control, ArgI, ArgI negative control. Far left and right: GeneRuler 1kb ladder.

Colony PCR: To test the usefulness of "colony PCR" as a method for screening candidate clones, parallel PCR reactions using the primer pair `COPCR1FWD + COPCRREV` was set up. Two PCR reactions was performed each using purified pSB-M1g plasmid DNA ($\sim$1 ng/μL) and colony material with colonies of transformed *E. coli* Dh5α containing the same plasmid and growing on an LA + Kan plate. The result of gel electrophoresis of the PCR products is shown in Figure 16.

Construction of GFP-LVA plasmids: The LVA degradation tag was successfully appended to the GFP ORF in plasmid pSB-M1g to create plasmid

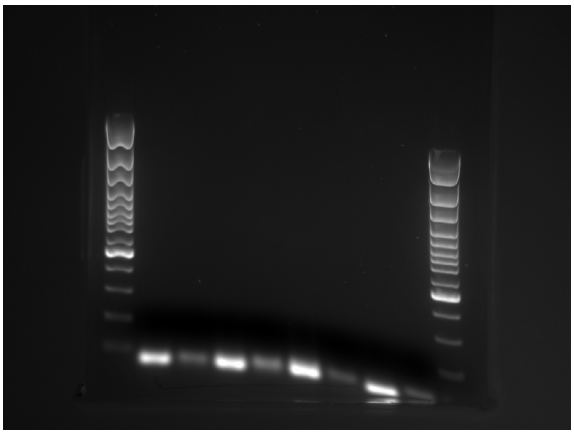


Figure 15: Gel electrophoresis of PCR products containing promoter sequences generated by sequence overlap extension (SOE) by PCR. From left to right, pairs of PCR and negative control samples for IraP, LivJ, His and Thr.

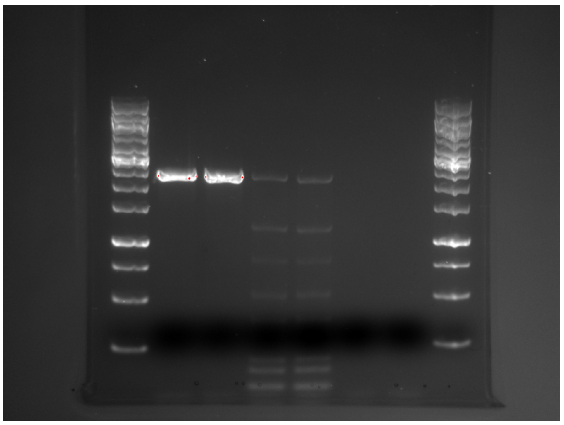


Figure 16: Gel electrophoresis of PCR products from colony PCR test. Left two lanes: Amplicons from PCR reactions using plasmid DNA as template. Right two lanes: Amplicons from PCR reactions using colony material dissolved in H_2O as template. Far left and right: GeneRuler 1kb ladder.

pJP-11. Sequence analysis confirmed the presence of the LVA tag. To allow ligation of the confirmed promoter sequences to GFP-LVA, plasmids pJP-11 and pJP-1 to pJP-6 were digested with restriction enzymes PciI and XhoI. DNA fragments were separated and isolated by gel electrophoresis. The visualized fragments after digestion and gel electrophoresis are shown in Figure 17. The major fragment from plasmid pJP-11 was then ligated with the minor fragments from the various promoter-bearing plasmids. The result of gel electrophoresis of plasmid DNA samples from resulting clones for LivJ, ArgI and IraP, digested with enzymes PciI + XhoI and showing expected fragment sizes, are shown in Figure 18.

The resulting plasmids containing the greA, LacUV5, ArgI, IraP and LivJ promoter sequences were named pJP-13, pJP-14, pJP-5, pJP-16 and pJP-17, respectively.

For plasmid pJP-1, gel electrophoresis revealed an additional unexpected band, and no GFP-LVA plasmid was obtained for this promoter. As this pattern can not readily be explained by the known sequence features of the pJP-1 plasmid. The plasmid was re-transformed from the original sequenced plasmid sample and additionally re-grown from the stored glycerol stock to determine if this was caused by a contaminated or misidentified sample. However, the restriction pattern was the same for all samples tested.

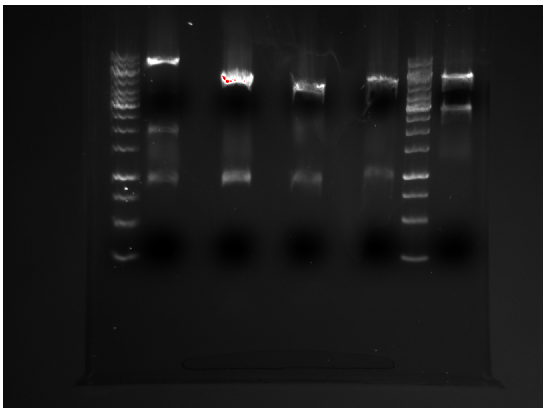


Figure 17: Gel electrophoresis of plasmid DNA cut with restriction enzymes PciI, XhoI. From left to right, pJP-1 (rrnBp1), pJP-2 (GreA), pJP-3 (lacUV5), putative ThrABC plasmid, and pJP-11 (XylS/Pm-GFP-LVA). The first three samples show a band at approximately 1000 bp, matching the expected size of the plasmid fragment flanked by PciI and XhoI restriction sites. The band pattern for plasmid pJP-1 additionaly shows an unexpected band at a size of approximately 2000 bp. The major fragment also appear larger than for the other samples, although there is some variation also among these. The pattern for plasmid pJP-11 shows two bands at their expected positions.

Sequence analysis: A varying number of candidate clones were obtained for each promoter construct for which plasmid DNA was sequenced. Sequencing data is shown in Appendix A. The complete, desired promoter

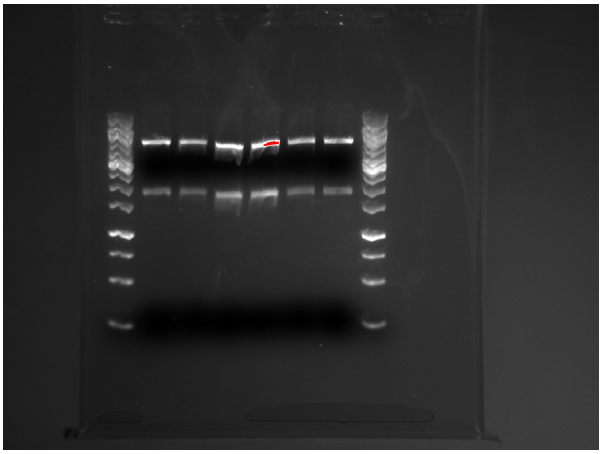


Figure 18: Result of gel electrophoresis of plasmid DNA isolated from candidate clones for promoter-GFP-LVA constructs and digested with enzymes PciI + XhoI. In pairs of two from left to right: Samples from plasmids constructed using fragments from pJP-6 (LivJ), pJP-4 (ArgI) and pJP-5 (IraP) plasmids. The fragment sizes were as expected for all samples.

insert sequences in plasmids pJP-2 (GreA), pJP-3 (LacUV5), pJP-4 (ArgI), pJP-5 (IraP), pJP-6 (LivJ), along with the LVA sequence in plasmid pJP-11, were cofirmed in this way. For plasmid pJP-1 (rrnBp1), the obtained sequence data indicated some mismatches between the expected and observed sequence in the promoter region. However, for this sample the quality of the chromatogram quality was lower than usual, and it is possible that the obtained sequence is not correct. The part of the chromatogram covering the promoter region is shown in Figure A.1. The observed differences constituted 4 nucleotide substitutions at three locations in the 73bp promoter region. However, the suspected changes seem unlikely to have had much, if any, effect on the function of the promoter. An alignment of the expected (Query) and observed (Sbjct) sequences is shown below.

```
Query  1    GTTGCGCGGTCAGAAAATTATTTTAAATTTCCTCTTGTCAGGCCGGAATAACTCCCTATA  60
            ||||||||| |  ||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  205  GTTGCGCGGGCTTAAAATTATTTTAAATTTCCTCTTGTCAGGCCGTAATAACTCCCTATA  264

Query  61   ATGCGCCACCACT  73
            |||||||||||||
Sbjct  265  ATGCGCCACCACT  277
```

His promoter: Several candidate clones for the *his* promoter were sequenced, but none appeared to contain a complete and correct sequence. The nucleotide sequences obtained from samples 59DG30, 59DG46 and 59DG48 are shown in Appendix A. In sequenced sample 59DG30, the first 3 bp of the promoter sequence follow immediately after the AgeI site, but a 7bp fragment (ATAAAAT) is missing. For sample 59DG46 as in 59DG30, a significant gap with respect to the desired promoter sequence is present in the reported sequence, but the gap appears later.

Aligning sequences59DG30 and 59DG48 shows nearly identical sequences (981/990(99%) matches, 3/990(0%) gaps). As the two sequencing reactions gave the same result, this further strengthens the hypothesis that sequencing errors are not to blame, and that the clones differ because of sequence differences present in the PCR amplicons.

Thr promoter: Four candidate clones for the ThrABC promoter construct were obtained, and two were selected for sequencing. The sequencing reaction for the first candidate gave no usable data, while the sequence result for the second sample is shown in Appendix A. Alignment of the Thr promoter 73 bp sequence against Phred-generated .fasta using BLAST yields a 72/73 bp match with 1 gap. The peaks in the chromatogram were clearly resolved, giving no obvious indication of sequencing error. The sequence alignment is shown below.

```
Query  1    AACTGGTTACCTGCCGTGAGTAAATTAAAATTTTATTGACTTAGGTCACTAAATACTTTA  60
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  199  AACTGGT-ACCTGCCGTGAGTAAATTAAAATTTTATTGACTTAGGTCACTAAATACTTTA  257

Query  61   ACCAATATAGGCA  73
            |||||||||||||
Sbjct  258  ACCAATATAGGCA  270
```

All four clones have been retained, and further sequencing may thus be performed at a later date to determine if one of the clones contains the complete, correct sequence.

4.3 Promoter activities

As a proxy for promoter activities, fluorescence from GFP was measured using a microplate reader in several experiments. The results are presented below.

Accumulated expression of GFP from promoters: To investigate the overall activity of promoters and lifetime-accumulation of GFP during culturing to stationary phase, overnight cultures transformed with each confirmed promoter-GFP construct were inoculated, incubated overnight and fluorescence measured the following day. As GFP is stable at timeframe of several days, little to no loss of GFP fluorescence is expected between cessation of growth and the time of measurement. Fluorescence values were adjusted by dividing arbitrary fluorescence units (AU) by cell density (OD600) to give a the values(AU/OD600) shown in Figure 19. The results serve mainly to compare relative promoter strengths (overall promoter activities), and differential effects of the growth medium. For growth in rich medium (LB), the rrnBp1 promoter showed the highest amount of GFP production per OD

unit, and this promoter also displayed the biggest difference in production between the two culture conditions. For the other promoters, the difference was less pronounced. The consitutive LacUV5 promoter and Under growth in M9 medium, the per-OD fluorescence from rrnBp1 and LacUV5-driven GFP expression were comparable.

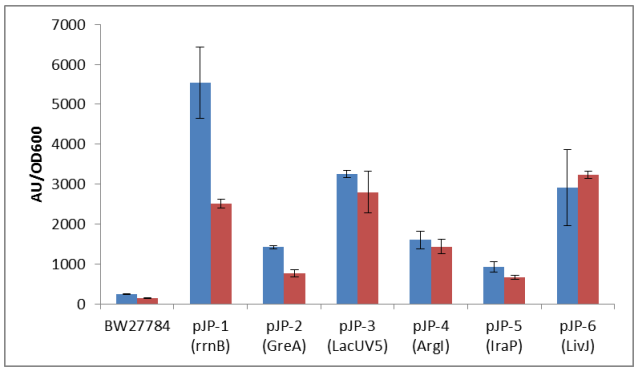


Figure 19: Fluorescence divided by OD600, as measured in overnight cultures growing in LB (blue bars) and M9 (red bars) medium. In rich medium, the rRNA promoter rrnBp1 shows the greatest values. The rrnBp1 promoter also shows the greatest difference between rich (LB) and minimal (M9) medium. Somewhat surprisingly, about the same level of activity is indicated for the amino acid operon promoter LivJ as for the constitutive promoter LacUV5.

Effect of LVA sequence on GFP stability: The effect of the LVA degradation tag on the stability of GFP expressed from the LacUV5 promoter was investigated, as described under materials and methods. After treatment with chloramphenicol to stop protein production, fluorescence was measured every 2 min for 30 minutes. The experiment was performed in triplicate, and the data normalized to starting values are shown in figure 20. The starting fluorescence value were significantly higher for the culture producing unmodified GFP, indicating that either less GFP-LVA than GFP was produced or that accelerated protein degradation kept GFP-LVA from reaching the same concentrations as GFP. The uncertainties in the relative fluorescence levels increase with time, but there is a clear separation between the two trajectories.

GFP production time profiles: To determine the time necessary for significant protein expression, fluorescence was measured starting immediately after for cells growing in LB medium. The results are shown in Figure 21. The relative levels of fluorescence observed by production from the different promoters are similar to that measured in cultures incubated overnight. The rRNA promoter rrnBp1 shows the strongest production and reaches an AU/OD600 value of approximately 1.5 times the initial value at 2h, meaning that GFP production could reliably be inferred from a measurement at

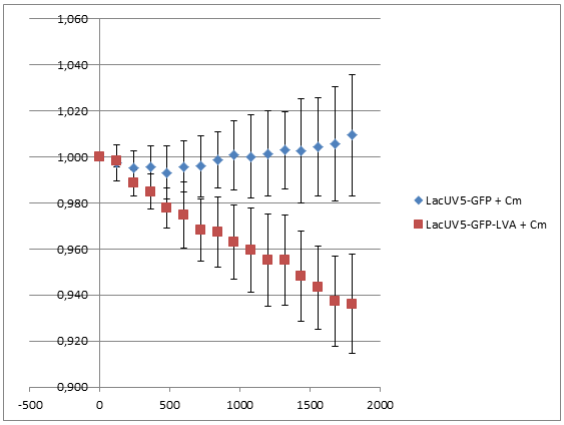


Figure 20: Relative fluorescence as compared to initial values, as measured for GFP and GFP-LVA expressed from the LacUV5 constitutive promoter, after treatment with chloramphenicol to stop protein production. Fluorescence from GFP increases while fluorescence from GFP-LVA decreases, indicating a difference in stability between the two protein variants.

this time point. Interestingly, the AU/OD600 of all the remaining cultures decrease for the first 3 hours. It appears that production of GFP does not outpace cell growth in this phase. As fluorescence increases from these cultures, the LacUV5 constitutive promoter again takes second place after the rrnBp1 promoter, while the iraP promoter shows the weakest response, trailing barely above the control culture curve. As similar starting fluorescence is measured for the non-GFP producing control culture (BW27784) as for the GFP-producing cultures, it appears that only a small amount of the initial fluorescence is caused by GFP carryover from the inoculating culture. However, some carryover does appear to affect the initial values, as the curves for the strong rRNA promoter rrnBp1 and the constitutive promoter LacUV5 are already positively displaced at the starting time.

Effect of induction of a (p)pGpp synthase: *E. coli* strain BW27784 was transformed with plasmid pUM9 allowing arabinose-inducible expression of the ppGpp synthase protein RelA, followed by co-transformation with plasmids pJP-1 to pJP-6. Overnight cultures were diluted 1:100 and incubated at 30 ℃ in the microplate reader, measuring optical density and fluorescence every 15 minutes.

To allow normalization of measured fluorescence by taking into account cell density, the optical density of (absorbance) of microplate cultures at 600 nm was measured during the fluorescence time-series experiments. The resulting data is shown in Figure 22. It should be noted that OD600 values shown cannot be compared with values as obtained using a conventional spectrophotometer with a 1 cm pathlength. While little variation in growth

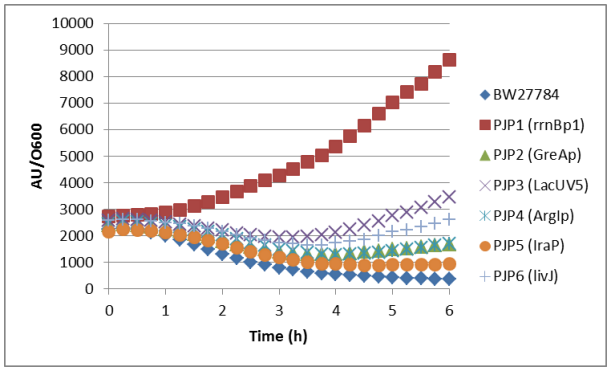


Figure 21: Fluorescence (arbitrary units) divided by OD600 as measured after dilution of LB-medium overnight cultures diluted 1:100 into fresh medium. The curve for the rrnBp1 promoter shows appreciable increase after 1 hour, while AU/OD600 decreases for the remaining promoters, albeit less so than for the control culture not producing any GFP. From 3.5 h onwards, the AU/OD600 values increase for the remaining promoters except IraP. The overall pattern of per-OD unit expression strength is similar to that seen in overnight cultures, with the highest value stemming from the rrnBp1 promoter, followed by LacUV5 and LivJ, and with the least amount produced by IraP.

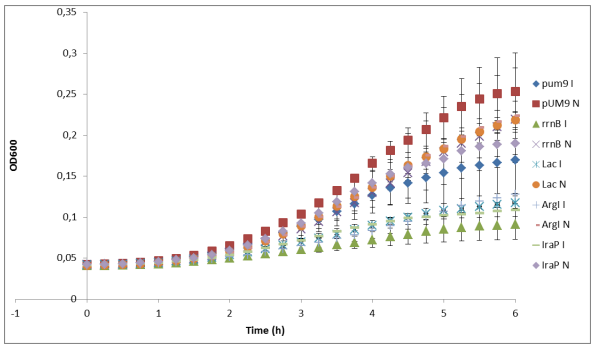


Figure 22: Growth curves for microplate cultures during (p)ppGpp induction experiments. Plasmid pUM9 encodes the RelSeq ppGpp synthase protein expressed from the arabinose-inducible pBAD promoter. *E. coli* BW27784 transformed with plasmid pUM9 only or pUM9 and a promoter-GFP construct were grown in induced (I) and control (N) cultures. RelSeq production was induced by addition of 0.5% arabinose to induced cultures. Results shown are averages of three replicate experiments.

rate is observed in the first 2 h, by 3 h the pattern has emerged. As expected, for cultures in which production of (p)ppGpp was induced by addition of arabinose to the growth medium, growth was slower. For all pairs of induced and non-induced cultures, the growth of cells transformed with pUM9 only was higher than that of cells transformed with pUM9 and an additional, GFP-producing plasmid. The maximal growth was observed for the non-induced culture transformed with pUM9, while the least growth was observed for the culturing harboring the rrnBp1-GFP plasmid (pJP-1) and overproducing (p)ppGpp.

Graphs showing measured fluorescence units divided by OD600 are shown in Figure 23 and Figure 24. The data in both graphs are from same experiments, but are shown separately to decrease clutter. Figure 23 shows that, as before, GFP production from the rRNA promoter rrnBp1 is several-fold larger than the other promoters in the non-induced cultures. Expression of GFP from rrnBp1 is clearly inhibited in the induced culture, indicating that ppGpp accumulates, but the effect is apparent only after several hours. For the constitutive promoter LacUV5, the curves for the two cultures track each other closely the first two hours, diverge slightly and then cross each other later.

In Figure 24, less strong promoters are shown, requiring a decrease in the y-axis scale. The differences between the induced and control cultures are small, but discernable: For both the ArgI and IraP promoters the induced cultures show increased AU/OD600 values compared to un-induced cultures. The differences between the induced and non-induced cultures are small before 2 h, are largest between 3-4 h and then decrease as the AU/OD600 values for the induced and non-induced cultures converge. The overall trend for all cultures is negative. The induced and non-induced cultures of the pUM9-only transformant show little deviation from each other although the two cultures have significantly different growth rates, indicating that the effect of growth rates on the background fluorescence do not skew the AU/OD600 values in a big way.

5 Discussion

Main results: The primary goal of constructing sequence-confirmed, GFP-expressing constructs was completed for six promoters. For two promoters, His and Thr, clones were also obtained, but with sequence deviations. For the rRNA promoter rrnBp1, some possible sequence deviations were also observed, but in this case the quality of sequence data was low, and the

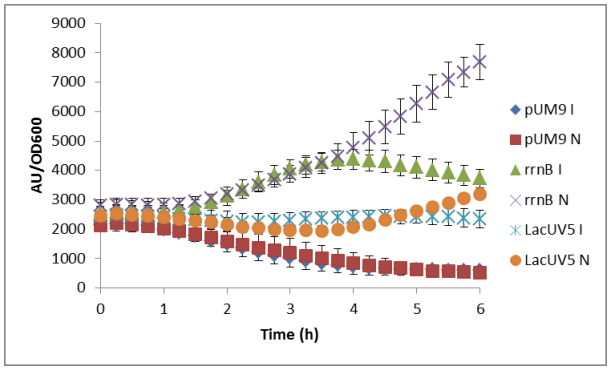


Figure 23: arbitrary units)/OD600 values showing the effect of ppGpp accumulation caused by induction of ppGpp synthase production in induced cultures (I) versus control cultures (N) In the induced cultures, expression of a fragment of the ppGpp synthase RelSeq protein from *Streptococcus equisimilis*, placed under control of the pBAD promoter on plasmid pUM9 was induced by addition of 0.5% arabinose to the growth medium

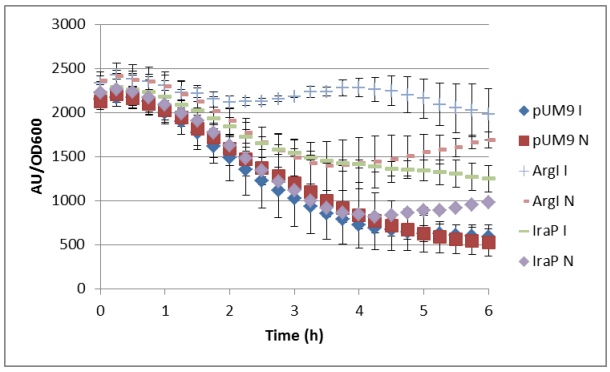


Figure 24: Fluorescence (arbitrary units)/OD600 values showing the effect of ppGpp accumulation caused by induction of ppGpp synthase production in induced cultures (I) versus control cultures (N). In the induced cultures, expression of a fragment of the ppGpp synthase RelSeq protein from *Streptococcus equisimilis*, placed under control of the pBAD promoter on plasmid pUM9 was induced by addition of 0.5% arabinose to the growth medium.

promoter appears to have functioned appropriately. Several constructs for production of LVA-tagged GFP variants were also made, and the LVA peptide sequence was found to decrease the stability of the GFP produced from one such construct. These constructs may potentially be of more use for studying gene regulation dynamics than those constructs producing the more stable, unmodified GFP.

Fluorescence in cultures incubated overnight was measured to determine the relative and overall output of the promoters accumulated from inoculation until stationary phase. As expected, the rRNA promoter rrnBp1 yielded the highest level of fluorescence, followed by the constitutive promoter LacUV5. Unfortunately, measured fluorescence levels can not be directly correlated to expression levels, as phenomena such as protein aggregations can reduce the amount of fluorescence observed from individual proteins.[97] The most likely effect of this is that for high expression levels fluorescence will not be proportional to produced GFP, and it may be important to be aware of this fact when interpreting results. Second to the rrnBp1 promoter in overall GFP production, the LacUV5 and LivJ promoter showed comparable fluorescence levels in both LB and M9 medium.

Measurements of GFP production were also performed in freshly inoculated cultures. These showed the same general pattern with respect to relative promoter strengths as did the overnight cultures. The measurements demonstrated that in rich growth medium GFP production from all the tested promoters is detectable during growth in the exponential phase.

An experiment where (p)ppGpp was accumulated by expressing a ppGpp synthase showed that the *argI* and *iraP* promoters were positively regulated by this intervention, according to expectations. However, for both promoters both the overall activity levels, and the absolute changes in expression, were low. The rrnBp1 promoter again showed large expression, and a large effect from (p)ppGpp inhibition, giving a much larger signal separation than for the previous two promoters. For the LacUV5 promoter, no obvious response to (p)ppGpp was observed, again as expected for this constitutive promoter.

Observing these results, it seems clear that in addition to the relative, fold-change effect of (p)ppGpp on promoter activity, the overall promoter activity levels may also be important for the choice of suitable promoter for expressing the reporter gene. The *livJ* promoter was included in the (p)ppGpp induction experiment, but showed overall higher activity levels in the previous experiments, comparable to the constitutive LacUV5 promoter. Going forward, the livJ promoter may thus be a better candidate and seems worth investigating in the same way.

The effect of the LVA degradation tag on GFP stability was assayed by stopping protein translation with chloramphenicol and measuring fluorescence levels over 30 minutes. Keeping with the expectation, fluorescence from the LVA-tagged variant decreased during the experiment. However, this decrease was less than that previously reported by Anderson et al. In the cell culture expressing the stable GFP variant, fluorescence was observed to increase even after chloramphenicol treatment. This may plausibly be explained by continued maturation of the fluorescent protein fluorophore after cessation of translation, imperfect halting of protein translation, or a combination. In the case that continued protein maturation was the major contributor to the increase in fluorescence, this may also serve to explain the slower than expected decay in fluorescence from the GFP-LVA variant. Further experiments may be in order to elucidate the kinetics of the specific GFP variant under the relevant conditions for later experiments where the reporters are to be employed.

SLIC failure: Attempts at cloning using SLIC were unsuccessful. This might have to do with the small size of the DNA fragments employed. Excessive exonuclease activity is a possible reason, and solutions might seek to decrease the exonuclease activity of the DNA polymerase used in the method. Decreasing the reaction time in the exonuclease digestion step. It is unclear whether quenching the reaction on ice and adding a nucleotide have comparable effects. As SLIC offers the potential for substantial time-savings for routine cloning, it might be of some interest to develop a protocol for use of the method with short fragments, if this is possible.

It is also a possibility that stable single-stranded DNA structures in the plasmid may interfere with annealing of the complementary overhangs. In this case, the issue would be intrinsic to the vector itself. Obtaining a working protocol for SLIC with the current vector used here would allow quicker construction of additional reporters, and so might be worh pursuing. Whether size is the major issue may be investigated by repeating the cloning procedure using larger DNA fragments.

Possibility of undesired mutations: It is known that modified strains which over-produce (p)ppGpp are prone to reversal mutations. As ppGpp inhibits growth, a selective pressure to decrease any excessive alarmone production may exist, especially in fast-growing populations. Here,(p)ppGpp was overproduced by inducing expression of the ppGpp synthase RelA from a plasmid, using the promoter pBAD and arabinose as inducer. Although pBAD is nominally a tight promoter, giving little expression in its un-induced state, compensatory mutations either on the plasmid or in the chromosome of transformed bacteria can not be ruled out. No special precautions were

taken here, but for rigorousness in future experiments, all (p)ppGpp overproducing cell lines should be screened with appropriate methods and efforts taken to preserve the integrity of the desired genotypes.

The strains were obtained and handled without obvious issues, but their identity was not confirmed independently. Prudence suggests that received materials be verified before use. For strains with known genetic modifications, screening for known inserts by PCR may help dispel doubts.

As (p)ppGpp decreases growth, cells not subject to its effects may be expected to outgrow those that are, and by time overtake the population in any mixed culture. Levels of (p)ppGpp have been inversely linked to growth rate.[118] If the difference in (p)ppGpp levels is large, we might expect this overtaking to happen rapidly. However, if the production of (p)ppGpp in induced cells is moderate, we might see two (for the simplest case, assuming cells are either induced or non-induced, and that all induced cells are induced at the same level) sub-populations growing at somewhat different rates.

As such, the moderate reduction in growth observed here at the bulk cell population level, could be due to a mix of faster-growing and slower-growing cells. In a potential use case for FACS, cells might be sorted by fluorescence as indication of (p)ppGpp production, followed by characterisation of the growth rates of the individual sub-populations. Mathematical modelling could also be employed to calculate the growth curves one would be likely to observe for different scenarios involving different levels and proportions of induction or arrest of growth rate. As (p)ppGpp has wide-reaching effects on cellular metabolism, a systems biology perspective with modelling of the gross observed parameters coupled with single-cell measurements to capture the variability in the cell population might lead to further insights.

Biological considerations The plasmids pUM9 and pUM66 were recently used in a study of the differential regulation of pppGpp and ppGpp.[72] In this work, plasmid pUM9 was used exclusively to express a RelSeq (p)ppGpp synthase fragment, with ppGpp expected to dominate over pppGpp as a result. As the pUM66 plasmid favors accumulation of pppGpp, repeating the experiments described here using this plasmd may be worthwhile to assess the differences in outcome, if any.

The strain BW27784 was used to facilitate homogeneous induction of ppGpp synthase production in the bacterial population. A closely related strain, BW277846, was also available but was not used. The two strains differ only in the constitutive promoter used to drive expression of the arabinose transporter, and BW27784 contains the stronger promoter of the two. It may be interesting to observe if there would be differences. However, differences

might not be apparent on the bulk cell population level. Measurements at the single-cell level, for example by flow cytometry could potentially resolve such differences.

The majority of the experiments were carried out using a rich growth medium (LB). However, environmental factors and culturing regimes affect the physiology and gene expression of the cultured organism. As such, it would be desirable to repeat the described experiments using different growth media and culturing conditions, to explore variability in responses and test specific hypotheses. For example, for the promoters belonging to the amino acid biosynthesis operons, the effect of including various species and levels of amino acids in the growth medium would be of interest. In particular, characterization of the reporter systems should be carried out in similar physiological conditions as those in which the reporting system will later be used. If the goal is to correlate expression levels with ppGpp and/or pppGpp levels, it will be necessary to perform isolation and direct measurements of the nucleotides. As an additional factor to consider, exponential growth and planctonic growth in monoculture are convenient settings to experiment in, but may not be the most relevant for the ultimate issue. Growth in biofilms may be of more interest, and fluorescent reporters have previously been used in this context.[103, 119].

Technical considerations: In addition to the choice of biological parameters (growth medium, culturing conditions, etc.), the potential for variations due to the choice of measuring protocol should also be considered. To some degree, the two issues overlap when measurements are performed on cell cultures incubated in a microplate reader. The culturing conditions are then in part defined by the same parameters which are of importance for the measurement process (sample volume in well, cover on or off, shaking, etc). The use of 96-well plates theoretically allows for performing parallel measurements on 32 separate cultures in a single experiment, assuming that three wells are used per culture. If use of the framing wells is avoided to reduce sample loss by evaporation, this number is reduced to 20. While this is still a considerable number, sample handling times must also be taking into account when planning an experiment. It was found that diluting overnight culture samples and transferring the resulting diluted samples to a microwell plate in triplicate by manual pipetting required close to 20 minutes. Excessive handling time may potentially introduce a systematic bias. If experiments with long sample handling times are to be carried out, investigating the relevance effect may be prudent. The use of automatic sample and plate handling systems might facilitate reducing sample handling time and between-sample variability stemming from this. However, commercial systems for automatic

sample handling are expensive and may not be justifi. It is conceivable that customized hardware and software could be constructed to perform. [120] As an example of the potential of "do it yourself" hardware, constructed or assembled by the end-user, to reduce costs of routine laboratory operations, an openPCR thermocycler assembled from a kit purchased for $599 was successfully used for several PCR amplifications in the course of this project.

Possibilities for further development: Plasmids will generally display some copy number variability, translating to gene copy variability. For bulk cell populations, these differences may be expected to average out, but at the single-cell level, variability may become apparent. Modelling of genetic regulation and expression also becomes more difficult when there is an unknown number of gene copies present. Thus, it can be attractive to fix the gene copy number, by integrating the desired sequence into the bacterial chromosome - methods for this are readily available. Chromosomal integration can also remove or reduce the need for selective markers, thus potentially allowing more "natural" growth conditions. However, chromosomal integration requires substantially more work than working with plasmids, which also have the advantage of being easily transformed into a variety of strains. As such, use of chromosomal integration may be more appropriate for a after initial work with a larger variety of constructs hosted on plasmids. Preparing BioBrick-compatible vectors bearing promoter sequences alone or linked to a fluorescent reporter only may give increased re-use opportunities and allow utilization of the growing number of tools applicable to BioBrick-compatible parts, for example for chromosomal integration.[121]

Finally, a potential major issue with using fluorescent reporters for monitoring stringent response is that the machinery driving the reporter system - that is, the transcriptional and translational system - is also majorly affected by the stringent response and its regulators. This may make it hard to discern global effects on gene expression from specific effects. For example, knowledge about the changes in the expression in one gene will be more informative if the general status of the cell is known. This challenge could be approached by slightly more advanced reporter designs, incorporating several genetic elements. For example, a basic combination design would employ one fluorescent protein expressed by a constitutive promoter, and another fluorescent reporter regulated by a stringent response-associated pathway or regulator. The constitutive promoter would then act as an internal control and a reference to the overall transcription/translation activity of the cell.

6 Conclusion

Fluorescent reporter plasmids for use in investigating the stringent response have been constructed using several promoters with different responses to (p)ppGpp levels. The overall, relative promoter strengths and the effect on transcription by overproduction of (p)ppGpp were investigated. Consistent with what has previously been described in the literature, promoters associated with amino acid biosynthesis genes appeared to be activated by elevated (p)ppGpp levels, while the ribosomal RNA promoter rrnBp1 is inhibited. Elevation of (p)ppGpp levels was accomplished by expressing a (p)ppGpp synthase from a separate plasmid. The decrease in growth rate upon induction of the (p)ppGpp synthase was moderate, while an increase in per-cell fluorescence could be observed for the *argI* and *iraP* promoters. However, general activity levels of these two promoters were low, and the relative difference between the induced- and non-induced states small. Regulation of GFP production by the ribosomal promoter rrnBp1, having a much larger general activity level, gave a much larger relative difference. If positive regulation by ppGpp and a large relative difference in observed fluorescence between high-ppGpp-and low-ppGpp states are desired, the *livJ* promoter, showing a higher general activity level, may be a better candidate.

Cloning by the recently developed SLIC method was attempted, but failed to yield any transformants. It is suggested that this may be due to short length of the sequences used, but properties of the linking sequences could also be a factor. As SLIC has potential for large time savings compared to conventional cloning, it may be of interest to pursue usability and reproducibility of the method also for very short (approximately 100 bp) sequences.

Some suggestions for further work have been given. In particular, the use of single-cell measurements by methods such as flow cytometry appear important to gain insight into population variability. Additionally, the reporter constructs might be extended or combined to give multi-part systems with separate promoters driving expression of separate fluorescent proteins, enabling read-out of several signals at once.

References

[1] Rolf Wagner. *ppGpp Signaling*, pages 395–414. Wiley-VCH Verlag GmbH & Co. KGaA, 2010.

[2] StephanieM. Amato, MehmetA. Orman, and MarkP. Brynildsen. Metabolic control of persister formation in *Escherichia coli*. *Molecular Cell*, 50(4):475 – 487, 2013.

[3] Zachary D. Dalebroux, Sarah L. Svensson, Erin C. Gaynor, and Michele S. Swanson. ppGpp conjures bacterial virulence. *Microbiology and Molecular Biology Reviews*, 74(2):171–199, 2010.

[4] Ezequiel Wexselblatt, Yaara Oppenheimer-Shaanan, Ilana Kaspy, et al. Relacin, a novel antibacterial agent targeting the stringent response. *PLoS Pathog*, 8(9):e1002925, 09 2012.

[5] Undine Mechold, Helen Murphy, Larissa Brown, and Michael Cashel. Intramolecular regulation of the opposing (p)ppGpp catalytic activities of RelSeq, the Rel/Spo enzyme from *Streptococcus equisimilis*. *Journal of Bacteriology*, 184(11):2878–2888, 2002.

[6] KOZO OCHI. Occurrence of the stringent response in streptomyces sp. and its significance for the initiation of morphological and physiological differentiation. *Journal of General Microbiology*, 132(9):2621–2631, 1986.

[7] Hyun-Woo Rhee, Chang-Ro Lee, Seung-Hyon Cho, et al. Selective fluorescent chemosensor for the bacterial alarmone (p)ppgpp. *Journal of the American Chemical Society*, 130(3):784–785, 2008.

[8] Xi Juan Zhao and Cheng Zhi Huang. Selective fluorometric detection of pyrophosphate and stringent alarmone with copper(ii)–2,6-bis(2-benzimidazolyl)pyridine complex. *Biosensors and Bioelectronics*, 30(1):282 – 286, 2011.

[9] Pu Zhang, Yi Wang, Yong Chang, Zu Hong Xiong, and Cheng Zhi Huang. Highly selective detection of bacterial alarmone ppgpp with an off–on fluorescent probe of copper-mediated silver nanoclusters. *Biosensors and Bioelectronics*, 49(0):433 – 437, 2013.

[10] Hidde de Jong, Caroline Ranquet, Delphine Ropers, Corinne Pinel, and Johannes Geiselmann. Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. *BMC Systems Biology*, 4(1):55, 2010.

[11] JD Watson and FH Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171:737–8, 1953.

[12] Oswald T Avery, Colin M MacLeod, and Maclyn McCarty. Studies on the chemical nature of the substance inducing transformation of pneumococcal types : Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type iii. *Journal of Experimental Medicine*, 79:137–158, 1944.

[13] F. Sanger and A.R. Coulson. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94(3):441 – 448, 1975.

[14] Cheuk-Wai Kan, Christopher P. Fredlake, Erin A. S. Doherty, and Annelise E. Barron. DNA sequencing and genotyping in miniaturized electrophoresis systems. *Electrophoresis*, 25(21-22):3564–3588, 2004.

[15] Martin Thanbichler and Lucy Shapiro. Chromosome organization and segregation in bacteria. *Journal of Structural Biology*, 156(2):292 – 303, 2006.

[16] Gloria del Solar, Rafael Giraldo, María Jesús Ruiz-Echevarría, Manuel Espinosa, and Ramón Díaz-Orejas. Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews*, 62(2):434–464, 1998.

[17] Neil Willetts. The genetics of transmissible plasmids. *Annu. Rev. Genet.*, 6:257–268, 1972.

[18] Naoto Yoshida and Misa Sato. Plasmid uptake by bacteria: a comparison of methods and efficiencies. *Applied Microbiology and Biotechnology*, 83(5):791–798, 2009.

[19] Fabian Svara and Daniel Rankin. The evolution of plasmid-carried antibiotic resistance. *BMC Evolutionary Biology*, 11(1):130, 2011.

[20] Luca Gerosa, Karl Kochanowski, Matthias Heinemann, and Uwe Sauer. Dissecting specific and global transcriptional regulation of bacterial gene expression. *Molecular Systems Biology*, 9:658, 2013.

[21] Tanja M. Gruber and Carol A. Gross. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology*, 57(1):441–466, 2003. PMID: 14527287.

[22] Haiyuan Yu and Mark Gerstein. Genomic analysis of the hierarchical structure of regulatory networks. *Proceedings of the National Academy of SciencesYu2006*, 103(40):14724–14731, 2006.

[23] David L Nelson and Michael M Cox. *Principles of Biochemistry*. W. H. Freeman, 2008.

[24] William S. Klug, Michael R. Cummings, Charlotte Spencer, Michael A. Palladino, and Sarah M. Ward. *Concepts of Genetics*. Pearson Education, 2009.

[25] Katarzyna Potrykus, Helen Murphy, Nadege Philippe, and Michael Cashel. ppGpp is the major source of growth rate control in *E. coli*. *Environmental Microbiology*, 13(3):563–575, 2011.

[26] Zachary D Dalebroux and Michelse S Swanson. ppGpp: magic beyond RNA polymerase. *Nature Reviews Microbiology*, 10:203–212, 2012.

[27] Mike Cashel and Jonathan Gallant. Two compounds implicated in the function of the RC gene of *Escherichia coli*. *Nature*, 221:838–841, 1969.

[28] Michael Cashel. Regulation of bacterial ppGpp and pppGpp. *Annual Review of Microbiology*, 29:301–318, 1975.

[29] Matthew F. Traxler, Sean M. Summers, Huyen-Tran Nguyen, et al. The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Molecular Microbiology*, 68(5):1128–1148, 2008.

[30] Gerhard Mittenhuber. Comparative genomics and evolution of genes encoding bacterial (p)ppGpp synthetases/hydrolases (the Rel, RelA and SpoT proteins). *J Mol Microbiol Biotechnol*, 3:585–600, 2001.

[31] Jun Wu and Jianping Xie. Magic spot: (p) ppGpp. *Journal of Cellular Physiology*, 220(2):297–302, 2009.

[32] Vikas Jain, Manish Kumar, and Dipankar Chatterji. ppGpp: Stringent response and survival. *The Journal of Microbiology*, 44:1–10, 2006.

[33] Dimpy Kalia, Gokce Merey, Shizuka Nakayama, et al. Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chem. Soc. Rev.*, 42:305–341, 2013.

[34] Cara C. Boutte and Sean Crosson. Bacterial lifestyle shapes stringent response activation. *Trends in Microbiology*, 21(4):174 – 180, 2013.

[35] Gemma C. Atkinson and Sandra L. Baldauf. Evolution of elongation factor G and the origins of mitochondrial and chloroplast forms. *Molecular Biology and Evolution*, 28(3):1281–1292, 2011.

[36] Thomas M Wendrich, Gregor Blaha, Daniel N Wilson, Mohamed A Marahiel, and Knud H Nierhaus. Dissection of the mechanism for the stringent factor RelA. *Molecular Cell*, 10(4):779 – 788, 2002.

[37] Viktoriya Shyp, Stoyan Tankov, Andrey Ermakov, et al. Positive allosteric feedback regulation of the stringent response enzyme RelA by its product. *EMBO reports*, 13:835–839, 2012.

[38] Karsten Tedin and Hans Bremer. Toxic effects of high levels of ppGpp in *Escherichia coli* are relieved by rpoB mutations. *The Journal of Biological Chemistry*, 267:2337–2344, 1992.

[39] Gemma C Atkinson, Tanel Tenson, and Vasili Hauryliuk. The RelA/SpoT homolog (rsh) superfamily: Distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLos ONE*, 6:e23479, 2011.

[40] G S Johnson, C R Adler, J J Collins, and D Court. Role of the spoT gene product and manganese ion in the metabolism of guanosine 5'-diphosphate 3'-diphosphate in *Escherichia coli*. *Journal of Biological Chemistry*, 254(12):5483–7, 1979.

[41] Beny Spira, Xuye Hu, and Thomas Ferenci. Strain variation in ppGpp concentration and RpoS levels in laboratory strains of *Escherichia coli* K-12. *Microbiology*, 154(9):2887–2895, 2008.

[42] Harald Weber, Tino Polen, Johanna Heuveling, Volker F. Wendisch, and Regine Hengge. Genome-wide analysis of the general stress response network in *Escherichia coli*: Sigma s-dependent genes, promoters, and sigma factor selectivity. *Journal of Bacteriology*, 187(5):1591–1603, 2005.

[43] H Xiao, M Kalman, K Ikehara, et al. Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. *Journal of Biological Chemistry*, 266(9):5980–5990, 1991.

[44] Melanie M. Barker, Tamas Gaal, Cathleen A. Josaitis, and Richard L. Gourse. Mechanism of regulation of transcription initiation by ppGpp. i. effects of ppGpp on transcription initiation *in vivo* and *in vitro*. *Journal of Molecular Biology*, 305(4):673 – 688, 2001.

[45] PC Loewen, B Hu, J Strutinsky, and R Sparling. Regulation in the rpoS regulon of *Escherichia coli*. *Journal of Microbiology*, 44:707–717, 1998.

[46] D R Gentry, V J Hernandez, L H Nguyen, D B Jensen, and M Cashel. Synthesis of the stationary-phase sigma factor sigma S is positively regulated by ppGpp. *Journal of Bacteriology*, 175(24):7982–7989, 1993.

[47] Matthew Hirsch and Thomas Elliott. Role of ppGpp in rpoS stationary-phase regulation in *Escherichia coli*. *Journal of Bacteriology*, 184(18):5077–5087, 2002.

[48] Alexandre Bougdour, Christofer Cunning, Patrick Jean Baptiste, Thomas Elliott, and Susan Gottesman. Multiple pathways for regulation of RpoS stability in *Escherichia coli* via the action of multiple anti-adaptors. *Molecular Microbiology*, 68(2):298–313, 2008.

[49] Alexandre Bougdour and Susan Gottesman. ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. *Proceedings of the National Academy of Sciences*, 104(31):12896–12901, 2007.

[50] Houra Merrikh, Alexander E Ferrazoli, and Susant T Lovett. Growth phase and (p)ppgpp control of IraD, a regulator of RpoS stability, in *Escherichia coli*. *Journal of Bacteriology*, 191:7436–7446, 2009.

[51] Larissa Brown, Daniel Gentry, Thomas Elliott, and Michael Cashel. DksA affects ppGpp induction of RpoS at a translational level. *Journal of Bacteriology*, 184(16):4455–4465, 2002.

[52] Patrick P. Dennis, Mans Ehrenberg, and Hans Bremer. Control of rRNA synthesis in *Escherichia coli*: a systems biology approach. *Microbiology and Molecular Biology Reviews*, 68(4):639–668, 2004.

[53] David A. Schneider and Richard L. Gourse. Changes in the concentrations of guanosine 5-diphosphate 3-diphosphate and the initiating nucleoside triphosphate account for inhibition of rRNA transcription in fructose-1,6-diphosphate aldolase (fda) mutants. *Journal of Bacteriology*, 185(20):6192–6194, 2003.

[54] Huijun Zhi, Xiangdong Wang, Julio E. Cabrera, Reid C. Johnson, and Ding Jun Jin. Fis stabilizes the interaction between RNA polymerase and the ribosomal promoter rrnB P1, leading to transcriptional activation. *Journal of Biological Chemistry*, 278(47):47340–47349, 2003.

[55] Remus Thei Dame, Claire Wyman, Reinhild Wurm, Rolf Wagner, and Nora Goosen. Structural basis for H-NS-mediated trapping of RNA Polymerase in the open initiation complex at the rrnb P1. *Journal of Biological Chemistry*, 277(3):2146–2150, 2002.

[56] X Zhang, P Dennis, M Ehrenberg, and H Bremer. Kinetic properties of rrn promoters i *Escherichia coli*. *Biochimie*, 84(10):981 – 996, 2002.

[57] P. P. Dennis, M. Ehrenberg, D. Fange, and H. Bremer. Varying rate of RNA chain elongation during rrn transcription in *Escherichia coli*. *Journal of Bacteriology*, 191:3740–3746, 2009.

[58] Ulla Vogel and Kaj Frank Jensen. Effects of the antiterminator BoxA on transcription elongation kinetics and ppGpp inhibition of transcription elongation in *Escherichia coli*. *Journal of Biological Chemistry*, 270(31):18335–18340, 1995.

[59] C L Squires, J Greenblatt, J Li, C Condon, and C L Squires. Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. *Proceedings of the National Academy of Sciences*, 90(3):970–974, 1993.

[60] Stefan Klumpp and Terence Hwa. Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. *Proceedings of the National Academy of Sciences*, 105(47):18159–18164, 2008.

[61] T Heinrich, C Condon, T Pfeiffer, and R K Hartmann. Point mutations in the leader boxA of a plasmid-encoded *Escherichia coli* rrnB operon cause defective antitermination *in vivo*. *Journal of Bacteriology*, 177(13):3793–800, 1995.

[62] Hans Bremer and Patrick P Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. In *EcoSal - Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, 2008.

[63] U Vogel and K F Jensen. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *Journal of Bacteriology*, 176(10):2807–2813, 1994.

[64] Bertil Gummesson, Lisa U Magnusson, Martin Lovmar, et al. Increased RNA polymerase availability directs resources towards growth at the expense of maintenance. *The EMBO Journal*, 28:2209 – 2219, 2009.

[65] Tim Durfee, Anne-Marie Hansen, Huijun Zhi, Frederick R. Blattner, and Ding Jun Jin. Transcription profiling of the stringent response in *Escherichia coli*. *Journal of Bacteriology*, 190:1084–1096, 2008.

[66] BJ Paul, MB Berkmen, and RL Gourse. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci U S A*, 102:7823–8, 2005.

[67] P J Kang and E A Craig. Identification and characterization of a new *Escherichia coli* gene that is a dosage-dependent suppressor of a dnaK deletion mutation. *Journal of Bacteriology*, 172(4):2055–2064, 1990.

[68] Brian J. Paul, Melanie M. Barker, Wilma Ross, et al. DksA: A critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, 118(3):311 – 322, 2004.

[69] Lisa U. Magnusson, Bertil Gummesson, Predrag Joksimovi, Anne Farewell, and Thomas Nystrm. Identical, independent, and opposing roles of ppGpp and DksA in *Escherichia coli*. *Journal of Bacteriology*, 189:5193–5202, 2007.

[70] Pete Chandrangsu, Justin J. Lemke, and Richard L. Gourse. The dksA promoter is negatively feedback regulated by DksA and ppGpp. *Molecular Microbiology*, 80(5):1337–1348, 2011.

[71] Bertil Gummesson, Martin Lovmar, and Thomas Nyström. A proximal promoter element required for positive transcriptional control by ppGpp and DksA during the stringent response. *Journal of Biological Chemistry*, 2013.

[72] U Mechold, K Potrykus, H Murphy, KS Murakami, and M Cashel. Differential regulation by ppGpp versus pppGpp in *Escherichia coli*. *Nucleic Acids Research*, 41(12):6175–89, 2013.

[73] Christine A. Hirvonen, Wilma Ross, Christopher E. Wozniak, et al. Contributions of UP elements and the transcription factor FIS to expression from the seven rrn P1 promoters in *Escherichia coli*. *Journal of Bacteriology*, 183(21):6305–6314, 2001.

[74] Karsten Tedin, Angela Witte, Gunter Reisinger, Werner Lubitz, and Udo Blasi. Evaluation of the *Escherichia coli* ribosomal rrnB P1 promoter and phage-derived lysis genes for the use in a biological containment system: A concept study. *Journal of Biotechnology*, 39(2):137 – 148, 1995.

[75] Xiangyang Zhang, Sung-Tzu Liang, and Hans Bremer. Feedback control ofribosome synthesis in *Escherichia coli* is dependent oneightcritical amino acids. *Biochimie*, 88(9):1145 – 1155, 2006.

[76] C A Josaitis, T Gaal, and R L Gourse. Stringent control and growth-rate-dependent control have nonidentical promoter sequence requirements. *Proceedings of the National Academy of Sciences*, 92(4):1117–1121, 1995.

[77] Richard J. Noel and William S. Reznikoff. Structural studies of lacUV5-RNA polymerase interactions *in Vitro* : Ethylation interference and missing nucleoside analysis. *Journal of Biological Chemistry*, 275(11):7708–7712, 2000.

[78] Mitchell Lewis. The *lac* repressor. *Comptes Rendus Biologies*, 328(6):521 – 548, 2005.

[79] Ingrid M. Keseler, Amanda Mackie, Martin Peralta-Gil, et al. EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Research*, Database issue, 2012.

[80] Katarzyna Potrykus, Daniel Vinella, Helen Murphy, et al. Antagonistic regulation of *Escherichia coli* ribosomal RNA rrnB P1 promoter activity by GreA and DksA. *Journal of Biological Chemistry*, 281(22):15238–15248, 2006.

[81] Francine Toulmhristine Mosrin-Huaman, Jason Sparkowski, et al. GreA and greB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. *The EMBO Journal*, 19:6863–6859, 2000.

[82] K Potrykus, H Murphy, JA Epstein, and M Cashel. Imprecise transcription termination within *Escherichia coli* greA leader gives rise to an array of short transcripts, graL. *Nucleic Acids Research*, 38:1636–51, 2010.

[83] Luciana Gualdi, Letizia Tagliabue, and Paolo Landini. Biofilm formation-gene expression relay system in *Escherichia coli*: Modulation of sigma s-dependent gene expression by the CsgD regulatory protein via sigma s protein stabilization. *Journal of Bacteriology*, 189(22):8034–8043, 2007.

[84] C Legrain, V Stalon, and N Glansdorff. Escherichia coli ornithine carbamolytransferase isoenzymes: evolutionary significance and the isolation of lambdaargF and lambdaargI transducing bacteriophages. *Journal of Bacteriology*, 128(1):35–38, 1976.

[85] M D Adams, L M Wagner, T J Graddis, et al. Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli*. *Journal of Biological Chemistry*, 265(20):11436–11443, 1990.

[86] Michael R Green and Joseph Sambrook. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2012.

[87] Mamie Z Li and Stephen J Elledge. Harnessing homologous recombination *in vitro* to generate recombinant dna via SLIC. *Nature Methods*, 4:251–256, 2007.

[88] Nathan J. Hillson. *Design and Analysis of Biomolecular Circuits:*, chapter DNA Assembly Method Standardization for Synthetic Biomolecular Circuits and Systems, pages 295–314. Springer Verlag, 2011.

[89] Judith Scholz, Huseyin Besir, Claudia Strasser, and Sabine Suppmann. A new method to customize protein expression vectors for fast, efficient and background free parallel cloning. *BMC Biotechnology*, 13(1):12, 2013.

[90] Guy Cox. *Optical Imaging Techniques in Cell Biology*. CRC Press, 2012.

[91] Osamu Shimomura, Frank H. Johnson, and Yo Saiga. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239, 1962.

[92] JC March, G Rao, and WE Bentley. Biotechnological applications of green fluorescent protein. *Appl. Microbiol. Biotechnol.*, 62:303–315, 2003.

[93] Martin C. Hansen, Robert J. Palmer, Camilla Udsen, David C. White, and Søren Molin. Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. *Microbiology*, 147(5):1383–1391, 2001.

[94] R Heim, D C Prasher, and R Y Tsien. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences*, 91(26):12501–12504, 1994.

[95] Brendan P. Cormack, Raphael H. Valdivia, and Stanley Falkow. Facs-optimized mutants of the green fluorescent protein (gfp). *Gene*, 173(1):33 – 38, 1996.

[96] Canghai Lu, C. Renee Albano, William E. Bentley, and Govind Rao. Differential rates of gene expression monitored by green fluorescent protein. *Biotechnology and Bioengineering*, 79(4):429–437, 2002.

[97] Spela Peternel, Vladka Gaberc-Porekar, and Radovan Komel. Bacterial growth conditions affect quality of GFP expressed inside inclusion bodies. *Acta Chim. Slov. 56,*, 56:860867, 2009.

[98] Jens Bo Andersen, Claus Sternberg, Lars Kongsbak Poulsen, et al. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and Environmental Microbiology*, 64(6):2240–2246, 1998.

[99] Kenneth C Keiler, Patrick RH Waller, and Robert T Sauer. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, 271:990–993, 1996.

[100] Susan Gottesman, Eric Roche, YanNing Zhou, and Robert T. Sauer. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & Development*, 12(9):1338–1347, 1998.

[101] David A Wah, Igor Levchenko, Gabrielle E Rieckhof, et al. Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the aaa+ clpXP protease. *Molecular Cell*, 12(2):355 – 363, 2003.

[102] Oliver Purcell, Claire Grierson, Mario Bernardo, and Nigel Savery. Temperature dependence of ssrA-tag mediated protein degradation. *Journal of Biological Engineering*, 6(1):10, 2012.

[103] Claus Sternberg, Bjarke B. Christensen, Tove Johansen, et al. Distribution of bacterial growth activity in flow-chamber biofilms. *Applied and Environmental Microbiology*, 65(9):4108–4117, 1999.

[104] M Ormo, AB Cubitt, K Kallio, et al. Crystal structure of the *Aequorea victoria* Green Fluorescent Protein. *Science*, 273(528):1392–1395, 1996.

[105] J. Alex Appleman, Wilma Ross, Julia Salomon, and Richard L. Gourse. Activation of *Escherichia coli* rRNA transcription by FIS during a growth cycle. *Journal of Bacteriology*, 180(6):1525–1532, 1998.

[106] V J Hernandez and H Bremer. Characterization of RNA and DNA synthesis in *Escherichia coli* strains devoid of ppGpp. *Journal of Biological Chemistry*, 268(15):10851–62, 1993.

[107] Paul Primakoff and Stanley W. Artz. Positive control of *lac* operon expression *in vitro* by guanosine 5-diphosphate 3-diphosphate. *Proceedings of the National Academy of Sciences*, 76(4):1726–1730, 1979.

[108] Dietrich Stueber and Hermann Bujard. Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. *The EMBO Journal*, 1:1399–1404, 1982.

[109] Simone Balzer, Veronika Kucharova, Judith Megerle, et al. A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in *Escherichia coli*. *Microbial Cell Factories*, 12(1):26, 2013.

[110] David H. Figurski, Richard J. Meyer, and Donald R. Helinski. Suppression of colE1 replication properties by the inc p-1 plasmid RK2 in hybrid plasmids constructed in vitro. *Journal of Molecular Biology*, 133(3):295 – 318, 1979.

[111] Bethesda Research Laboratories. BRL pUC host: *Escherichia coli* DH5α competent cells. *Focus*, 8(2):9, 1986.

[112] Artem Khlebnikov, Kirill A. Datsenko, Tove Skaug, Barry L. Wanner, and Jay D. Keasling. Homogeneous expression of the PBAD promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology*, 147(12):3241–3247, 2001.

[113] L M Guzman, D Belin, M J Carson, and J Beckwith. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, 177(14):4121–30, 1995.

[114] Deborah A. Siegele and James C. Hu. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proceedings of the National Academy of Sciences*, 94(15):8168–8172, 1997.

[115] Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403 – 410, 1990.

[116] C T Chung, S L Niemela, and R H Miller. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences*, 86(7):2172–2175, 1989.

[117] Jae-Yeon Jeong, Hyung-Soon Yim, Ji-Young Ryu, et al. One-step sequence- and ligation-independent cloning (SLIC): Rapid and versatile cloning method for functional genomics studies. *Applied and Environmental Microbiology*, 2012.

[118] Danièle Joseleau-Petit, Danielle Thévenet, and Richard D'Arl. ppGpp concentration, growth without PBP2 activity, and growth-rate control in *Escherichia coli*. *Molecular Microbiology*, 13(5):911–917, 1994.

[119] Sren Mller, Claus Sternberg, Jens Bo Andersen, et al. In situ gene expression in mixed-culture biofilms: Evidence of metabolic interactions between community members. *Appl Environ Microbiol*, 64:721732., 1998.

[120] Joshua M. Pearce. Building research equipment with free, open-source hardware. *Science*, 337(6100):1303–1304, 2012.

[121] Susanna Zucca, Lorenzo Pasotti, Nicolo Politi, Maria Cusella De Angelis, and Paolo Magni. A standard vector for the chromosomal integration and characterization of BioBrick parts in *Escherichia coli*. *Journal of Biological Engineering*, 7(1):12, 2013.

A Nucleotide sequences

pSB-M1g sequence: Plasmid pSB-M1g was used for construction of the novel reporter constructs described in this report. The nucleotide sequence of pSB-M1g according to records at the NTNU Department of Biotechnology is shown below, with additional annotations. pSB-M1g was prepared by Simone Balzer.[109]

```
LOCUS       pSB-M1g      7828 bp    DNA   CIRCULAR SYN       26-MAY-2013
DEFINITION  Ligation of Fragment 2 into GFP pTA16
ACCESSION   pSB-M1g
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown
            Unclassified.
REFERENCE   1  (bases 1 to 7828)
  AUTHORS   Self
  JOURNAL   Unpublished.
COMMENT     SECID/File created by SciEd Central, Scientific & Educational Software
FEATURES             Location/Qualifiers
     CDS             2..724
                     /gene="GFPmut3"
     misc_feature    100..119
                     /gene="COPCRREV"
                     /product="Binding site, Colony PCR REV primer"
     misc_feature    230..249
                     /gene="GFP-END-FWD"
                     /product="Binding site for GFP-END-FWD"
     misc_feature    652..669
                     /gene="Seq6"
                     /product="Sequencing primer binding site"
     misc_feature    701..721
                     /gene="GFP-END-REV"
                     /product="Binding site for GFP-END-REV primer"
     misc_feature    721
                     /gene="Stop"
     misc_feature    complement (1041..1191)
                     /gene="t"
                     /product="transkripsjonsterminator"
     CDS             complement (1385..2534)
                     /gene="trfA"
                     /product="trfA"
     misc_feature    complement (2704)
                     /gene="Pneo"
                     /product="Pneo"
     CDS             complement (2932..3457)
                     /gene="oriV"
                     /product="oriV"
     CDS             complement (4303..5082)
                     /gene="'kan"
                     /product="kan"
     CDS             complement (5085..5118)
                     /gene="kan'"
                     /product="Kanamycin resistance gene"
     misc_feature    complement (5434..5499)
                     /gene="oriT"
                     /product="oriT for conjugal transfer"
     misc_feature    5641..5660
                     /gene="COPCRFWD2"
                     /product="Binding Site for Colony PCR FWD 2 primer"
     misc_feature    5772..5789
                     /gene="Seq5"
                     /product="Binding site for pJP-1 Seq5 sequencing primer (Promoter sequence)"
     misc_feature    5970..5989
                     /gene="COPCR1FWD"
                     /product="Binding Site, Colony PCR FWD priemr"
     misc_feature    5973..5990
                     /gene="Seq5B"
                     /product="Sequencing primer suggested by CM"
     misc_feature    5981..6000
                     /gene="pSB-REV2"
                     /product="Binding site for pSB-REV2 primer"
     misc_feature    5995..6014
                     /gene="SLIC Linker B"
     CDS             complement (6039..7004)
```

```
                     /gene="xylS"
                     /product="gene for regulatory transcription protein XylS"
     misc_feature    7145..7289
                     /gene="rrnBT1T2"
                     /product="terminator"
     misc_feature    7764..7801
                     /gene="Pm"
                     /product="Pm promoter from the P. putida TOL plasmid pWWO"
     misc_feature    7800..7819
                     /gene="SLIC Linker A"
     misc_feature    7816..7820
                     /gene="RBS"
BASE COUNT     1927 a   2022 c   1971 g   1908 t
ORIGIN
        1 TATGGTACCA AGTAAAGGAG AAGAACTTTT CACTGGAGTT GTCCCAATTC TTGTTGAATT
       61 AGATGGTGAT GTTAATGGGC ACAAATTTTC TGTCAGTGGA GAGGGTGAAG GTGATGCAAC
      121 ATACGGAAAA CTTACCCTTA AATTTATTTG CACTACTGGA AAACTACCTG TTCCATGGCC
      181 AACACTTGTC ACTACTTTCG CGTATGGTCT TCAATGCTTT GCGAGATACC CAGATCACAT
      241 GAAGCAGCAT GACTTTTTCA AGAGTGCCAT GCCCGAAGGT TATGTACAGG AAAGAACTAT
      301 ATTTTTCAAA GATGACGGGA ACTACAAGAC ACGTGCTGAA GTCAAGTTTG AAGGTGATAC
      361 CCTTGTTAAT AGAATCGAGT TAAAAGGTAT TGATTTTAAA GAAGATGGAA ACATTCTTGG
      421 ACACAAATTG GAATACAACT ATAACTCACA CAATGTATAC ATCATGGCAG ACAAACAAAA
      481 GAATGGAATC AAAGTTAACT TCAAAATTAG ACACAACATT GAAGATGGAA GCGTTCAACT
      541 AGCAGACCAT TATCAACAAA ATACTCCAAT TGGCGATGGC CCTGTCCTTT TACCAGACAA
      601 CCATTACCTG TCCACACAAT CTGCCCTTTC GAAAGATCCC AACGAAAAGA GAGACCACAT
      661 GGTCCTTCTT GAGTTTGTAA CAGCTGCTGG GATTACACAT GGCATGGATG AACTATACAA
      721 ATAAGGATCC TCTAGCTAGA GTCAGCTTTA TGCTTGTAAA CCGTTTTGTG AAAAAATTTT
      781 TAAAATAAAA AAGGGGACCT CTAGGGTCCC CAATTAATTA GTAATATAAT CTATTAAAGG
      841 TCATTCAAAA GGTCATCCAC CGGATCAGCT TAGTAAAGCC CTCGCTAGAT TTTAATGCGG
      901 ATGTTGCGAT TACTTCGCCA ACTATTGCGA TAACAAGAAA AAGCCAGCCT TTCATGATAT
      961 ATCTCCCAAT TTGTGTAGGG CTTATTATGC ACGCTTAAAA ATAATAAAAG CAGACTTGAC
     1021 CTGATAGTTT GGCTGTGAGC AATTATGTGC TTAGTGCATC TAACGCTTGA GTTAAGCCGC
     1081 GCCGCGAAGC GGCGTCGGCT TGAACGAATT GTTAGACATT ATTTGCCGAC TACCAAGGAT
     1141 CGGGCCTTGA TGTTACCCGA GAGCTTGGCA CCCAGCCTGC GCGAGCAGGG GAATTGATCC
     1201 GGTGGATGAC CTTTTGAATG ACCTTTAATA GATTATATTA CTAATTAATT GGGGACCCTA
     1261 GAGGTCCCCT TTTTTATTTT AAAAATTTTT TCACAAAACG GTTTACAAGC ATAAAGCTGA
     1321 CCCTCTAGCA AGCTTGCGAT GCAGGTGGCT GCTGAACCCC CAGCCGGAAC TGACCCCACA
     1381 AGGCCCTAGC GTTTGCAATG CACCAGGTCA TCATTGACCC AGGCGTGTTC CACCAGGCCG
     1441 CTGCCTCGCA ACTCTTCGCA GGCTTCGCCG ACCTGCTCGC GCCACTTCTT CACGCGGGTG
     1501 GAATCCGATC CGCACATGAG GCGGAAGGTT TCCAGCTTGA GCGGGTACGG CTCCCGGTGC
     1561 GAGCTGAAAT AGTCGAACAT CCGTCGGGCC GTCGGCGACA GCTTGCGGTA CTTCTCCCAT
     1621 GTGAATTTCG TGTAGTGGTC GCCAGCAAAC AGCACGACGA TTTCCTCGTC GATCAGGACC
     1681 TGGCAACGGG ACGTTTTCTT GCCACGGTCC AGGACGCGGA AGCGGTGCAG CAGCGACACC
     1741 GATTCCAGGT GCCCAACGCG GTCGGACGTG AAGCCCATCG CCGTCGCCTG TAGGCGCGAC
     1801 AGGCATTCCT CGGCCTTCGT GTAATACCGG CCATTGATCG ACCAGCCCAG GTCCTGGCAA
     1861 AGCTCGTAGA ACGTGAAGGT GATCGGCTCG CCGATAGGGG TGCGCTTCGC GTACTCCAAC
     1921 ACCTGCTGCC ACACCAGTTC GTCATCGTCG GCCCGCAGCT CGACGCCGGT GTAGGTGATC
     1981 TTCACGTCCT TGTTGACGTG GAAAATGACC TTGTTTTGCA GCGCCTCGCG CGGGATTTTC
     2041 TTGTTGCGCG TGGTGAACAG GGCAGAGCGG GCCGTGTCGT TTGGCATCGC TCGCATCGTG
     2101 TCCGGCCACG GCGCAATATC GAACAAGGAA AGCTGCATTT CCTTGATCTG CTGCTTCGTG
     2161 TGTTTCAGCA ACGCGGCCTG CTTGGCCTCG CTGACCTGTT TTGCCAGGTC CTCGCCGGCG
     2221 GTTTTTCGCT TCTTGGTCGT CATAGTTCCT CGCGTGTCGA TGGTCATCGA CTTCGCCAAA
     2281 CCTGCCGCCT CCTGTTCGAG ACGACGCGAA CGCTCCACGG CGGCCGATGG CGCGGGCAGG
     2341 GCAGGGGGAG CCAGTTGCAC GCTGTCGCGC TCGATCTTGG CCGTAGCTTG CTGGACCATC
     2401 GAGCCGACGG ACTGGAAGGT TTCGCGGGGC GCACGCATGA CGGTGCGGCT TGCGATGGTT
     2461 TCGGCATCCT CGGCGGAAAA CCCCGCGTCG ATCAGTTCTT GCCTGTATGC CTTCCGGTCA
     2521 AACGTCCGAT TCATTCACCC TCCTTGCGGG ATTGCCCCGG AATTAATTCC CCGGATCGAT
     2581 CCGTCGATCT TGATCCCCTG CGCCATCAGA TCCTTGGCGG CAAGAAAGCC ATCCAGTTTA
     2641 CTTTGCAGGG CTTCCCAACC TTACCAGAGG GCGCCCCAGC TGGCAATTCC GGTTCGCTTG
     2701 CTGTCCATAA AACCGCCCAG TCTAGCTATC GCCATGTAAG CCCACTGCAA GCTACCTGCT
     2761 TTCTCTTTGC GCTTGCGTTT TCCCTTGTCC AGATAGCCCA GTAGCTGACA TTCATCCGGG
     2821 GTCAGCACCG TTTCTGCGGA CTGGCTTTCT ACGTGGCTGC CATTTTTGGG GTGAGGTCGT
     2881 TCGCGGCCGA GGGGCGCAGC CCCTGGGGGG ATGGGGTGCC GCGTTAGCGG GCCGGGAGGG
     2941 TTCGAGAAGG GGGGGCACCC CCCTTCGGCG TGCGCGGTCA CGCGCCAGGG CGCAGCCCTG
     3001 GTTAAAAACA AGGTTTATAA ATATTGGTTT AAAAGCAGGT TAAAAGACAG GTTAGCGGTG
     3061 GCCGAAAAAC GGGCGGAAAC CCTTGCAAAT GCTGGATTTT CTGCCTGTGG ACAGCCCCTC
     3121 AAATGTCAAT AGGTGCGCCC CTCATCTGTC ATCACTCTGC CCCTCAAGTG TCAAGGATCG
     3181 CGCCCCTCAT CTGTCAGTAG TCGCGCCCCT CAAGTGTCAA TACCGCAGGG CACTTATCCC
     3241 CAGGCTTGTC CACATCATCT GTGGGAAACT CGCGTAAAAT CAGGCGTTTT CGCCGATTTG
     3301 CGAGGCTGGC CAGCTCCACG TCGCCGGCCG AAATCGAGCC TGCCCCTCAT CTGTCAACGC
     3361 CGCGCCGGGT GAGTCGGCCC CTCAAGTGTC AACGTCCGCC CCTCATCTGT CAGTGAGGGC
     3421 CAAGTTTTCC GCGTGGTATC CACAACGCCG GCGGCCCTAC ATGGCTCTGC TGTAGTGAGT
     3481 GGGTTGCGCT CCGGCAGCGG TCCTGATCCC CCGCAGAAAA AAAGGATCTC AAGAAGATCC
     3541 TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT
     3601 GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
     3661 TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG
     3721 TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT GACTCCCCGT
     3781 CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG CAATGATACC
     3841 GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC
     3901 CGAGCGCAGA AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG
     3961 GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC TGCAGGGGGG GGGGGGCGCT
     4021 GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCCTGAATC GCCCCATCAT
     4081 CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT GTTGTAGGTG GACCAGTTGG
```

```
    4141 TGATTTTGAA CTTTTGCTTT GCCACGGAAC GGTCTGCGTT GTCGGGAAGA TGCGTGATCT
    4201 GATCCTTCAA CTCAGCAAAA GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG
    4261 CGTAATGCTC TGCCAGTGTT ACAACCAATT AACCAATTCT GATTAGAAAA ACTCATCGAG
    4321 CATCAAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT TTTGAAAAAG
    4381 CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC CATAGGATGG CAAGATCCTG
    4441 GTATCGGTCT GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAATT TCCCCTCGTC
    4501 AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGACG ACTGAATCCG GTGAGAATGG
    4561 CAAAAGCTTA TGCATTTCTT TCCAGACTTG TTCAACAGGC CAGCCATTAC GCTCGTCATC
    4621 AAAATCACTC GCATCAACCA AACCGTTATT CATTCGTGAT TGCGCCTGAG CGAGACGAAA
    4681 TACGCGATCG CTGTTAAAAG GACAATTACA AACAGGAATC GAATGCAACC GGCGCAGGAA
    4741 CACTGCCAGC GCATCAACAA TATTTTCACC TGAATCAGGA TATTCTTCTA ATACCTGGAA
    4801 TGCTGTTTTC CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG TACGGATAAA
    4861 ATGCTTGATG GTCGGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCATCTCATC
    4921 TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG GCGCATCGGG
    4981 CTTCCCATAC AATCGATAGA TTGTCGCACC TGATTGCCCG ACATTATCGC GAGCCCATTT
    5041 ATACCCATAT AAATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AAGACGTTTC
    5101 CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG ACAGTTTTAT
    5161 TGTTCATGAT GATATATTTT TATCTTGTGC AATGTAACAT CAGAGATTTT GAGACACAAC
    5221 GTGGCTTTCC CCCCCCCCCC TGCAGGTCGA CGGATCTTTT CCGCTGCATA ACCCTGCTTC
    5281 GGGGTCATTA TAGCGATTTT TTCGGTATAT CCATCCTTTT TCGCACGATA TACAGGATTT
    5341 TGCCAAAGGG TTCGTGTAGA CTTTCCTTGG TGTATCCAAC GGCGTCAGCC GGGCAGGATA
    5401 GGTGAAGTAG GCCCACCCGC GAGCGGGTGT TCCTTCTTCA CTGTCCCTTA TTCGCACCTG
    5461 GCGGTGCTCA ACGGGAATCC TGCTCTGCGA GGCTGGCCGA TAAGCTCTAA GAAACCATTA
    5521 TTATCATGAC ATTAACCTAT AAAAATAGGC GTATCACGAG GCCCTTTCGT CTTCAAGAAT
    5581 TAATTCACTG GCCGTCGTTT TACAACGTCG TGACTGGGAA AACCCTGGCG TTACCCAACT
    5641 TAATCGCCTT GCAGCACATC CCCCTTTCGC CAGCAGATCC ACATCCTTGA AGGCCGCAGC
    5701 GACGAGCAGA AGGAAACCCT CATTCGGGAA GTCAGCGAGG CCATCTCGCG CTCCCTGGAT
    5761 GCGCCGCTGA CCAGCGTGCG AGTGATTATC ACGGAGATGG CCAAGGGCCA CTTCGGCATC
    5821 GGCGGCGAAC TGGCCAGCAA GGTCAGACGC TGAAGTGGAG ATGCCCAAGG GCACTTCGGG
    5881 TCGAGGAACC CGACCTGCAT TGGGACGCGG CCACGGAGAG CGCGGGCAAA CGCCGGCACT
    5941 ATAGCCAGTG GAGTTTGTAA AACGCTATTT CAGAGCTTGG AGAGTGTCTA AGAAAGCCGG
    6001 GCGATGCCAA CCGGTCCCTT CTTCGGCTAC GTTCGTAATC AAGCCACTTC CTTTTTGCAT
    6061 TGACGCAGGG TGTCGGAAGG CAACTCGCCG AACGCGCTCC TATAGTTTTC AGCGAAGCGT
    6121 CCCAAATGTA AGAAGCCGTA GTCTAGGGCT ATCTCAGTTA TACTACGCAC ATTGGCACTG
    6181 GGATCGTTCA AGCAGGCGCG GATGCTTTCG AGCTTGCGGT TGCGGATGTA GTTCTTCGGC
    6241 GTGGTGCCGG CATGCTTCTC GAACAAATTG TAGAGCGAGC GTGGACTCAT CATCGCCAGC
    6301 TCCGCTAACC GCTCAAGGCT GATATTCCGT TTGAGATTCT CCTCAATGAA TTGAACGACT
    6361 CGCTCGAAAG ACGGGTTACC TTTGCTGAAA ATTTCACGGC TGACATTGCT GCCCAGCATT
    6421 TCGAGCAGCT TGGAAGCGAT GATCCCCGCA TAGTGCTCTT GGACCCGAGG CATCGACTTT
    6481 GTATGTTCCG CTTCGTCACA AACTAACCCG AGTAGATTGA TAAAGCCATC GAGTTGCTGG
    6541 AGATTGTGTC GCGCGGCGAA ACGGATACCC TCCCTCGGCT TGTGCCAATT GTTGTCACTG
    6601 CATGCCCGAT CAAGGACCAC TGAGGGCAAT TTAACGATAA ATTTCTCGCA ATCTTCTGAA
    6661 TAGGTCAGGT CGGCTTGGTC ATCCGGATTG AGCAGCAATA GTTCGCCCGG CGCAAAATAG
    6721 TGCTCCTGGC CATGGCCACG CCACAGGCAA TGGCCTTTGA GTATTATTTG CAGATGATAA
    6781 CAGGTCTCTA ATCCAGGCGA GATTACCCTC ACGCTACCGC CGTAGCTGAT TCGACACAGG
    6841 TCGAGGCATC CGAAGATTCT GTGGTGCAGC CTGCCTGCCG GGCGCCCGCC CTTGGGCAGG
    6901 CGAATAGAGT GCGTACCGAC ATACTGGTTA ACATAATCGG AGACTGCATA GGGCTCGGCG
    6961 TGGACGAAGA TCTGACTTTT CTCGTTCAAT AAGCAAAAAT CCATAGTTCA CGGTTCTCTT
    7021 ATTTTAATGT GAGCTCTTGG TGTGATGTAG AAAGGCGCCA AGTCGATGAA AATGCAGGAA
    7081 TTAATTCGAG ATCCCCCCCT GGCGGATGAG AGAAGATTTT CAGCCTGATA CAGATTAAAT
    7141 CAGAACGCAG AAGCGGTCTG ATAAAACAGA ATTTGCCTGG CGGCAGTAGC GCGGTGGTCC
    7201 CACCTGACCC CATGCCGAAC TCAGAAGTGA AACGCCGTAG CGCCGATGGT AGTGTGGGGT
    7261 CTCCCCATGC GAGAGTAGGG AACTGCCAGG CATCAAATAA AACGAAAGGC TCAGTCGAAA
    7321 GACTGGGCCT TTCGTTTTAT CTGTTGTTTG TCGGTGAACG CTCTCCTGAG TAGGACAAAT
    7381 CCGCCGGGAG CGGATTTGAA CGTTGCGAAG CAACGGCCCG GAGGGTGGCG GGCAGGACGC
    7441 CCGCCATAAA CTGCCAGGCA TCAAATTAAG CAGAAGGCCA TCCTGACGGA TGGCCTTTTT
    7501 GCGTAGATCC GGTCGAGGCC GGTAGCGGAG CTATCCAACG GCGGTATACC AGGAAAACAC
    7561 ACAGCAGGTA CATCAGAACA GTACCATGAC TGAAGAACAA ATAGTTTTTT CCTGATCCAT
    7621 AAAGCAGAAC GGCCTGCTCC ATGACAAATC TGGCTCCCCA ACTAATGCCC CATGCAGCCA
    7681 GCATAACCAG CATAAACGTG TCCGGTTTGA TAGGGATAAG TCCAGCCTTG CAAGAAGCGG
    7741 ATACAGGAGT GCAAAAAATG GCTATCTCTA GAAAGGCCTA CCCCTTAGGC TTTATGCAAC
    7801 ATGTACAATA ATAATGGAGT CATGAACA
//
```

Partial nucleotide sequences of plasmid DNA samples analyzed using the LightRun sequencing service of GATC Biotech are shown below. With exception of the pJP-1 sample 76IE58, all sequences represent bases as reported (received .fasta files) by GATC. For 76IE58, the sequence was exported from the CodonCode Aligner program having called the bases by reading the raw sequence trace (.abi) file, and after applying automatic trimming of sequence ends.

pJP-1 (rrnBp1):

>76IE58_clipped.fasta
CGACTTCTTTCTGGGCTAATGGCCAGCATGTCAGACGCTGCAGTGGAGATGCCCAATGGCACCTCGTTACGAGGAACCCGATCTGCATTGGGACGCGGCC
ACGGAGAGCGCGGGCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAAC
CGGTGTTGCGCGGGCTTAAAATTATTTTAAATTTCCTCTTGTCAGGCCGTAATAACTCCCTATAATGCGCCACCACTACATGTACAATTATAATGGAGTC
ATGAACATATGGTACCAAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTATATGGTGATGTTAATGGGCACAAATTTTCTGT
CAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACT
ACTTTCGCGTATGGTCATCAATGCTTTGCGAGATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAA
GAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAACGTGATACCCTTGTTACTAGAATCGAGATTAAAGGTATAGA
TTTCAAAGATGATGGAAACATTCTTGGACATCAATTGAAATACNACTATAACTCCCCTAATGTATACATCATGGCAGACTAACAAAATAATGGATTCACA
GTTCACTTCAAAATTAGACACATCATCGAAGATGGAATCGTTCAACTAGCAGACCATTATCATCACATACTCCAATCGTCGATCGCCCTTTCCTTTTACC
ATACACCTTTACTGTCCACACATTCTCCCTTTCCTAT

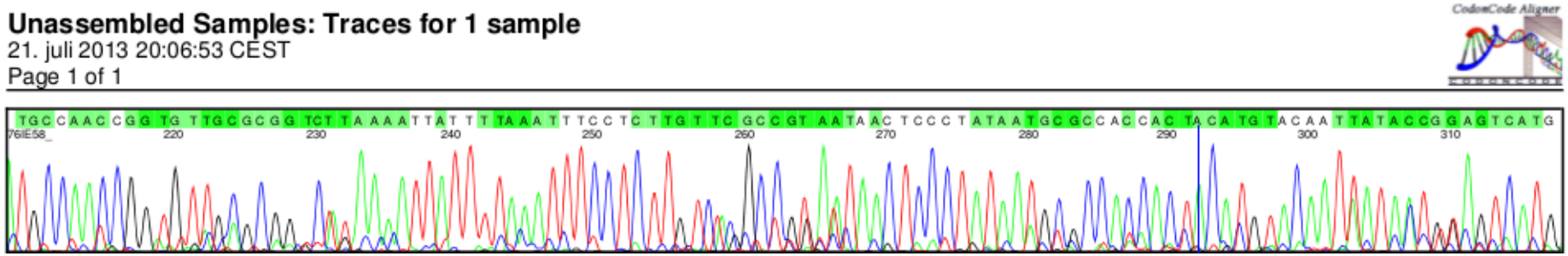


Figure A.1: Chromatogram for pJP-1 plasmid DNA.

pJP-2 (GreAp)

```
>11199099.seq - ID: 76IE57- on 2013/3/21-12:47:16 automatically edited with PhredPhrap, start with base no.: 31
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
ctggccaGCAGGTCagacGCtgaAGTGGagATgcccaagggcacnnnnngtcgaGGaAcCcgaCCTGCAttgggacGCGGcCAcggagaGCGCgGGCAaA
cGCcGgCActatagcCagtggagTTTGtAAAACGCTATTTCagaGCTTggaGagtgTCtAagaAagcCGGGGCGaTGCCAACCGgtgGcgcaacGCccta
TAAAGTAAACGAtgaCCCtTCGGGAACTTCAgggtAAAATGACtATACATGtACAATAATAATGGagtCATGAACATATggtACCAAgtAAAGGAGAAGA
ACTTTTCacTggaGTTGTCCCAATTCTTGTTGAATTaGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATac
ggAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGgtcTTCAATGCTTTGCGA
GATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAgGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTA
CAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAaGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGAcac
AAATTggaATACAACTATAACTCACACAATGTATACATcATGgcaGaca
```

pJP-3 (LacUV5):

```
>11373459.seq - ID: 59DG34- on 2013/4/25-0:38:46 automatically edited with PhredPhrap, start with base no.: 33
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
tgGCCnnAcAGGtcataCGCTGAAGTGGAGATGCCCAAgGGCannnngGgTCgAGGaacCCGACCTGCATtgGgACGCGGccACggagagCGCGGGcAAA
CGCCGGCactataGcCagtGGagtTtgtAAAACGctaTTTCagaGCTTggAGaGtGTCTAAGAAAGCCGGGGCGATGCCAACCGgtgCACCCCAGGCTTT
ACAcTTTATGCTTCCGGCTCgtATAATGTGTGGAACAtgTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAGGAGAAGAACTTTCACTGG
AGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACC
CTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGAGATACCCAGATC
ACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGC
TGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATAC
AACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTC
AACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTtTACCAGACAACCATTancttgTCCACACAATCtngCCctTTcgAAA
AGATCCCAACGAAnaagAGAGACCACAtggnccctTCTTgaaGTTTGtaACAanctgctgGggaTtnnncATGGnnatggGATG
```

pJP-4 (ArgIp):

```
>11316983.seq - ID: 59DG27- on 2013/4/15-21:20:3 automatically edited with PhredPhrap, start with base no.: 41
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
AnGTCAGACGCTGAAGTGGAGATGCCCAAGGGCACTTcGGGTCGAGGAACCCGACCTGCATTGGGACGCGGCCACGGAGAGCGCGGGCAAACGCCGGCAC
TATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTGCTTTAGACTTGCAAATGAATAA
TCATCCATATAAATTGAATTTTAACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAA
TTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTAT
TTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGAGATACCCAGATCACATGAAGCAG
CATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGT
TTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTC
ACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGAC
CATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAA
AGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAAGGATCCTCTAGCTAGAGtncAGC
TTTATGCTTGTAAaccnnttt
```

pJP-5 (iraP):

```
>11338250.seq - ID: 59DG33- on 2013/4/18-18:2:41 automatically edited with PhredPhrap, start with base no.: 26
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
cggcgactgGCcagCAGGTCAGACGCTGAAGTGGAGATGCCCAAGGGCAcncnnggtcgagGAACCCGACCTGCATTGGGACGCGGCCACGgagAGCGCG
GGCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTGCTGGTAA
TCAAACAAAAAATATTTGCGCAAAGTATTTCCTTTGTCATAAAAATAATACTTACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAG
GAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGC
AACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGC
TTTGCGAGATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACG
GGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCT
TGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAAC
ATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTtTACCAGACAACCATTACCTGTCCACAC
AATCTGCCCTTTCGAAAGATCCCAACGAAnAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATA
CAAATAAGGATCCTCTAGCTAGa
```

pJP-6 (LivJp):

```
>11426501.seq - ID: 59DG44- on 2013/5/6-22:9:53 automatically edited with PhredPhrap, start with base no.: 24
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
TCgGcGGCGAacTGGCcaGCAAGGTCaGACGCTGAAGTGGAGATGCCCAAGGGCACTTCGGGTCGAGGAACCCGACCTGCATTGGGACGCGGCCACGGAG
AGCGCGGGCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTAT
TGTTAATAAACTGTCAAAATAGCTATTCCAATATCATAAAAATCGGGTATGTTTTAGCAACATGTTGCGCTGGTCGGTGAGCCGCTTCAGCGCGGCAAAT
ATCTGGTACGGTCACGGTACCGATAACCCGTGGGATGAAGCCGTACAGCTGGTGTTGCCTTCGCTCTACCTGCCGCTGGATATTCCGGAAGATATGCGCA
CCGCGCGTCTGACCTCCAGCGAAAAACACCGTATTGTTGAACGCGTGATCCGCCGCGTCAATGAACGCATTCCGGTGGCTTACCTGACCAACAAAGCGTG
GTTCTGCGGCCATGAATTTTACGTCGATGAACGCGTGCTGGTGCCGCGCTCGCCGATTGGTGAACTGATCAACAATAAATTTGCCGGACTTATCAGCAAG
CAACCGCAGCATATTTTAGATATGTGTACTGGTAGCGGCTGCATCGCCATTGCCTGTGCTTATGCCTTCCCGGATGCAGAAGTCGACGCGGTGGATATCT
CTCCAGACGCGCTGGCGGTTGCTGAACAGAACATCGAAGAaCACGGTCTGATCCACAACGTCATTCCGATTCGTTCCGATCTGtTCCGCGAcTTGcCGAA
AGTGCAgTACGACCTGATTGTCACTAACCCGcCGTATGTCgATGCGGAAgAtATGTCcgAnCTGCcnAACgAATACCGCCaCGAGcCGGAACTgGGcCTG
GCATCTGGcaCTGACgGcCTGAAACTGACGCGTCgCATTCTCGGtAACgCGGcAgAtTacCtTngcTGAtGAtggnGtGnTGAtTTg
```

pJP-11 (GFP-LVA):

>11372291.seq - ID: 59DG42- on 2013/4/24-23:11:32 automatically edited with PhredPhrap, start with base no.: 21
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
atggCatggatgaACTATACaAAGCAGCAAACGACGAAAACTACGCTTTagtAGCTtaacttAAGGGATCCTCTAGctaGAGTCAGCTTTATGCTTGTAA
ACCGTTtTGTGAAAAAATTTTTAAAATAAAAAAGGGgaCCTCTAGGgTCCCCAAtTAATTAgtAATATAATCTATTAAAGGTCATTCAAAAGGTCATCca
ccGGATCAGCTTagtAAAGcccTCgctAGATTTTAATGCGGATGTTGCGATTACTTCGCCAACTATTgCgATAACAAGAAAAAGCCAGCCTTTCATGATA
TATCTCCCAATTTGTGTAGGgctTATTATGCACGCtTAAAAATAATAAAAGCAGACTTgaCCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCAT
CTAACGCTTGAGTtaAGCCGCGCCGcGAAGCGGCGTcGgCTTGaacGAATTGTTannnATTATTTGCCGACTACCTTGCAaggAtnngGCCTTGATGTTA
Cc

His candidate clones:

59DG30:

>11338096.seq - ID: 59DG30- on 2013/4/18-18:13:18 automatically edited with PhredPhrap, start with base no.: 25
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
cggcgactGGnCcagCAAGGTCAGACGCTgaagtGGAGATGCCCAAGGGCacnatcgGgtCGAGGAAcCCGACCTGCATTgGgACGCGGCCACGgaGAGC
GCGGGCAAACGCCGGCAcTATAGCCAGTGGAGTTTGtAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGgtGCCAT
ATAAAAAAGCCCTTGCTTTCTAACgtGAAAGTGGTTTAGGTTAAAAGACACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACA
TACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTG
CGAGATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAA
CTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAnACATTCTTGGA
CACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTG
AAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGAcaACCATTACCTGTCCACACAATC
TGCcCTTTCGAAAGATCCCAACGAAnAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATacAA
ATAAGGATCCTCTAGCTAGa

59DG46:

>11486243.seq - ID: 59DG46- on 2013/5/17-21:33:55 automatically edited with PhredPhrap, start with base no.: 28
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
ggcgactggCcagCAGGTCAGACGCtgAAGTGGAGATGCCCAAGGGCAnnnngggtcGAGGAACCCGACctgCATtgGGACGCGGCCACGgaGAGCGCGG
GCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTgGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTGCCATAAAA
TAAAAGCCCTTGCTTTCTAACGTGAAAGTGGTTTAGGTTAAAAGACACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAGGAGAAGA
ACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATAC
GGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGA
GATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTA
CAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACAC
AAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAG
ATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGC
CCTTTCGAAAGATCCCAACGAAnAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAaatAA
GGATCCTCTAGCTAGa

59DG47:
>11486366.seq - ID: 59DG47- on 2013/5/17-23:21:49 automatically edited with PhredPhrap, start with base no.: 21
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
aTCggcGGCGAnTGGCcnGCAAGGTCAGACGCTGAAGTGGAGATGCCCAAGGGCAcTTCGGGTCGAGGAACCCGACCTGCATTGGGACGCGGCCACGGAG
AGCGCGGGCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTGC
CATAAAATATATAAAAAAGCCTTGCTTTCTAACGTGAAAGTGGTTTAGAAAGACACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAA
GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATG
CAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATG
CTTTGCGAGATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGAC
GGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTC
TTGGACACAAATTGgAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGgAATCaAAGTTAACTTCaAAATTAGACACAA
CATTGAAgAtGGAAGCGTTCAACTAGCAGAcCATTATCAACAAAATACTCCAATTGGCGATGGCCCtGTCCTTTTACCAGACAACCATTACCTGTCCACA
CAATCTGCCCTTtCGAAAgATCCCaACGAAAAgAgAganCACaTGGnCCTTCTtGAgTTTntAACAgCTGCTGGGAtTACaCaTgGnatgGntGaAcTAT
acAAaTAagnAtCtCTAgCTAgA

59DG48:

>11486401.seq - ID: 59DG48- on 2013/5/17-23:21:47 automatically edited with PhredPhrap, start with base no.: 31
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
anTGGCcnGCAAGGTCagACGCTGAAGTGGAGATGCCCAAGGGcACTTCGGGTCGAGGAACCCGACCTGCATTGGGACGCGGCCACGGAGAGCGCGGGCA
AACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAGAAAGCCGGGGCGATGCCAACCGGTGCCATATAAAAA
AGCCCTTGCTTTCTAACGTGAAAGTGGTTTAGGTTAAAAGACACATGTACAATAATAATGGAGTCATGAACATATGGTACCAAGTAAAGGAGAAGAACTT
TTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAA
AACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGAGATA
CCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAG
ACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAAT
TGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAaTGnAATCAAAGTTAACTTCAAAATTAGACACaACATTGAAGATGG
AAGCGtTCAACTAGCAgAnCATTATCAACAAAATACTCCAaTTGGCGATGGcCCTGTCCtTTTAccaGAcAACCaTTAcCTGTCCACACAATCTGcCCTT
TCgAAAGATCCCaACGAAAAgAGAgAccACatnGgtCCTTCtTGAGTTtGtAACAGCTGCTGGgatTacaCATGGCaTGGntGaacTataCAATa

Thr candidate clone:

```
>11486356.seq - ID: 59DG53- on 2013/5/17-23:21:48 automatically edited with PhredPhrap, start with base no.: 21
Internal Params: Windowsize: 20, Goodqual: 19, Badqual: 10, Minseqlength: 50, nbadelimit: 1
TCnngCgGCgAcTGgcCaGCAAGGTCAGACGCTGAAGTGGAGATGCCCAAGGGcACTtCnGGTCGAGGAACCCGACCTGCATTGGGACGCGGCCACGGAG
AGCGCGGGCAAACGCCGGCACTATAGCCAGTGGAGTTTGTAAAACGCTATTTCAGAGCTTGGAGAGTGTCTAAgAAAGCCGGGGCGATGCCAACCGGTAA
CTGGTACCTGCCGTGAGTAAATTAAAATTTTATTGACTTAGGTCACTAAATACTTTAACCAATATAGGCAACATGTACAATAATAATGGAGTCATGAACA
TATGGTACCAAGTAAAGGAgAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGA
GAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCG
CGTATGGTCTTCAATGCTTTGCGAGATACCCAGATCACATGAAGCAGCATGACTTTTTCAAGAGTGCCaTGCCCGAAgGTTATGTACAGGAAAGAActat
aTTTTTCAAAGATGACGGGAACTaCaagAcacgTGCTGAagTCAAGTTTGAAGGtGATACCCTTGtTAATAnaatCgAGTtaaaaGGTAtTGAtTttt
```

B Abbreviations

bp: Base pairs

ppGpp: Guanosine tetraphosphate

pppGpp Guanosine pentaphosphate

DNA: Deoxyribonucleic acid

E. coli: *Escherichia coli*

Fis: Factor for Inversion Stimulation

RNA: Ribonucleic acid

mRNA: Messenger RNA

rRNA: Ribosomal RNA

tRNA: Transfer RNA

RBS: Ribosomal binding site

GFP: Green Fluorescent Protein

GTP: Guanosine tetraphosphate

GDP: Guanosine diphosphate

NTNU: Norwegian University of Science and Technology

OD600: Optical density at 600 nm

RNAP: RNA Polymerase

iGEM: International Genetically Engineered Machine competition

PCR: Polymerase chain reaction

SOE: Sequence Overlap Extension

LB: Lysogeny Broth

RSH: RelA/SpoT Homologue

SLIC: Sequence- and Ligation-Independent Cloning