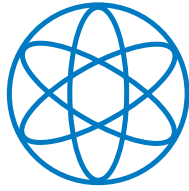# Evaluation of an E. coli Cell Extract Prepared by Lysozyme-Assisted Sonication and Transcriptional Interference in Toehold Switch-Based Circuits

Elisabeth Falgenhauer

Technische Universität München
TUM School of Natural Sciences

Evaluation of an E. coli Cell Extract Prepared by Lysozyme-Assisted Sonication and Transcriptional Interference in Toehold Switch-Based RNA Circuits

Elisabeth Falgenhauer

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzende: Prof. Dr. Karen Alim

Prüfer der Dissertation:
1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Ulrich Gerland

Die Dissertation wurde am 15.09.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 07.11.2022 angenommen

Alles verändert sich mit dem,
der neben einem ist
oder neben einem fehlt.

(Sylke-Maria Pohl)

Contents

List of publications

Peer-reviewed publications that are part of this cumulative thesis

<u>**E. Falgenhauer**</u>, A. Mückl, M. Schwarz-Schilling, and F. C. Simmel
Transcriptional interference in toehold switch-based RNA circuits.
ACS SyntheticBiology, vol. 11, no. 5, pp.1735-1745, 2022

<u>**E. Falgenhauer**</u>, S. von Schönberg, C. Meng, A. Mückl, K. Vogele, Q. Emslander, C. Ludwig, and F. C. Simmel
Evaluation of an E. coli cell extract prepared by lysozyme-assisted sonication via gene expression, phage assembly and proteomics.
Chembiochem,vol. 22, no. 18, pp. 2805-2813, 2021

Other peer-reviewed publications

J. Müller, A. C. Jäkel, J. Richter, M. Eder, <u>**E. Falgenhauer**</u>, and F. C. Simmel
Bacterial Growth, Communication, and Guided Chemotaxis in 3D-Bioprinted Hydrogel Environments.
ACS Applied Materials and Interfaces, vol.14, no.14, pp.15871-15880, 2022

K. Vogele, <u>**E. Falgenhauer**</u>, S. von Schönberg, F. C. Simmel, and T. Pirzer
Small antisense DNA-based gene silencing enables cell-free bacteriophage manipulation and genome replication.
ACS synthetic biology, vol. 10, no. 3, pp. 459-465, 2021

J. List, <u>**E. Falgenhauer**</u>, E. Kopperger, G. Pardatscher, and F. C. Simmel
Long-range movement of large mechanically interlocked DNA nanostructures.
Nature communications, vol. 7, no. 1, pp. 1-7, 2016

Zusammenfassung

Zellextraktprojekt Die synthetische Biologie fokusiert sich darauf, biologisch basierte Bauteile und Systeme weiterzuentwickeln und mit neuartigen Funktionen auszustatten. Hierfür werden regulatorische Genexpressionselemente, die aus der Natur stammen oder von Grund auf neu entwickelt werden, zu genetischen Schaltungen oder auch größeren Netzwerken verknüpft. Als Alternative zu *in vivo*-Studien, können genetische Schaltungen auch innerhalb von zellfreien Genexpressionssystemen untersucht werden. Zellfreie Experimente können unter Verwendung von linearen DNA-Templaten durchgeführt und notwendige Chemikalien einfach zu den Reaktionen hinzugefügt werden, was den Design-Konstruktions-Test-Zyklus deutlich beschleunigt. Nach der anfänglichen Verwendung von Rohextrakten wurden über die vergangenen Jahrzehnte einige Protokolle zur Zellextraktherstellung getested und optimiert, allerdings gestaltet sich die Reproduzierbarkeit zwischen verschiedenen Chargen und Laboren schwierig. Einer der entscheidenden Schritte während der Zellextraktherstellung ist die Zelllyse, welche die Qualität und die Quantitiät der extrahierten Enzyme beeinflusst. In dieser Arbeit wurde ein *E. coli*-Zellextrakt hergestellt, wobei für die Lyse eine Kombination aus Lysozym-Inkubation und Ultraschall-Behandlung verwendet wurde. Zur Bewertung der Qualität wurden einzelne Reporterproteine z.B. YFP bis zu Konzentrationen von 0.6 mg/ml zellfrei exprimiert. Zusätzlich wurden T7-Bakteriophagen bis zu Titern von 10^8 PFU/ml assembliert. Quantitative Proteomanalysen wurden verwendet um die Extrakte untereinander und mit einem komerziellen Produkt zu vergleichen. Obwohl die Unterschiede, die sich in den selbst produzierten Extrakten durch

übernommen von Falgenhauer et al. [1]

IX
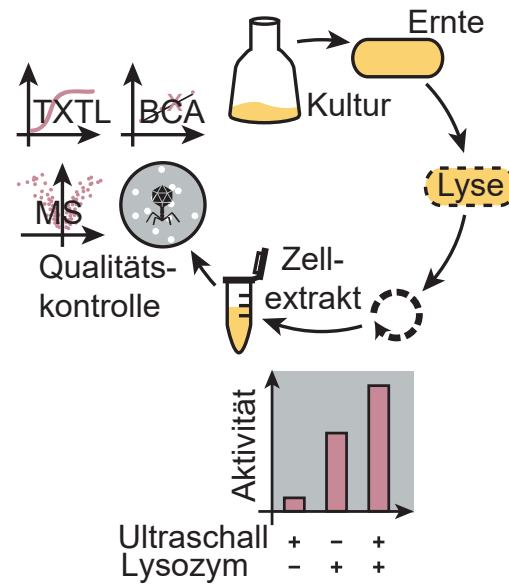
die Verwendung von unterschiedlichen Lysebedingungen ergeben, überraschend klein waren, konnte trotzdem ein Anstieg in der freigesetzten Menge an DNA-bindenden Enzymen für eine steigende Zahl an Ultraschallzyklen festgestellt werden. Es wurde außerdem beobachtet, dass in den selbst hergestellten Zellextrakten Proteine, die im Karbohydrat-Metabolismus, in der Glykolyse, sowie in Aminosäure- und Nukleotid-Stoffwechselwegen beteiligt sind, stärker vertreten sind, während Proteine, die der RNA-Modifikation und Prozessierung, der DNA-Modifikation und Replikation, der Transkriptionsregulation, -initiation und -termination, der Translation, sowie dem Zitratzyklus zugeordnet werden können, im kommerziellen Extrakt höher angereichert sind.

Nach erfolgreicher Hochskalierung des Batchvolumens könnte das vorgestellte Zellextrakt zur kommerziellen Produktion von medizinisch relevanten Phagen verwendet werden, welche anstelle von Antibiotika zur Bekämpfung von multiresistenten Keimen eingestzt werden könnten.

Transkriptionsinterferenz- und Toeholdswitch-projekt Wegen des beschleunigten Design-Konstruktions-Test-Zyklus werden genetische Netzwerke z.B. Logikgatter häufig in zellfreien Systemen getestet und optimiert. Allerdings gibt es viele interessante Anwendungsfelder *in vivo*, die die Funktion der Netzwerke auch im Zielorganismus vorraussetzen. Die Netzwerkcharakteristiken können nämlich durch Wechselwirkungen mit dem Host-Metabolismus beeinflusst werden, der ebenfalls ein kompliziertes, genetisches Netzwerk darstellt. Organismen haben eine sehr große Bandbreite von orthogonalen, regulatorischen Elementen evolviert, um Wechselwirkungen zwischen verschiedenen, metabolischen

Toehold-Switch Riboregulator + Trigger & Anti-Trigger in verschiedenen Promotoranordnungen

TI-Effekt modulierte Regulationsfunktion — 2-Input-2-Output Logikgatter

übernommen von Falgenhauer et al. [2]

Pfaden zu vermeiden. Unter anderen ist Genregulation basierend auf regulatorischer RNA ein sehr wichtiger Mechanismus in Zellen, welcher immer häufiger auch für genetische Schaltungen in der synthetischen Biologie Verwendung findet. Ein Beispiel
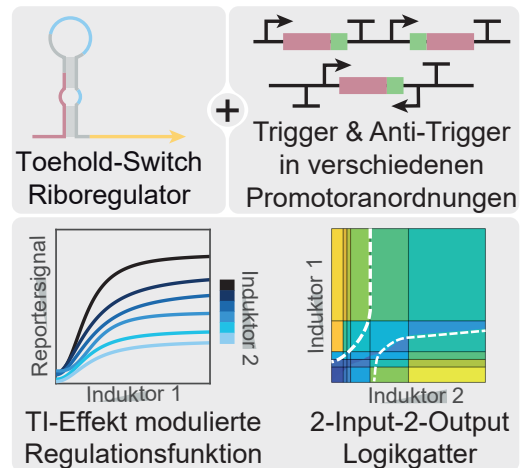
für einen rational designten, post-transkriptionellen Riboregulator ist ein Toehold-Switch. Eingebunden in die 5'-untranslatierte Region von mRNA-Molekülen, blockieren Toehold-Switches mit Hilfe von Sekundärstruktur die Ribosomen-Bindestelle (RBS) und verhindern damit das Binden von Ribosomen. Ein Trigger-Molekül (Tr) kann die RBS freisetzen, indem es bindet und eine Strangverdrängungsreaktion auslöst (toehold-mediated strand displacement). Unter Verwendung einer gegen den Trigger gerichteten Anti-Sense-RNA (Anti-Trigger, AT) kann der Prozess umgekehrt werden und Translation wieder ausgeschaltet werden.

In dieser Arbeit wurden *in vivo*-Schaltungen mit zwei induzierbaren Promotoren, die Tr- und AT-Transkription regulieren, entworfen und getestet. Der dazugehörige Toehold-Switch befand sich am selben DNA-Strang (*in cis*) unter einem konstitutiven Promoter und regulierte die Expression eines Reporterproteins. Da die Sequenzkomplementarität von Tr und AT auch kompakte Designs mit Anti-Sense-Transkription ermöglicht, konnten vielfältige Designmöglichkeiten für die Promoteranordnung untersucht werden. Das dynamische Verhalten der getesteten Systeme wurde durch Transkriptionsinterferenz, welche hauptsächlich auf nah benachbarte, aktive Promotoren und einer daraus resultierenden erhöhten, lokalen RNAP-Konzentration zurückzuführen ist, beeinflusst. Der stärkere Promoter im System konnte den größten Teil der vorhandenen Polymerasen rekrutieren und wurde dadurch hochreguliert. Der Schwellwert für den Abstand der beteiligten Promotoren, bei dem der Effekt noch zu beobachten war, war größer, wenn die Promoterstärken höher waren. Basierend auf den Erkenntnissen wurden die besten Promotoranordnungen ausgewählt und zu einem zwei-Input-zwei-Output-Logikgatter zusammengestellt [2].

Alternative Designs mit fluoreszenten Aptameren anstelle von Trigger und Anti-Trigger können dabei helfen, die Transkriptionsinterferenzeffekte bei drei aktiven Promotoren besser zu verstehen. Die gewonnenen Erkenntnisse können bei der Konstruktion von weiteren genetischen Netzwerken mit vergleichbaren Promoteranordnungen helfen, die Systemcharakteristiken zu optimieren.

Abstract

Cell extract project Synthetic biology focuses on the design and construction of biologically based parts and systems with engineered and novel functions. Thereby regulatory gene expression elements found in nature or designed from scratch are organized in circuits or greater networks. Alternatively to *in vivo* studies genetic circuits can also be studied within cell-free gene expression systems. Cell-free experiments can be performed using linear DNA templates and the reaction can be easily supplemented with chemicals, which speeds up the design-build-test-cycle. Over the past decades, starting from crude cell extracts, various cell extract preparation protocols were tested and optimized, but cell-free systems lack reproduciblility between different batches and labs. One of the crucial steps

adapted from Falgenhauer et al. [1]

during the preparation of cell extract is the cell lysis procedure, which determines the quantity and quality of the extracted enzymes. In this study we prepared *E. coli* cell extract using a combination of lysozyme incubation and sonication cycles for the lysis step. As quality measure, the cell-free expression of single reporter proteins e.g. YFP at concentrations up to 0.6 mg/ml was demonstrated. Additionally T7 bacteriophages up to titers of 10^8 PFU/ml were assembled. State-of-the-art quantitative proteomics was used to compare the produced extracts with each other and with a commercial extract. The differences in protein composition, which were a result of different lysis settings used for the preparation of our self-made extracts, were surprisingly small, but we were able to observe an increase in the release of DNA-binding proteins for increasing sonication cycle numbers. Proteins taking part in carbohydrate metabolism, glycolysis, amino acid and nucleotide related pathways
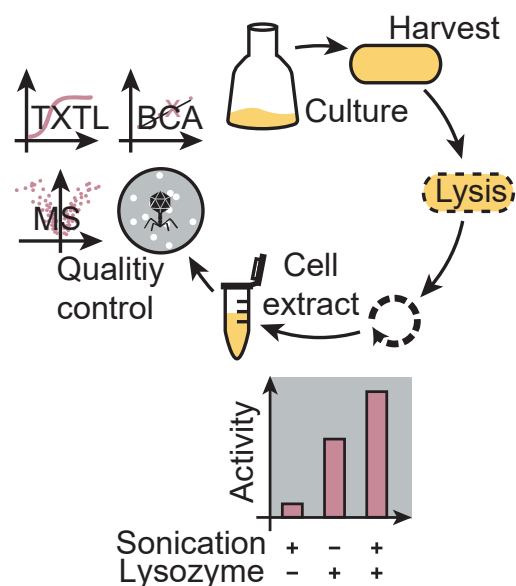
were found to be more abundant in our extracts, while proteins related to RNA modification and processing, DNA modification and replication, transcription regulation, initiation, termination, translation and the TCA cycle were found to be enriched in the commercial extract [1].

After upscaling the batch volume, the presented cell extract could be used for the commercial production of medically relevant phages to replace antibiotics in the treatment of patients with multi resistant bacterial infections.

Transcriptional interference and toehold switch project Due to the sped up design-bulid-test cycle genetic circuits are often tested and optimized within cell-free systems. However there are various interesting applications *in vivo*, which need the circuits to function also within the target organism. The circuit characteristics might be altered *in vivo* due to interactions with the host metabolism, which is by itself a complicated genetic network. Organisms have evolved a variety of molecular implementations of orthogonal regulatory elements to avoid crosstalk between different metabolic pathways. Amongst others gene regulation based on regulatory RNA is an important mechanism, which is increasingly used also

Toehold switch riboregulator

$+$

Trigger & anti-trigger in different promoter arrangements

Reporter signal

Inducer 2

Inducer 1

Inducer 1

Inducer 2

TI effect modulated regulation function

2-input 2-output logic gate

adapted from Falgenhauer et al. [2]

for regulatory circuits in synthetic biology. One example for a rationally designed, post-transcriptional riboregulator is a toehold switch. Toehold switches, placed in the 5' untranslated region of mRNA molecules, sequester via a secondary structure the ribosome-binding site (RBS) and inhibit ribosomes from binding. A trigger RNA molecule (Tr) can release the RBS upon binding via a toehold-mediated strand displacement process. Using antisense RNA against trigger molecules (anti-trigger RNA, AT), the process can be reversed and translation can be switched OFF again. In this thesis circuits were designed and tested *in vivo*, which utilized two inducible promoters that regulated trigger and anti-trigger transcription. The cognate toehold switch was *cis*-encoded downstream of a constitutive promoter and regulated the
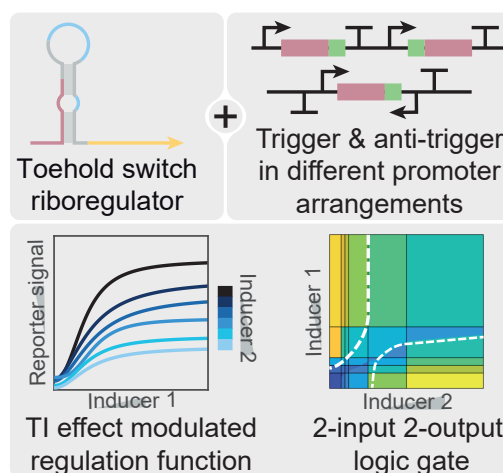
expression of a reporter protein. As the sequence complementarity of Tr and AT also allowed for compact designs with antisense transcription, various design options for the arrangement of the promoters were explored. The dynamic behaviour of the tested systems was influenced by transcriptional interference effects, which were mainly based on the close proximity of two active promoters and the resulting increased local RNAP concentration. The stronger promoter could recruit the major portion of these RNAPs and was thereby upregulated. The threshold distance, for which this effect still could be observed, was larger when the participating promoters were stronger. Based on this insights, the optimum promoter arrangements were combined to implement a two-input two-output logic gate with two toehold switches and two reporters [2].

Alternative designs, in which trigger and anti-trigger are replaced by fluorescent aptamers, could be used in future projects to study transcriptional interference effects for three active promoters in greater detail. One could benefit of these results in the construction of genetic circuits with similar promoter arrangements.

Chapter 1

Synthetic biology

1.1 In vivo versus cell-free synthetic biology

What is synthetic biology? Synthetic biology is an interdisciplinary field, which gives an engineering perspective to biology and aims for the design and construction of complex, biologically based parts, devices and systems, which display novel functions. This approach is applied to all levels of biological structures ranging from biomolecules to whole cells and organisms [3]. With the help of the most prominent model organism, the gram-negative bacterium *Escherichia coli* (*E. coli*), the roots of synthetic biology were set back in the 1960s (Figure 1.1A), when Jacob and Monod [4] discovered the regulation mechanism of the *lac* operon. Today these transcription factors are basic components in synthetic circuits.

The assembly of circuits was enabled by the development of cloning techniques based on DNA restriction enzymes discovered in 1969 [5] and DNA ligases successfully isolated by several groups in the 1960s [6][7][8][9][10]. Finally further advances in genetics including the polymerase chain reaction (PCR) [11] and DNA sequencing methods [12] enabled the amplification of DNA fragements and the verification of the cloning success respectively.

As the basis of the field was then set, synthetic biology could step further in the modules era (Figure 1.1B), in which existing biological systems were modified to perform desired tasks. In this top-down approach characterized biological elements like promoters, ribosome binding sites (RBS), genes of interest and terminators collected as standard biological parts [13] can be assembled to genetic modules and transferred to the host cell (chassis) [14]. The modules can be plugged together to perform as bistable switches [15] [16], oscillators [17][18] and cascades [19] for example. As the basic components and also the complete modules turned out to show a different behaviour upon connection to other building blocks, which is known as context dependency (see section 1.3 for details), researchers developed different

design strategies: Circuits were designed and tested in iterative cycles to refine the performance until the requirements were met. In other approaches, different circuit variants were tested in parallel [20] or direct evolution was used as optimization method [14].

Even though it has been demonstrated that feedback and feed-forward motifs can provide robustness to noise or genetic context [21], the assembly of systems with higher complexity in the current era of synthetic biology (Figure 1.1C) needs the implementation of insulating modules and devices. After achieving these goals whole pathways might be engineered and environmental or therapeutic applications might be found [21].
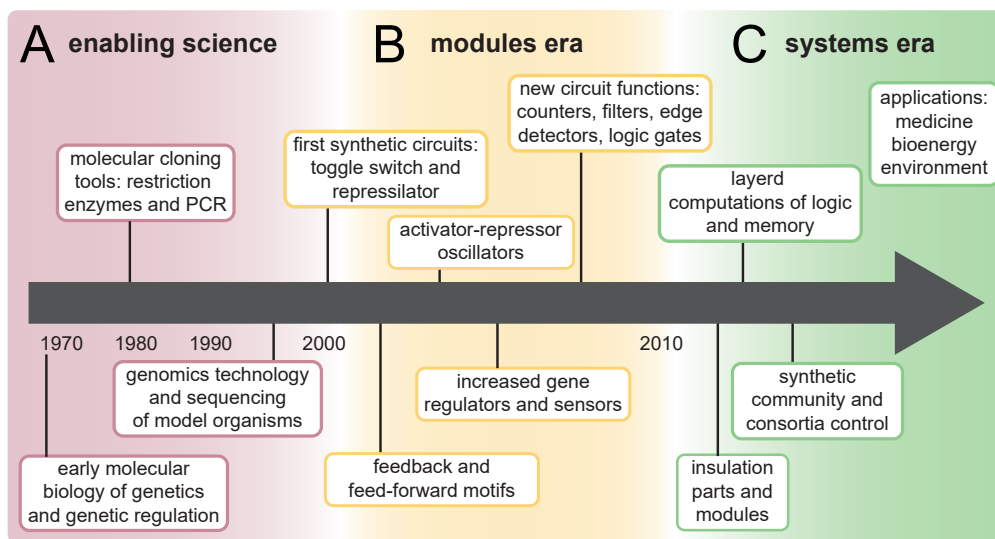
Figure 1.1: Timeline of synthetic biology. (A) The discovery of molecular cloning tools and gene regulation mechanisms (e.g. the lac operon) were the basis for the development of the field of synthetic biology. (B) In the modules era of synthetic biology first synthetic circuits were successfully assembled. (C) Nowadays researchers focus on the assembly of lager systems based on functional modules, but biomolecular systems often lack modularity, which makes insulating devices necessary. The figure was reproduced with permission from [21].

Researchers also started in parallel to follow a bottom-up approach, where biological and artificial building blocks are assembled to biomimetic systems or artificial cells. This was for example demonstrated by Weiss et al. [22], who created a high-throughput microfluidic method to generate stable liposomes, which could be sequentially loaded with biomolecules. The method might be forward engineered to

create more complex systems and ideally synthetic cells.

In vivo synthetic biology   The characterization of genetic networks *in vivo* e.g. in *E. coli* is beneficial as the (bacterial) cell maintains the DNA templates over numerous cell cycles. These are mainly introduced as compatible, circular plasmids with specific copy numbers. In addition, the cell provides the whole expression machinery, cofactors, chaperons and other components, which are needed for the circuit to operate. To optimize the work with bacteria, numerous strains with engineered genomes are available. The modifications include amongst others the deletion of nucleases (e.g. Endonuclease I; genotype: endA1), the reduction of proteolysis of expressed proteins (e.g. mutation in outer membrane protein protease VII, genotype: ompT) and the reduction of unwanted recombination of cloned genes (e.g. inactivated recA, genotype: recA1). Also additional proteins such as a T7 polymerase (genotype: DE3) or repressor proteins (e.g. TetR, genotype: Tn10) can be integrated into the genome [23]. As a result the stability of nucleic acids and proteins is increased and high gene expression yields are established. However the circuit performance can be *in vivo* restricted due to unpredictable genetic crosstalk and resource sharing between the circuit and the endogenous systems present in the host [24]. It remains also a challenge to determine simultaneously numerous biochemical parameters and to monitor the changes in cellular states [25].

Cell-free synthetic biology   Although *in vivo* projects are dominant in the field of synthetic biology, cell-free gene expression systems (CFS) turned out to be a useful tool for various applications, which were reviewed by Silverman et al. [26]. For example CFS played an important role in understanding biological basics like the genetic code [27], DNA replication [28], and eukaryotic translation [29]. CFS were also utilized for the production of proteins at high yields [30], rapid prototyping of regulatory elements for synthetic biology [31], and the re-engineering of metabolic pathways [32].
Cell-free systems are interesting as they give direct access to the gene expression machinery of a cell without having transport barriers and thus allow for a simple utilization, manipulation, and complementation of the reaction. Even linear DNA fragments can be used as templates within a CFS as they can be prevented from degradation by the RecBCD complex under addition of pause elements (chi sites) or a inhibitory protein (GamS) speeding up the design-build-test cycle [33]. Unwanted interference with the host metabolism is eliminated by removing nucleic acids of the host and cell viability is not required any longer, which enables also the

addition or expression of toxic compounds [34]. The lack of cellular growth and its associated complex gene regulatory processes and metabolism makes also the systems conceptually simpler to understand and quantitate [1].

In order to perform cell-free gene expression reactions, the necessary biochemical machinery has to be extracted from live cells in an active form and combined with supplements (transcription-translation (TXTL) buffer containing precursor molecules and chemical fuels) [1]. There are two extraction alternatives: the first one is to reconstitute the TXTL system from individually purified components. A reconstituted system like the "protein synthesis using recombinant elements" (PURE) system [35], which is commercially available, is expensive to produce, but one has control over each component and one can omit unwanted enzymes, which would degrade nucleic acids or hydrolyze nucleoside triphosphates. A successful approach to reduce costs and preparation time was shown by Grasemann et al. [36], who His-tagged each individual component and co-purified all of them in a single pot.

The alternative is the preparation of a crude (S30) cell extract containing the whole soluble fraction of macromolecules including the TXTL apparatus. Most of the current preparation protocols are based on a procedure first described by Zubay [37], but protein expression yields could be considerably increased by optimizing the preparation protocol or buffer composition[30][38]. Technical innovations [39] [40][41] and a better understanding of metabolic processes within the extract and energy regeneration systems supported the optimization process [42][43]. Also high-throughput methods and machine-learning tools were used in the optimization processes recently [44][45][46].

However at the moment cell-free protein synthesis lacks reproducibility between different labs and batches (batch-to-batch variation), which is a result of different preparation protocols, the experience of the person who prepares the extract/reaction and the missing repository of data and cell-free expression materials. In future this problem can be solved by using a fully automated cell-free expression process beginning with the extract preparation to the final sample preparation [47]. In view of these batch-to-batch variations it seems to be very attractive to use commercial available extracts. One example is the myTXTL kit (Arbor Biosciences), which is prepared of Rosetta 2 (DE3) cells cultured at 40 °C instead of 37 °C [33], which results in higher chaperone concentrations and therefore in a beneficial effect on the gene expression yield [48]. Accordingly GFP expression levels of up to 4 mg/ml were reached in non-fed batch-mode and up to 8 mg/ml in a semi-continous synthetic cell stetting consisting of cell extract loaded liposomes. Also T7 bacteriophages up to a titer of 10^{13} PFU/ml were assembled using this kit [33]. However besides the fact,

that these commercial products are usually quite expensive per reaction, one has to chose the optimal kit for the desired purpose carefully as the gene expression kinetics and yield varies between the different distributors largely as it was demonstrated by Burrington et al. [49].

Additionally, the in house production of cell extracts increases flexibility as genome engineered strains can be used for the preparation and thereby specific requirements (e.g. lack of repressor proteins) can be met. The basic steps for cell extract preparation include cultivation of the cells, followed by cell harvesting and washing steps. The obtained cell pellet can be flash frozen and stored or one continues directly with cell lysis and the clarification of the cell suspension (Figure 1.2). In addition, a run-off reaction is performed, during which residual DNA and RNA is degraded and bound ribosomes are released. This is commonly followed by a buffer exchange via dialysis in order to provide optimum reaction conditions for gene expression. Preparation details vary between different labs and there are also protocols which omit some of the steps as exemplified by a high throughput protocol with just three steps [50].

One of the most critical steps during the production of an active cell extract is cell lysis. The cell wall has to be substantially robust as the osmotic pressure form the inside, which is about 5 atmospheres for *E. coli* [51] but can reach up to 20 atmospheres for *Micrococcus lysodeikticus* [52] and additionally depends on the morphology of the cell wall [53] and on the growth conditions of the cell [54], would otherwise burst the cells. Therefore one has to apply enough engery during cell lysis to break up the cell, but simultaneously avoid the degradation of enzymes needed for the protein expression. Accordingly, a wide range of lysis methods have been tested and improved for bacterial cell extract preparation in the past. The pros and cons of the different lysis methods with respect to the integrity of different biomolecules and cell structures, scalability and prevalence of the method are summarized in the following section.
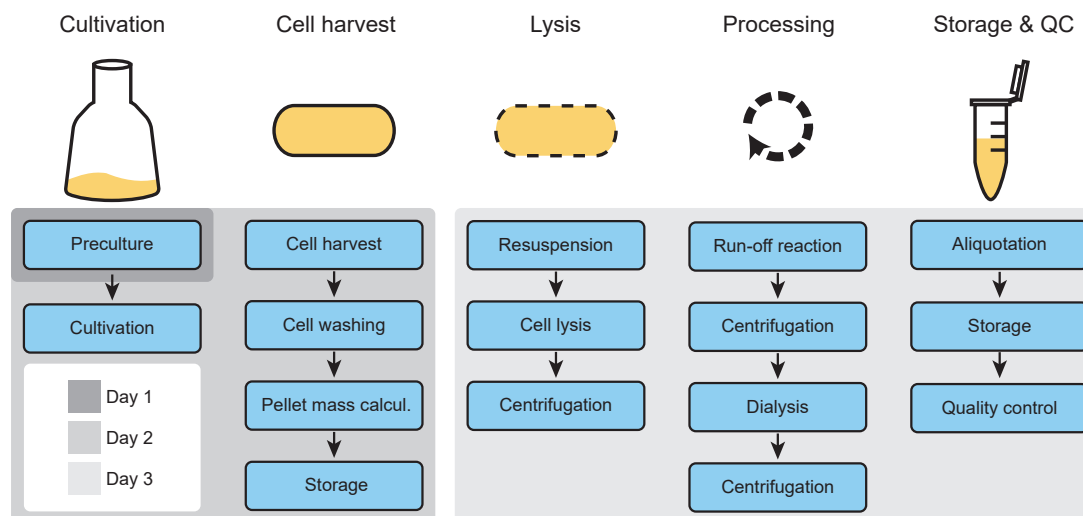
Figure 1.2: Overview of the cell extract preparation workflow, which includes cell culturing, cell harvest and washing, cell lysis, further processing steps and storage of the extract, followed by assessment of its quality.

1.2 Cell lysis methods

Bead beating Bead beating is an inexpensive, mechanical disruption method in which beads are added to a cell suspension and the tube is then shaken. The shaking in a frequency of up to 4500 rpm causes a collision between the beads and the cells and the disruption of cell walls. As a result nucleic acids and proteins are extracted, which was applied for plant tissues [55], yeast [56], organs [57] and gram positive and negative bacteria [58]. The particle size generated during the process depends on the bead diameter, as smaller beads have a larger surface area per unit mass and therefore provide a more frequent contact with the sample [58]. The bead diameter can range from microns to centimeters, but for bacteria mainly glass or zirconia beads of a diameter of 0.5 or 1 mm are used. The lysis efficiency also increases with increasing shaking frequencies and disruption times. Extended disruption times result in a higher fragmentation of DNA, which is later anyways degraded, but also in an increased heat production, which might be harmful for the extracted proteins [58]. Even though also bigger sample tubes are commercially available and usually used for hard tissue homogenization (e.g. 50 ml tubes, OMNI International), this disruption method is in the context of bacterial lysis usually

performed in 2 ml sample tubes, which limits the batch scalability. To avoid protein degradation during the shaking process the sample tubes should be filled without any remaining air bubbles. The cell suspension viscosity upon addition of beads increases strongly, which slows the filling speed of the tubes down and makes the method time consuming [59]. However Sun et al. [59] presented in 2013 a cell extract preparation protocol based on bead beating for subsequent cell-free protein syntesis, with a redution of 98% in costs compared to commercial systems.

Figure 1.3: A bacterial cell suspension containing glass beads is filled in a 2 ml bead beating tube. The suspension shows high viscosity. The figure was adapted with permission from [59].

High pressure cell disruptors In high pressure cell disruptors cells are exposed to high pressure which is released through a valve. The disruption is caused by the pressure difference, cavitation and shear stress [52]. Although the system is cooled, the lysis comes along with a high temperature increase in the sample [60], which can have a detrimental effect on the released proteins. High pressure cell disruptors allow the regulation of pressure that is applied, the velocity of cell suspension release and the number of passages. In addition it was reported by Uhlmann et al. [61] that the geometry of the seat, valve and impact ring have an influence on the cell lysis and the activity of the resulting extract, which limits the transferability of optimized protocols for one device to another device. For example they reported a similar protein composition for extracts produced with an EmulsiFlex-B15 high pressure homogenizer (Avestin Europe GmbH, Germany) or with a French press (SLM Instruments), but they observed significant differences in gene expression capability of the extracts. They observed a about 3 fold higher fluorescence signal of a expressed GFP for the Avestin cell extract, but the total concentration of expressed GFP (inactive and active) determined using a SDS PAGE was higher for the French press samples. To overcome this problem they constructed a continuously operating

system with interchangeable geometry.

Ultrasound sonication Also widely available lab instruments are ultrasound tip sonicators, which are commonly used for cell lysis via cavitation in protein purification protocols. A typical sonicator (Figure 1.4A) consists of a generator with a control unit for the desired settings and a probe transducer with a tip. The tip is submerged in the cell supsension and transmits its own oscillations, which are typically applied in pulses with defined ON and OFF cycles, to the sample. The pressure waves in the sample create bubbles, which collapse and generate shock waves rupturing the cell walls. As this process generates heat, the sample is cooled in an ice bath. Additionally, the ultrasonic waves cause a fluid flow in the sample known as acoustic streaming. Starting at the tip a liquid jet protrudes downwards, is reflected by the bottom of the tube, goes upwards again and turns near the vessel neck. Besides this circulation zone also dead zones can exist, where the fluid will stay idle (Figure 1.4B). As a proper mixing of the sample is anyway desired to avoid parallel over and under sonication in the same sample, Ferdous et al. [62] simulated the resulting circulation pattern depending on the tube diameters and the depth of the submerged tip in the sample. They obtained a circulation zone mixing over 90 % of the sample volume for a 50 ml Falcon tube filled with 15 ml cell suspension and a tip depth from the top of 20-30 %. Other tubes turned out to have also the greatest circulation zone for this tip depth, but did not reach 90 % due to higher length to diameter aspect ratios of the tubes (Figure 1.4C). On the other hand the generated heat can be better dissipated to the ice bath for a greater area to volume ratio, which is the smallest for the Falcon tubes. Ferdous et al. also concluded that the extent of cavitation (lysis efficiency) is dependent on the power input, but to high power densities should be avoided as they come along with a higher heat production. Also prolonged ultrasound exposure times causing the formation of hydrogen, hydroxide and peroxide radicals are not desired. In summary the activity of the final extract not only depends on the total energy input determing the lysis efficiency, but also on the rate at which the energy is supplied determing the activity of the extracted enzymes. These parameters can be regulated via the sonication amplitude, the pulse duration, and the total sonication time [62].
Ferdous et al. [62] validated their simulation results based on the earlier study of Kwon and Jewett [50], who developed a high throughput crude extract preparation method based on sonication using a 20 kHz device. In their study Kwon and Jewett already found a correlation between optimal cell-free protein synthesis yield and the sonication energy input depending on the volume of the cell suspension used. Small

volumes were found to be very sensitive to energy input, whereas larger volumes showed a higher tolerance.
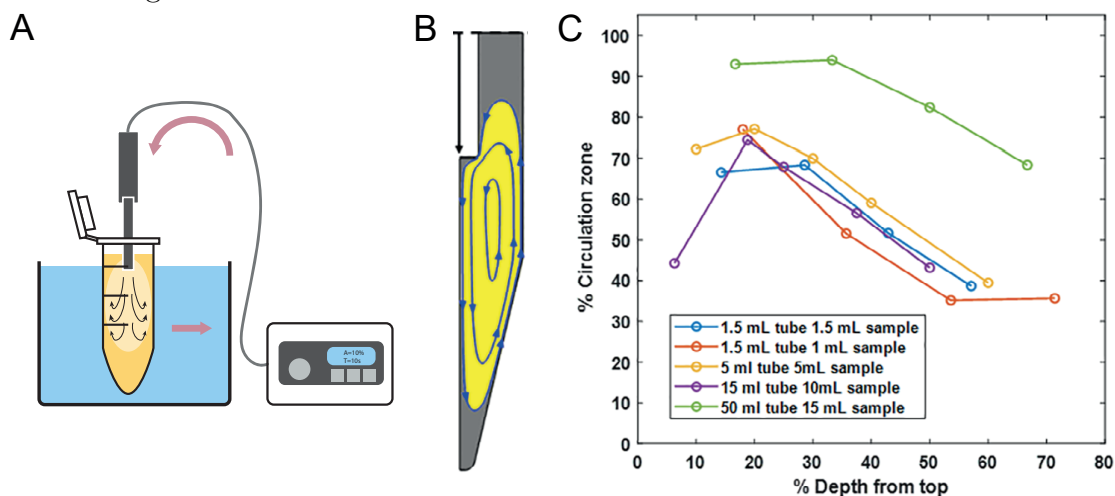
Figure 1.4: Tip sonicator. (A) The sonicator tip is submerged in the cell suspension, which is cooled in an ice bath. Pulse, amplitude and duration of the sonication process are set with the controller. (B) Sonication induces a fluid flow due to acoustic streaming. The resulting circulation zone (yellow) and dead zone (gray) depending on the sample tube dimensions and the depth of the submerged tip in the sample were simulated by Ferdous et al. [62] (C) They obtained the best result for a 50 ml Falcon tube filled with 15 ml cell suspension. Other tubes showed the greatest circulation zone for a tip depth of 20 % from the top. Figure A was reproduced and B-C were adapted with permission from [62].

Enzymatic cell wall degradation   In 1922 Alexander Flemming discovered an enzyme referred to as lysozyme, which degrades the petidoglycane layer of bacteria. Lysozyme is a small protein consisting of 129 amino acids crosslinked over four disulfide bonds and can be purified from chicken eggs [63].

Its substrates, peptidoglycanes are marcromolecules in the cell wall of gram positive and gram negative bacteria. They give the bacterium stability against the cells turgor from the inside. Gram negative bacteria like *Escherichia coli* contain just a about 6 nm thin layer of peptidoglycane, which is embedded between two layers of membranes, whereas gram positive bacteria like *Staphylococcus aureus* contain a thick layer in the range of 15-30 nm (outer wall zone) outside of a single membrane. Peptidoglycane (Figure 1.5) consists of glycane strands made up of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues. The D-lactyl group of each MurNAc residue is substituted by a peptide

stem, which is used for crosslinkage of parallel chains [64]. Lyzoszmye binds up to six residues of the sugar chains, whereby the third binding pocket is inaccessible for MurNAc. It catalyzes the hydrolysis of the glycosidic bond between the C-1 of MurNAc and the C-4 of GlcNAc located in the forth and fifth binding pocket [63].
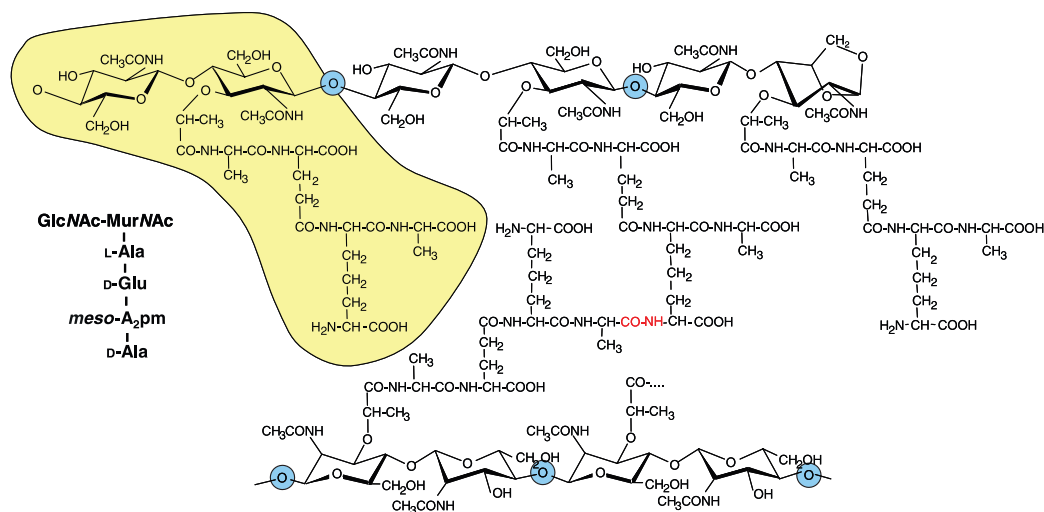
Figure 1.5: Structure of peptioglycan of *E. coli*. The glycane strands are made up of alternating GlcNAc and MurNAc residues. The basic subunit consists out of a disaccharid tetrapeptide (yellow), which is also shown with conventional amino acid and hexosamine abbreviations on the left-hand side. In the middle a cross-linkage between two glycane strands via the peptide stems is shown in red. Potential lysozyme hydrolysis sites are highlighted in blue. However lysozyme binds up to six residues of the sugar chains beginning with GlcNAc and hydrolyzes the glycosidic bond after the forth residue (not shown). The figure was adapted with permission from [64] (minor modifications were made).

The first attempts to produce cell lysates from *E. coli* via enzymatic cell wall degradation using 0.1 mg/mL lysozyme – without sonication – did not result in extracts with appreciable cell-free protein synthesis [60]. The peptidoglycan layer of gram-negative bacteria is not very accessible to the enzyme, but Irvin et al. [65] demonstrated that appropriate buffer conditions during cell lysis with lysozyme can enhance the accessibility. In addition, Fujiwara and Doi [66] showed that a protocol involving lysozyme incubation combined with osmotic shock and freeze-thaw cycles does result in an active cell extract.

1.3 Gene expression yield in cell-free systems

Talking about cell-free gene expression within cell extracts, it is an important question how to determine the quality of the extract and which effects can have influence on the expression yields. As a standard measure usually the total protein content of a produced cell extract is determined using a Bradford or a bicinchoninic acid (BCA) assay or alternatively the protein composition is analyzed using state-of-the art quantitative proteomics [67]. Despite the fact that this methods gives a relative estimation about the lysis efficiency, the activity of the extracted enzymes can not be evaluated. For this reason typically also expression tests are performed. In the past a protein of interest was chloramphenicol acetyltransferase (CAT) [68], which activtiy and quantity was read out over radioactively labeled substrates and spectrophotometric measurements respectively. Nowadays mainly fluorescent proteins are expressed to determine the expression yield, which can be also combined with fluorescent aptamers to have a parallel read out of mRNA levels [69]. As over 100 different components are involved in the transcription and translation process [25], it is essential to have a basic knowledge of the gene expression kinetics within cell-free systems and factors, which have influence on it.

Kinetics in CFS

$$\frac{d}{dt} \text{mRNA(t)} = k_{TX} [\text{DNA}] - k_{deg} \, \text{mRNA(t)}$$

$$\frac{d}{dt} \text{prot}_{unfold}(t) = k_{TL} \, \text{mRNA(t)} - k_{mat} \, \text{prot}_{unfold}(t)$$

$$\frac{d}{dt} \text{prot}_{mat}(t) = k_{mat} \, \text{prot}_{unfold}(t)$$

Steady state $\dfrac{d}{dt} \text{mRNA(t)} = 0$:

$$\text{mRNA(t)} = \frac{k_{TX} [\text{DNA}]}{k_{deg}}$$

$$\frac{d}{dt} \text{prot}_{unfold}(t) + \frac{d}{dt} \text{prot}_{mat}(t) = \frac{k_{TL} \, k_{TX} [\text{DNA}]}{k_{deg}}$$

for $k_{mat} t, k_{deg} \, t \gg 1$

k_{TX}	Transcription rate	mRNA	mRNA concentration
k_{deg}	mRNA degradation rate	DNA	DNA concentration
k_{TL}	Translation rate	prot_{unfold}	Unfolded protein
k_{mat}	Protein maturation rate	prot_{mat}	Mature protein

Table 1.1: Simple model for TXTL reactions suggested by Siegal-Gaskins et al. [69].

Cell extract preparation causes both a dilution of the system by a factor of about 30 (protein concentration in *E. coli*: 300 mg/ml vs. in TXTL reaction : 10 mg/ml) and a partial inactivation or degradation of the extracted enzymes. Additionally membrane- or DNA-bound enzymes may be partially or fully lost during the extract preparation [67]. As a consequence the TXTL machinery concentration and rates are 1-2 orders of magnitude lower than *in vivo* [25].

In the ideal TXTL model (Table 1.1) with unlimited resources and unchanging reaction conditions (stable pH, no hydrolysis of NTPs, no degradation of enzymes, ...) the mRNA and protein concentration dynamics can be described by a set of ordinary differential equations with an exact solution.

Based on this model one would expect a constant RNA concentration in the steady state, which depends on the transcription and degradation rate as well as the DNA concentration. The actual course of the RNA concentration shows at the beginning Michaelis-Menten kinetics but the RNA synthesis stops after some time and one observes an exponential decay of the mRNA concentration [69]. The lifetime of RNA in TXTL samples is in the range of 12-18 min [25][69], but can be increased up to

16 h [70] via strong secondary structures, which protect from degradation.

The protein synthesis rate (premature and mature) is expected to be proportional to the mRNA concentration and therefore approach a constant in the steady state. Indeed this is a good estimation for the protein synthesis in the first hour, for later times the proportionality is lost [69] and the protein synthesis rate decays exponentially [25]. Protein degradation caused by the ClpXP complex exhibits zeroth-order kinetics and not first order kinetics as observed in cell populations [25].

For the total expressed protein concentration one can derive a proportional relationship to the plasmid concentration. However if a certain threshold DNA concentration is reached known as DNA capacity the linear dependency is lost in real samples. The maximum DNA capacity increases with decreasing promoter strength, but does not exceed 32 nM [69]. The expression levels also changes if the load on the TXTL system is increased by expressing more than a single protein. Even though there might be no direct coupling between the genes, the resources (NTPs, AA, TXTL machinery, ...) have to be shared resulting in a decrease in expression yield for each product. Siegal-Gaskins et al. [69] additionally revealed, that the variation (decrease) in RNA concentration under increasing load caused by a second reporter gene is similar for all tested template concentrations in the range of 1-10 nM. In contrast to this, the variation (decrease) in reporter expression (on the protein level) strongly depends on the template concentrations of both reporter genes. Small template concentrations turned out to show a low, high template concentrations a strong decrease in protein expression under increased load. This suggests that the translational resources are more limiting to the system performance.

Also a strong influence on gene expression yield of CFS have context dependencies of basic parts and modules, which are disscussed in the following section.

Context dependency of basic building blocks Despite the fact, that context dependencies also arise *in vivo*, only cell-free studies are discussed in this paragraph.

As mentioned at the beginning, cell-free gene expression systems have a wide range of applications, which are implemented in the optimal case with standardized and well characterized parts. In especially complex genetic networks need varying expression levels of single circuit components, which can be established either by adjusting the according DNA template concentration or by choosing parts with the desired characteristics (e.g. promoter strength). Unfortunately, while Chizzolini et al. [71] were able to tune the RNA expression level with the promoter strength in a T7 RNA

polymerase dependent system, they observed an increased variation in RNA levels for TXTL reactions based on the *E. coli* RNA polymerase. The difference is a result of the higher degree of complexity, which arises from the assembly process of the *E. coli* RNA polymerase from multiple subunits. Additionally the protein expression level depends not only on the strength of the ribosome binding site, but also on factors like the codon usage and the secondary structure of the mRNA, which may interfere with the multiple factors involved in translation initiation, elongation and termination. To overcome these context dependencies basic building blocks can be separated from each other using insulator parts. These can be standardized sequences up- and downstream of the -35 and -10 box of promoters respectively [72] or standard translation initiation elements [73]. It was also tested successfully to physically separate (cut) the 5' untranslated region of mRNAs by using ribozymes [74] or CRISPR [75]. Context dependency is an effect, which is not only observed for basic building blocks, but also for connected functional modules (greater assemblies, which are tested in isolation and then combined). The connection to a downstream system cause an additional reaction flux on molecules needed in the upstream process and therefore a change on the system dynamics, which is known as retroactivity. [76]

Regardless of the part and module characteristics also the composition of the TXTL buffer is crucial for a successful cell-free gene expression. Molecules below 10 kDa like nucleotides and amino acids are lost in the dialysis step during cell extract preparation and have to be added to the TXTL reaction again. Usually also secondary energy sources are included to enable energy regeneration, which turned out to be crucial for gene expression yield.

1.4 Energy regeneration pathways

In early studies protein synthesis in cell-free systems stopped after 20 min, which was the result of rapid depletion of adenosin triphosphate (ATP). Secondary energy sources such as phosphoenol pyruvate (PEP), creatine phosphate and acetylphosphate containing a high-energy phosphate bond in combination with appropriate kinases were then used for ATP regeneration. However these compounds are expensive and can be degraded in non-specific (uncoupled) phosphatase activities. As a result inorganic phosphate is accumulated, which inhibits protein synthesis by chelation of free magnesium ions [77][78]. To partly counteract this problem one can add inorganic phosphate and glucose to the culture medium and thereby suppress the expression of phosphatases during cell growth [79].

To avoid the inorganic phosphate accumulation from the beginning, Kim and Swartz tested pyruvate for the ATP regeneration. In their first approach [80], they added exogenous pyruvate oxidase to the system to convert pyruvate directly to acetylphosphate (Figure 1.6). Acetylphosphate can be further processed using the endogenous acetate kinase producing ATP and acetate and therefore has besides the ATP regeneration an negative effect on the pH in the sample. The protein yield could be further increased by inhibiting with oxalate a different pathway in which pyruvate is converted to PEP (this reaction consumes a ATP and converts it to AMP) [68]. However the addition of an exogenous enzyme and the need for molecular oxygen in the reaction makes it less interesting for applications.

In their next approach [77], they demonstrated, that pyruvate can also be metabolized under addition of the cofactors NAD and CoA making advantage of the endogenous enzymes pyruvate dehydrogenase (producing the intermediate acetyl-CoA under reduction of NAD) and phosphotransacetylase (producing again acetylphosphate). The NAD is regenerated under formation of lactate (from pyruvate), which affects together with the produced acetate the pH negatively. The expressed protein level under addition of pyruvate reached just about 70% of the level in presence of PEP, but they demonstrated that each intermediate during glycolysis starting with glucose-6-phospate to pyruvate can be used for energy regeneration. The combination of these results is known as PANOx system (PEP, amino acids, NAD, oxalic acid), which benefits from the two step conversion of pyruvate via NAD.

In their study they obtained poor results for glucose as energy source, potentially because the glucose-specific PTS enzyme II responsible for the conversion of glucose to glucose-6-phosphate upon glucose uptake in a cell, is membrane-bound and therefore lost during cell extract preparation (centrifugation steps) [67]. However Kim et al. [81] demonstrated the use of glucose as energy source with no need for addition of cofactors within a S12 extract (has no high speed centrifugation and dialysis step). One glycolysis intermediate two steps above PEP namly 3-PGA turned out to be a better alternative. Under addition of 3-PGA PEP is continuously synthesized at low levels and the unspecific degradation is avoided [42].

Also the disaccharide maltose was tested as a energy source. During its metabolism not only accumulation of inorganic phosphate is avoided but also its recycling is enabled. Maltose activates the glycolysis pathway and has an inhibitory effect at concentrations over 15-20 mM as again the produced acids lower the pH. In presence of maltose up to 10 h of protein synthesis were reported, which is by a factor of 2 higher compared to samples without maltose. [30]

Maltodextrin together with maltodextrin phosphorylase and phosphoglucomutase can also be added to a TXTL reaction. Maltodextrin is slowly cleaved into glucose-

1-phosphate (G1P) in the presence of inorganic phosphate (no ATP is consumed to produce G1P). G1P is further converted into G6P by phosphoglucomutase. It was shown that TXTL samples with this energy source produce higher protein yields than PEP, glucose, and glucose-6-phospahte. These TXTL samples also exhibited better controlled phospate levels and the smallest decrease in pH from 7.6 to 6.9 compared to the other three compounds. This polysaccharide approach was further improved by utilizing starch and glycogen, which need no addition of exogenous enzymes and resulted in a homeostatic maintenance of the reaction conditions and a protein synthesis over 12 hours yielding 1.7 mg/ml expressed protein [82].

It was also observed that glutamate, a intermediate of the TCA cycle, can be used as a energy source, which is converted in a oxidative phosphorylation potentially in inverted membran vesicles, which are generated in lysis methods with high shear forces (eg. high pressure homogenization) [83].
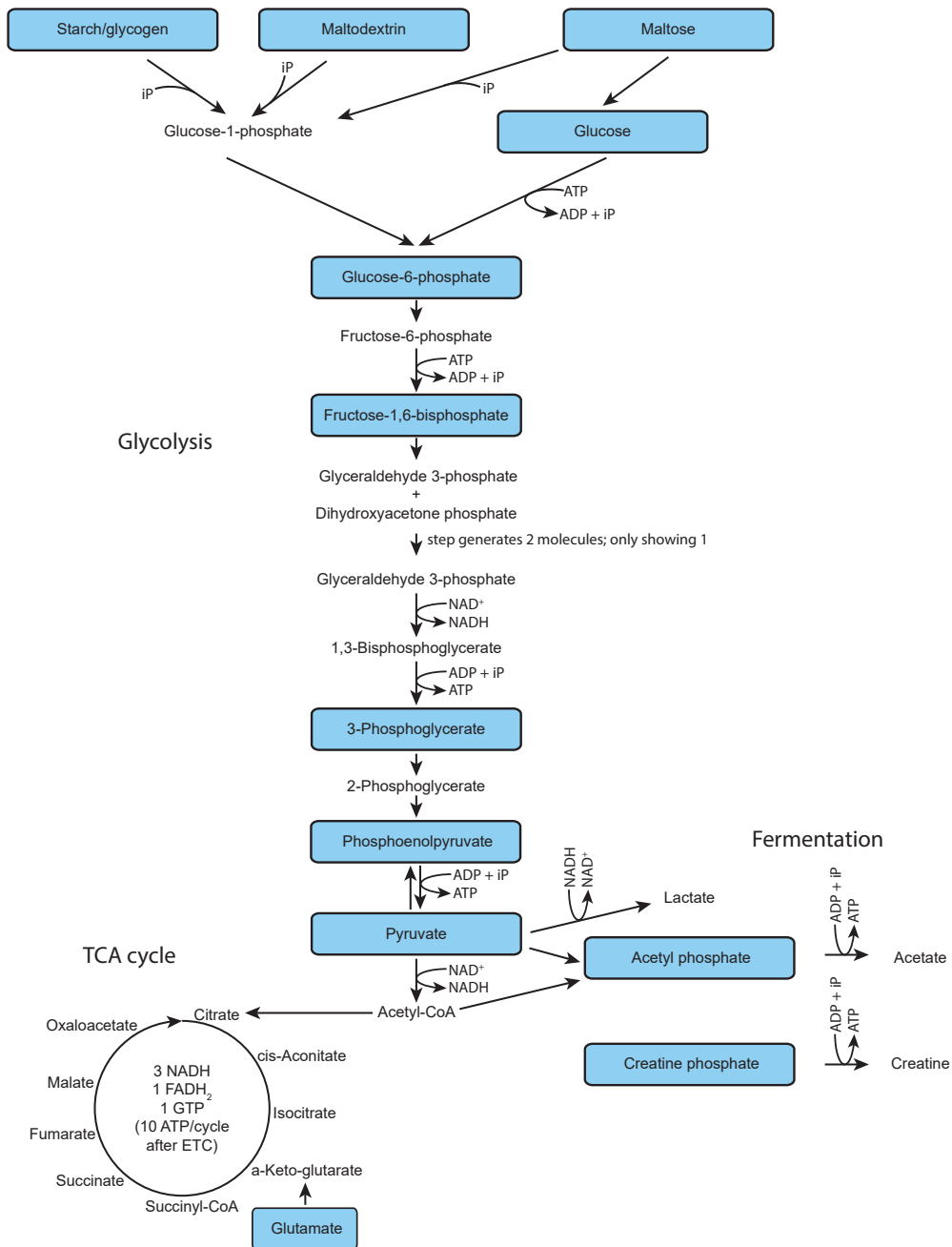
Figure 1.6: Simplified scheme of *E. coli* metabolism. Shown are glycolysis, TCA cycle and fermentation and highlighted in blue are energy sources, that have been tested for cell-free reactions (ETC electron transport chain). The figure was reproduced with permission from [83].

Chapter 2

Regulatory mechanisms

So far most of the implemented genetic circuits were constructed using protein-based regulatory mechanisms with response times set by the cell doubling time. The circuit response time can be sped up using RNA regulatory elements, which propagate the signals directly as RNAs and eliminate expression of intermediate proteins. The circuit response time then depends on the degradation time of the RNA. RNA circuits benefit from potentially simpler design rules, as they rely on predictable base-pairing interactions and can be characterized using qPCR and next-generation sequencing [84]. A huge variety of regulatory RNAs were discovered in the past and their regulatory mechanisms in eukaryotes and prokaryotes are presented in the following section to evaluate their applicability for the construction of synthetic circuits.

2.1 RNA-based regulatory mechanisms

Even though just a small fraction of the human genome directly codes for proteins [85], it has been found that it is almost completely transcribed, but the expression level of single gene products is cell- or tissue-specific. So besides the well known noncoding RNAs (ncRNAs) with house-keeping or infrastructural roles (rRNA, tRNA), there exist a huge variety of other ncRNAs, which have regulatory functions. They can be divided in different classes depending on their length (small and large ncRNAs) and function mechanism [86].

Small RNAs Small RNAs, which include microRNAs, small interfering RNAs (siRNAs) and Piwi-interacting RNAs with a length between 18 to 30 nucleotides,

can regulate gene expression via post-transcriptional gene silencing, chromatin-dependent gene silencing or RNA activation. They are supposed to regulate more than 30 % of genes in a cell and each sRNA can have multiple RNA targets, which are bound imperfect or perfect via base-pairing. Small RNAs have important roles in development and disease via regulation of cell differentiation, growth/proliferation, migration, apoptosis/death, metabolism and defense (i.e., CRISPR) [87] [88] [89].
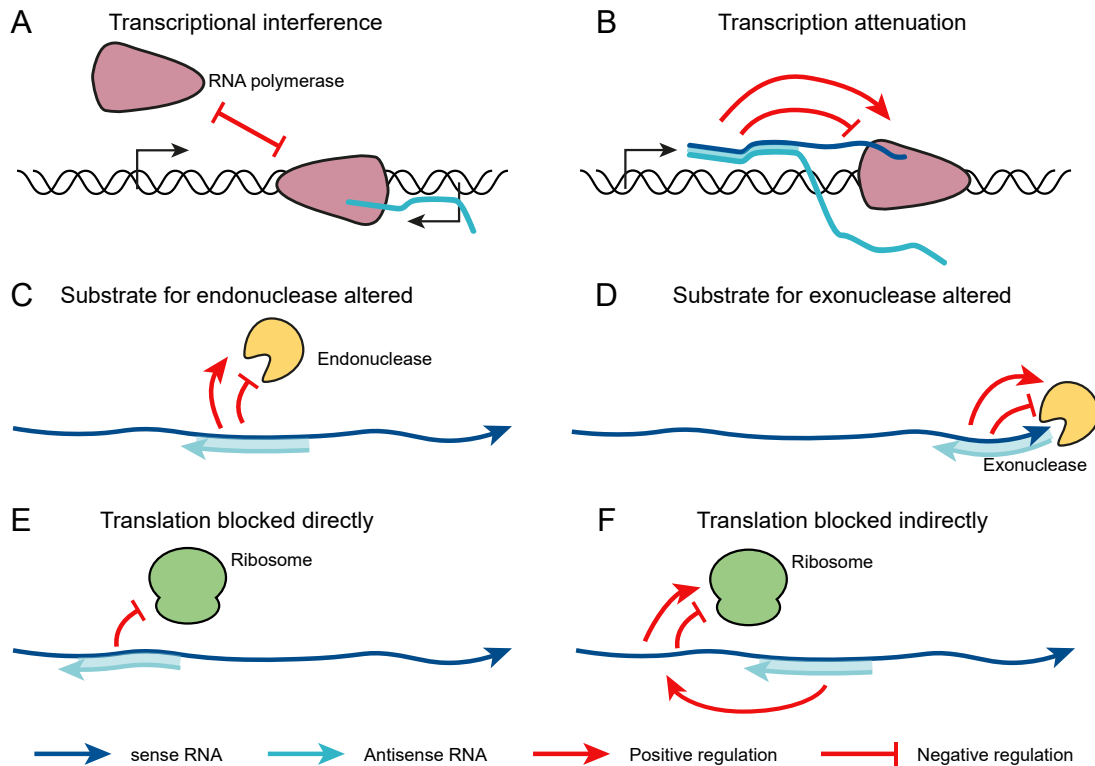
Figure 2.1: Regulatory mechanisms of RNAs with RNA targets. (A) *Cis*-encoded antisense RNA can cause transcriptional interference effects. (B) Base pairing of the sRNA to the target RNA can cause secondary structure changes, which result in transcription termination. (C-D) Hybridization of a sRNA to its target can influence the target RNA stability by generating or blocking endo- or exonuclease sites. (E) sRNA can bind the Shine Dalgarno sequence of the target RNA and thereby block translation directly. (F) By altering the target secondary structure translation can also be indirectly influenced. The figure was reproduced with permission from [90].

Cis-acting and trans-acting RNAs   Other noncoding RNAs with RNA targets, but unspecific length ranging from ten to thousand nucleotides, can either be *cis*-acting,

referred to as antisense RNAs, or *trans*-acting. Antisense RNAs are encoded on the DNA strand opposite to a gene and therefore perfect complementary to their target. They either overlap partially (3' or 5' end or somewhere in between) or fully with their target genes. If both antisense and target RNA are transcribed simultaneously, the regulatory function can also rely on the base-pairing independent transcriptional interference effect (see section 2.2 for details). In contrast, *trans*-acting RNAs usually have several binding targets and require often the RNA chaperone protein Hfq for base pairing as it is the case for the regulation of quorum sensing in *V. choerae* [91] or iron metabolism in *E. coli* [92] [90].

For both *cis*- and *trans*-acting RNAs the regulatory function can be promoted via various mechanisms (Figure 2.1). On the transcriptional level the RNA-RNA interaction can result in the formation of a terminator structure and therefore cause transcription attenuation. Post-transcriptionally, interactions with the translational machinery [93], or the regulation of alternative splicing [94] were reported. Also the modulation of the stability of their targets by generating or blocking endo- or exoribonuclease sites is conceivable. Alternatively the regulatory function can rely on either the binding to the Shine Dalgarno sequence or on a change in secondary structure and an indirect modulation of the accessibility to the RBS [90].

Regulatory mechanisms between RNA and proteins Long noncoding RNAs, small nucleolar RNAs and untranslated mRNA regions can also have proteins as their binding targets. Besides the ribosome and t-RNA-synthases about 5 % of the human proteom consists of RNA binding proteins [95], but they do not only receive from the RNA but also transmit to the RNA the regulatory effect. For example it was recently discovered for mammalian cell lines, that several metabolic enzymes involved in glycolysis, the tricarboxylic acid (TCA) cycle, lipid metabolism, and deoxynucleotide biosynthesis are RNA binding proteins [96]. Among these the thymidylate synthase autoregulates its own expression by binding at low substrate concentrations its own mRNA [97]. For the second interaction direction several mechanisms are conceivable. The bound RNA can partially or fully block the active site or cofactor binding pocket of the target protein. Alternatively it can bind to a different region of the protein and cause an allosteric effect. Thereby the enzymatic function or the interaction of the enzyme with structural elements such as membranes can be modulated. This was reported for protein kinase R, which is triggered to dimerize upon binding of a pathogen-drived RNA and is thereby activated [98]. A RNA can also bridge subunits or larger assemblies and thereby enable a superior metabolic flux as it was shown for complexes of glycolytic enzymes being sensitive to RNase digestion [99][96][100].

Riboswitches Other regulatory RNAs can sense metabolites via riboswitches. These are mRNAs containing an aptamer at their 5' untranslated region. Upon binding of a small ligand the aptamer undergoes conformational changes that lead to a change in translation initiation (the RBS is revealed or blocked, Figure 2.2A) [101], to transcriptional termination [102] or to a cleavage of the RNA [103][104]. Naturally occuring riboswitches serve as control elements for metabolic pathways sensing metabolites like coenzyme B1 [101], FMN [102], lysine [105] and guanine [106]. An example for a rationally designed riboswitch is the theophylline riboswitch, which is a combination of the theopylline aptamer and a helical communication module taking advantage of a ligand-dependent one-nucleotice slipping mechanism [107].

Regulatory RNAs as components for synthetic circuits Regulatory RNAs are increasingly used also as components for synthetic biological circuits [108]. As their function is mainly determined by predictable nucleic acid base-pairing interactions, they can be engineered more straightforwardly than protein-based regulators. One of the first examples for a synthetic post-transcriptional riboregulator was developed by Isaacs et al. [109]. Their riboregulators are mRNA strands, in which the ribosome binding site is sequestered within a hairpin structure, resulting in a strong reduction of translation initiation. In the presence of a RNA molecule partly complementary to the stem of the hairpin, referred to as *trans*-acting RNA, the hairpin opens and translation is activated (Figure 2.2B). However, this leads to a limited number of regulators due to sequence constraints and high crosstalk levels up to 20 % [110]. Building upon this work, Green and co-workers [111] later developed rationally designed riboregulators with a considerably higher dynamic range and high design flexibility. In their design, the RBS is hidden in the loop of an RNA hairpin, while the start codon is included as a bulge in its stem (Figure 2.2C). A single stranded toehold sequence precedes the hairpin structure at its 5' end. This toehold facilitates a toehold-mediated strand displacement reaction by the *trans*-activating trigger RNA (Tr), which reveals the RBS and start codon and therefore activates translation. TS can be deactivated again using an anti-trigger (AT) complementary to Tr. AT hybridizes to unbound Tr or displaces Tr from the TS, in case Tr and AT share a common toehold region. Toehold switches can be engineered to recognize different trigger RNAs as inputs, which already has led to several applications in

medical diagnostics [112]. Also logical negation (NOT gates) were implemented using toehold-switches [113].

In an alternative approach, Lucks and coworkers engineered naturally occurring transcriptional attenuators to obtain a set of inducible riboregulators termed STAR (small transcription activating RNA) regulators. In STAR mechanisms, either a terminator or an anti-anti-terminator sequence is put upstream of a gene. Inducible expression of an antisense (STAR) molecule then induces a conformational change in the terminator (or anti-anti-terminator), resulting in anti-termination and thus activation of transcription [114][115].
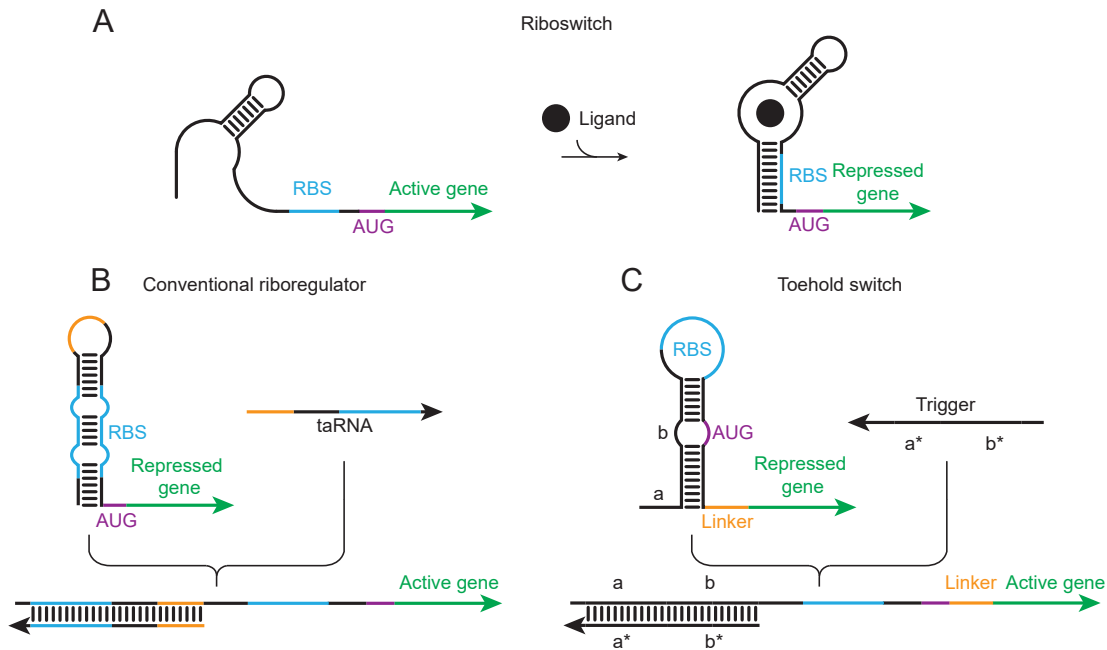
Figure 2.2: Riboswitches and riboregulators. (A) An aptamer located at the 5' untranslated region of a mRNA undergoes structural changes upon binding of a small ligand. Thereby the RBS is sequestered and translation is repressed, but also other mechanisms can apply (not shown). (B) In a conventional riboregulator base pairing to the RBS and start codon inhibit translation. A *trans*-acting RNA can open the hairpin and activate translation. In the illustrated example this is enabled via loop-linear interactions (orange highlighted). (C) In a toehold switch the RBS is located at the loop and the start codon at a bulge region in the stem of a hairpin structure. A toehold sequence is located at the 5' end of the toehold switch. Translation is activated in presence of a trigger RNA via a toehold-mediated strand displacement reaction. B and C were reproduced with permission from [111].

2.2 Transcriptional interference

Until now the performance of synthetic circuits was tuned by exchanging basic parts by ones with suitable characteristics (promoter strength, translation rate of RBS, ...). In our approach we want to evaluate whether one could take advantage of transcriptional interference effects to tune the circuit performance.
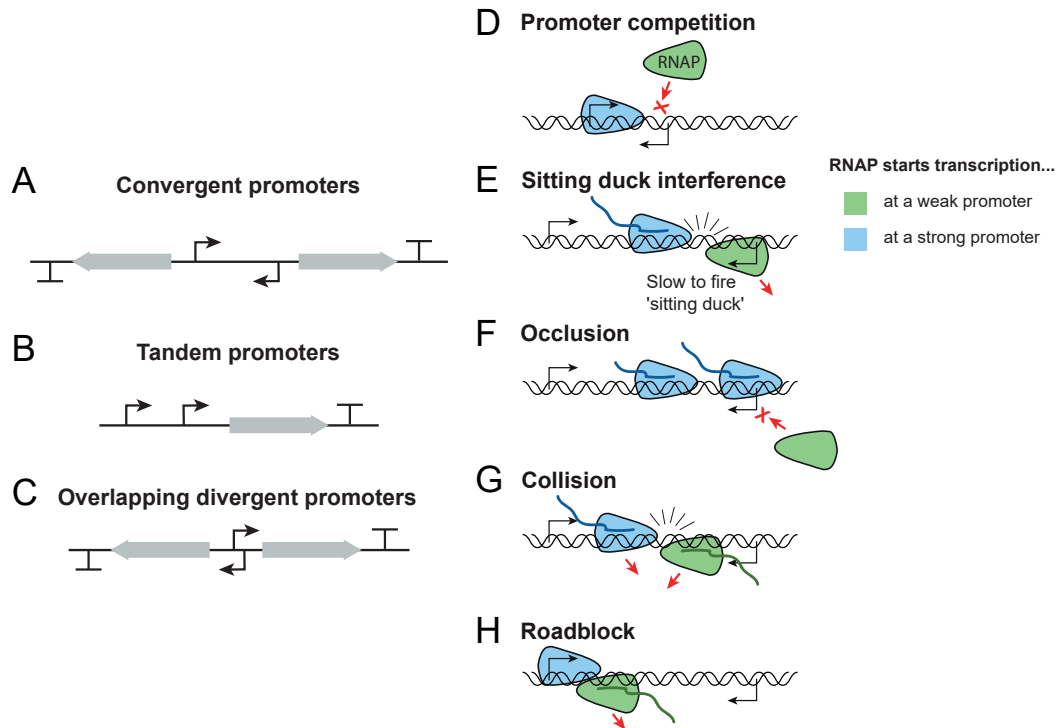
Figure 2.3: Different promoter arrangements and transcriptional interference effects. (A) Convergent promoters are face-to-face oriented. The coding sequence is fully or partially between the promoters. (B) Tandem promoters point in the same direction. The transcripts do not necessarily overlap. (C) Divergent promoters point in opposite directions. The promoter regions can overlap or be separated. (D) Overlapping promoters or promoters in close distance compete for the binding of a RNAP to the promoter sites. (E) A RNAP starting at a weak promoter, which is slow in escape of the promoter region, can be dislodged by a second polymerase. This effect is referred to as sitting duck interference. (F) A transcribing polymerase can occlude the binding of a second polymerase to a promoter region. (G) Converging RNAPs can collide and can cause a termination of one or both transcription processes. (D-H were reproduced with permission from [116]).

Transcriptional interference (TI) is defined as the direct negative influence of one

transcriptional process to a second one *in cis*. Usually this effect is asymmetric, as the strong (aggressive) promoter regulates the weak (sensitive) promoter down. TI is surmised to fulfill important functions in many cellular processes, both in prokaryotes and eukaryotes. Some examples were already revealed, these include e.g. zinc homeostasis [117] and entry into meiosis [118] in *S. cerevisiae*, embryonic developement in *D. melanogaster* [119], regulation among the murine β -like globin genes [120] and T-cell receptor alpha recombination [121] and in the orientation-dependent regulation of integrated HIV expression [122] in mammals. TI may also play an evolutionary role through its effect in head-on collisions between replisomes and RNAPs, which potentially increases the evolvability of convergently oriented genes through mutagenesis [123]. Moreover, several metabolic engineering tasks have been successfully implemented via convergent promoter designs [124] [125] [126].

Different mechanism apply depending on the orientation of the promoters and the promoter characteristics. Possible promoter arrangements are convergent, tandem or divergent promoters (Figure 2.3A-C). Convergent promoters are face-to-face oriented and overlap fully or partially in their transcripts. It is conceivable for overlapping transcripts that the TI effect arises due to tethering both RNAPs together via base pairing of the complementary transcripts. Callen et al. [127] could reject this hypothesis by inserting a RNAseIII site in the region, where the transcripts are supposed to base pair. The cutting during transcription showed just a small reduction of the TI effect. Tandem promoters transcribe in the same direction (co-directional) and their transcripts possibly but not necessarily overlap. Divergent promoters point in opposite directions (back-to-back). A special case are overlapping promoters, which can be of either type and share at least a part of their DNA sequence [116].

Several TI mechanisms are known (Figure 2.3D-H), which arise during different stages of transcription. The binding of a RNAP to the weak promoter can be hindered, which is either a result of promoter competition or occlusion. The **competition** effect is observed in close or overlapping promoters, where the binding of a RNAP to one promoter precludes the occupation of the second promoter. This effect also arises, if two promoters share the same enhancer site, but can be eliminated if the promoters are separated by a higher distance or a second enhancer site is introduced [116]. Similar to this is a so called **local RNAP concentration effect**. Active promoters increase the local RNAP concentration around their promoter sites. If two promoters are close they can in principle benefit from each other. However the stronger promoter acts as a local sink for the transcriptional machinery and is able to recruit the major fraction of the available polymerases, which regulates its own activity up and the weak promoter down [128][127].

In contrast **occlusion** was introduced to describe the phenomenon, that an elongat-

ing RNAP over a second promoter region limits the available time for RNAPs to bind to this region. The extent of interference, which is in most cases minor compared to other mechanisms, depends on the length of the DNA sequence of the second promoter, the promoter strength of the strong promoter and the passage time over the second promoter [116]. The latter can be increased by inserting pause elements as it was reported by Palmer et al [129].

Another interference effect, referred to as **sitting duck interference**, is known for open complex formation. A polymerase, which has already bound a promoter site but has not started elongation yet, can be hit by an elongation complex and thereby dislodged. This effect is the dominant mechanism for close but not overlapping promoters and is maximal if the ratio between open complex formation rate and transition rate from open to elongation complex is equal 1. Otherwise the dislodgement is unlikely as the open complex transits fast to the elongation complex (< 1) or the dislodged complex is rapidly replaced by another one (> 1). [116] So both, the occlusion and the sitting duck interference effect are based on the passage of a RNAP over the weak promoter region and therefore can be stopped by inserting a transcriptional terminator upstream of the second promoter.

For elongating polymerases two TI mechanisms exist: collision and roadblock effect. While one would not expect any collision effect in tandem promoters due to the co-directional transcription and the cooperativity of trailing RNAPs [130], the situation changes for convergent designs. The **collision** of converging elongation complexes can cause for one or both of the transcription processes a stalling or the premature termination of the polymerase. The exact mechanism is not known, but the interference may be mediated through DNA supercoiling rather than a direct collision of the transcriptional machinery [131]. It is conceivable, that stalled polymerases simply fall off, are rescued from host factors or resume transcription after a cooperative push through enabled by trailing polymerases. Collisions become more likely with increasing promoter distances and higher promoter activities [116].

The assumed requirements on promoter characteristics (promoter strength, and the aspect ratios of the RNAP binding rate and the firing rate) and distance for a efficient regulation by TI effects were confirmed by Sneppen et al. [132]. They introduced a mathematical model for TI effects based on the occlusion, collision and sitting duck mechanism. Connor et al [131] demonstrated that the collision effect can be tuned up to 38-fold through processitivity (uninterrupted transcription) control. They tested the antibiotic bicyclomycin and the phage protein Psu to inhibit Rho dependent transcription termination of untranslated regions initiated by NusG. By doing this the RNAP can collide with its converging pendant and cause an increased TI effect instead of being removed. NusG would usually bridge a RNAP and a ribosome dur-

ing the pioneering round of transcription, but recruits in absence of a co-translating ribosome Rho. Therefore also tuning of co-translation is valid target for processivity control.

Also well known as a TI mechanism is the **roadblock** effect. In principle all DNA binding proteins can act as a roadblock for a elongating polymerase, which then backtracks and falls off or is removed by the transcription-repair coupling factor Mfd [133]. Alternatively it again resumes transcription and reads through under cooperation with trailing RNAPs. To sucessfully read through a roadblock some ploymerases have to be accumulated in a row, which requires a long distance of the roadblock to the promoter and a high RNAP flux provided by a strong promoter. Whereas this effect is well known for the lac repressor, is is unlikely for an open complex to act as a roadblock [116] [128][134].

Recently, TI effects have gained the interest of synthetic biologists as a potentially engineerable tool to manipulate RNAP traffic on DNA [128] and to tune gene expression levels [135]. Using a variety of synthetic gene constructs, Bordoy et al. [128] investigated TI effects of convergent promoter designs in detail to assess their tunability and used them to regulate gene expression over several orders of magnitude. In a different study a series of two-input genetic devices that use the presence of a roadblocking protein to control gene expression was created [136]. On the downside, unwanted transcriptional interference also poses a design challenge for synthetic circuits of increasing complexity [137].

2.3 Overview of the thesis

This thesis includes both a cell-free and an *in vivo* synthetic biology project.

In the cell-free project we show the production of active cell extract for cell-free gene expression based on a combination of lysozyme incubation and sonication cylces for cell lysis. Compared to Shrestha et al. [60] we increased the lysozyme concentration and modified the reaction conditions for the first lysis step. For the second lysis part we tested low energy densities (low amplitude) and short pulse durations as recommended by Kwon and Jewett [50]. We determined the best settings in screening experiments (lysozyme concentration against sonication cycle number) for shaking flask and bioreactor cultivated cells and introduced four quality measures: The total protein content of the extract (prior to gene expression) determined using BCA assays helped us to asses the relative lysis efficiency between different extracts. In order

to reveal changes in protein composition caused by the different lysis settings and cultivation methods we performed also state-of-the-art quantitative proteomics analyses. Additionally, we expressed four different fluorescent reporter proteins, which gave information about the dependence of the gene expression yield on the mRNA and protein characteristics. To demonstrate the applicability of the produced extract for more complex systems, we also produced and assembled functional T7 bacteriophages.

In the *in vivo* project, we tested toehold switches and its cognate triggers and anti-triggers within different promoter arrangements and evaluated the effect of transcriptional interference on the system performance. In detail, we encoded a toehold switch controlling the expression of fluorescent reporter protein in *cis* with its cognate trigger (Tr) and anti-trigger RNA (AT). Whereas the toehold switch was expressed constitutively, Tr and AT were under the control of inducible promoters. In our first attempt we tested several tandem designs of the Tr and AT transcription cassettes with varying promoter distances and determined threshold distances for which local RNAP concentration effects arise. In the next step we took advantage of the sequence complementarity of Tr and AT and tested convergent designs with and without an excess of AT. Based on the results we combined the best performing designs and implemented a toehold switch-based two-input two-output logic gate in *E. coli*.

Chapter 3

Lysozyme-assisted sonication cell extract

Evaluation of an E. coli cell extract prepared by lysozyme-assisted sonication via gene expression, phage assembly and proteomics

E. Falgenhauer, S. von Schönberg, C. Meng, A. Mückl, K. Vogele, Q. Emslander, C. Ludwig, and F. C. Simmel

In this chapter, we show that lysozyme treatment with higher lysozyme concentrations and modified reaction conditions compared to Shrestha et al. [60] can be used to support a gentle sonication protocol. In particular we avoided an additional incubation step of the cell suspension at a physiologically relevant temperature (which can have a counterproductive result) by performing the lysozyme incubation on ice instead at 37 °C. As a result, the gene expression capability of our extract was considerably improved and comparable to a commercially available extract. In our experiments, we found that fluorescent proteins can be expressed from a constitutive promoter up to a concentration of 22 µM, which corresponds to 0.6 mg/ml. Using a different promoter with higher promoter strength can further increase the gene

expression yield. As another example, we produced and assembled functional T7 bacteriophages in our extract, which resulted in higher phage titers than with any other in vitro transcription/translation system we tested. In order to rationalize our observations, we analyzed selected extracts using state-of-the-art quantitative proteomics and focused on differences in protein composition resulting from different culture conditions and different lysis methods. Our analysis indicates a better release of DNA binding enzymes with increasing numbers of sonication cycles. Major differences were detected compared to a commercial extract, which is of particular interest as the extracts show a similar performance in cell-free gene expression experiments.

3.1 Results and discussion

3.1.1 Lysozyme-assisted sonication (LAS) as preparation method

An overview of our general workflow for lysozyme/sonication-based preparation of cell extract is shown in Figure 3.1. Cells are cultivated either in shaking flasks or a bioreactor and then harvested at the desired OD. The cell pellet is washed and the pellet mass is determined. The most important deviation from other protocols is the lysis step. Cells are incubated with lysozyme and afterwards subjected to several sonication cycles. The cell extract is clarified in centrifugation steps and further processed by perfoming a run-off reaction and a dialysis step. A detailed protocol is given in the methods section.
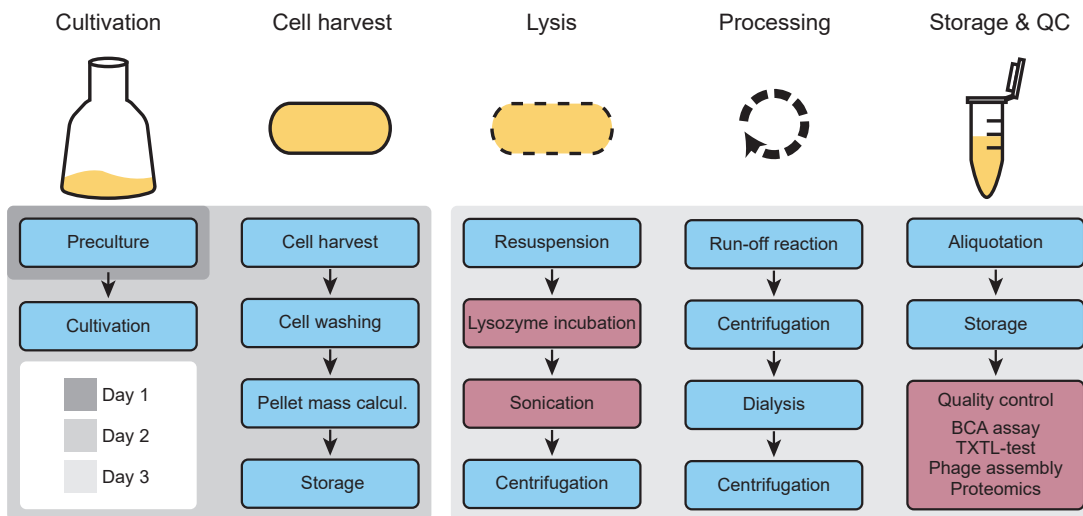
Figure 3.1: Overview of the cell extract preparation workflow. Cells cultured in shaking flasks or a bioreactor are harvested and washed. The pellet mass is determined and the cells are lysed via lysozyme incubation and sonication cycles. To asses the quality of the produced extract, the total protein content is determined, reporter proteins and phages are produced and proteomics analyses are performed.

3.1.2 Comparsion of shaking flask and bioreactor cultivation

Batch	Slope (10^{-4} 1/min)	Doubling Time (min)	k (1/h)
Shaking Flask	183± 4	16.4 ± 0.1	2.53 ± 0.06

Batch	Slope (10^{-4} 1/min)	Doubling Time (min)	k (1/h)
BR1	93 ± 3	32.3 ± 0.3	1.29 ± 0.04
BR2	95 ± 2	31.8 ± 0.2	1.31 ± 0.03
BR3	80 ± 7	37.7 ± 0.9	1.10 ± 0.09
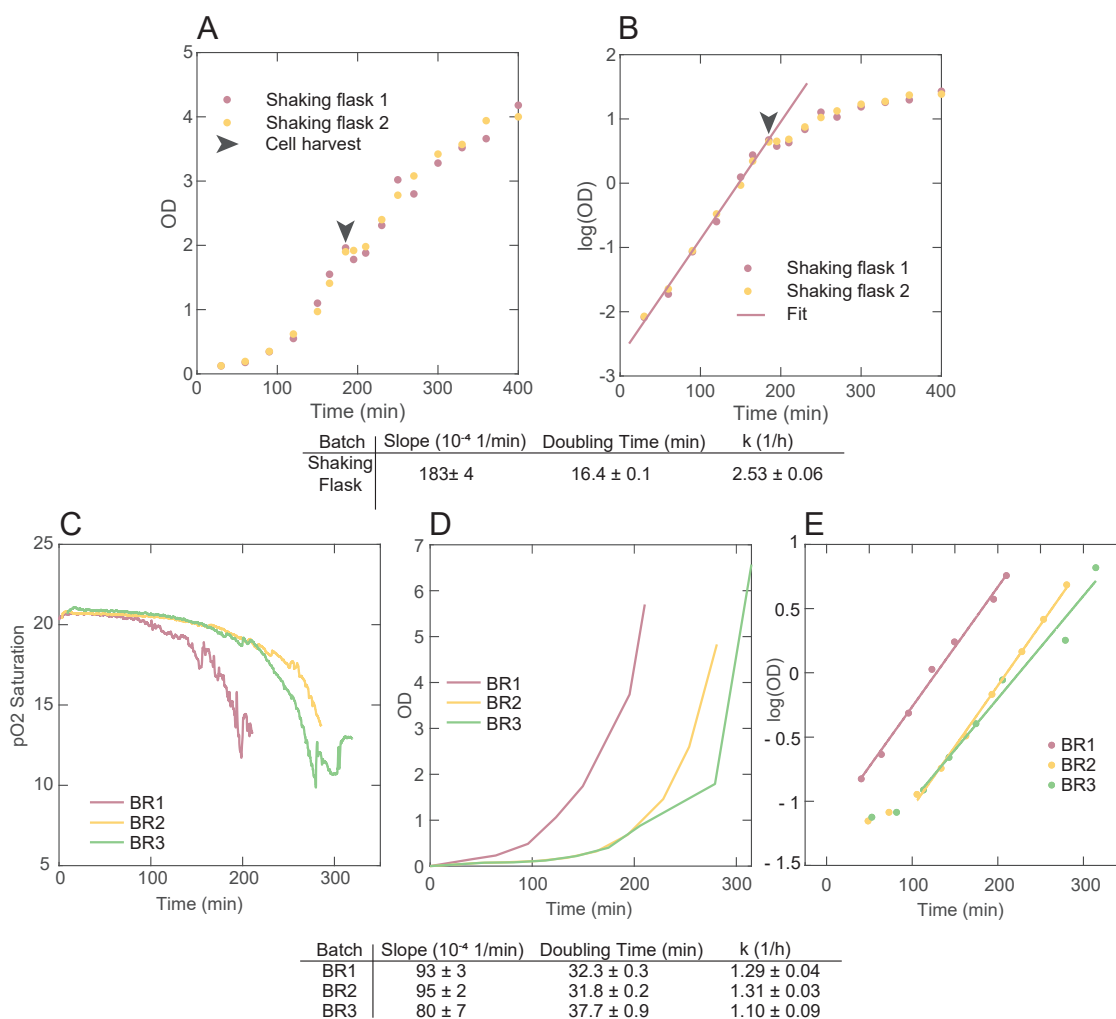
Figure 3.2: Growth curves for cultivation in shaking flasks and in a bioreactor. (A) pO2-saturation. The oxygen level was kept above 14% during the whole batch time by regulating the stirrer speed (500-1000 rpm) and the aeration rate (2-4 l/min) up. (B) OD. Batches BR2 and BR3 showed an elongated lag phase at the beginning. The desired OD was reached after about 200 min for batch BR1 and after about 280-300 min for batches BR2 and BR3. (C) Semi-logarithmic plot. The doubling times are between 32 and 38 min. (D) Growth curve for two different shaking flasks, cells are harvested between OD 1.8 and 2. (E) Semi-logarithmic plot. Bacteria were harvested in the late-log growth phase.

Cells were cultivated either in shaking flasks at 37 °C and 250 rpm or in a bioreactor.

The shaking flask cultures showed a typical logistic growth curve and were harvested in the late-log phase at an OD of 1.8-2, which yielded typically 2-2.5 g/L pellet mass (Figure 3.2A,B). During bioreactor (BR) cultivation glucose was fed with a constant rate of 0.85 g/(l·h) and the oxygen level was kept above 14% (Figure 3.2C) by regulating the stirrer speed (500-1000 rpm) and the aeration rate (2-4 l/min) up. Also the pH was monitored and kept constant, even though the culture medium was already buffered with potassium phosphate. As culture medium we chose 2xYT+P. 2xYT medium is very common for this purpose and is often supplemented with phosphate and also with glucose[50][59][82], which avoids phosphatase induction and results in reduced ATP hydrolysis activity [82]. Batches BR2 and BR3 showed an elongated lag phase at the beginning, but then resumed exponential growth at a similar rate as batch BR1 (Figure 3.2D-E). Due to the favorable growth conditions in the bioreactor, the exponential growth phase was prolonged and cells could be harvested after about 200 min for BR1 and 280-300 min for BR2 and BR3 at an OD between 5 to 6. As a result the biomass yield increased by a factor of 4 to 10 g/L.

3.1.3 Expression of proteins and protein content

In an initial study with shaking flask cultivation ("SF batches"), we coarsely screened the influence of lysozyme incubation against the number of sonication cycles. In the following we use, a shorthand notation for the lysis conditions, where Sx/Ly denotes a protocol with x sonication cycles at a lysozyme concentration of y mg/mL. For all extracts, the total protein content was determined by a Bicinchoninic acid (BCA) assay, while the efficacy of cell-free gene expression was tested by expressing the fluorescent reporter protein YFP (mVenus) from a constitutive promoter J23106.

Also other reporter proteins within the same expression cassette namely a RFP reporter (mScarlet-I), a GFP reporter (GFPmut3) and a CFP reporter (mTurqoise2) were tested. Appreciable differences in expression levels and yields in a range spanning one order of magnitude (Figure A.5) were observed. These may in part be explained by the optimized codon usage for the YFP reporter and differences in maturation path and time, quantum yield and other protein characteristics. Secondary structure and GC content of the corresponding mRNA can also have an influence on protein expression [71]. Nevertheless, the different reporters showed the same trends in the screening experiments for the different lysis settings. In addition, the maximum protein expression rate and the end level of expressed

protein concentration were positively correlated (Figure A.4). We therefore restrict the following discussion of gene expression efficiency to the concentration end level of the YFP reporter.

Samples, which were not incubated with lysozyme but lysed with sonication pulses (S5/L0) had the lowest mean protein content of around 6 mg/ml measured in the BCA assay (Figure 3.3A). These samples also generated the lowest fluorescence signal intensities in TXTL experiments (Figure 3.3B). The increase in the number of sonication cycles (from 5 to 15) resulted in an increase in the mean protein content (mean of 3 biological replicates) and also in an increase of the fluorescence end level almost by a factor of 3. Compared to these samples a higher protein content and therefore better lysis efficiency and also a higher TXTL fluorescence end level were observed for the S0/L0.5 and S0/L1 samples, which were not sonicated at all but incubated with lysozyme. Both tested lysozyme concentrations (0.5 mg/ml and 1 mg/ml) resulted in similar signal intensities.

As stated above, a combination of lysozyme incubation and sonication cycles should have a synergistic effect, as lysozyme is supposed to weaken the integrity of the cell walls and thus support sonication-induced cell lysis. Surprisingly, combined protocols did *not* result in an increased lysis efficiency (the mean cell extract protein content of three biological replicates saturated between 14 and 19 mg/ml, Figure 3.3A), but nevertheless in a twofold YFP increase in the TXTL test, when comparing the S5/L0.5 to the S0/L0.5 sample. The same end level for both tested lysozyme concentrations was observed (compare S5/L0.5 and S5/L1, Figure 3.3B). The fluorescence end level was reduced for the 15 cycle samples, suggesting an optimum intermediate number of sonication cycles, which was also confirmed using a t-test (Table A.6).

Next cells produced from a 2L fed-batch culture in a bioreactor ("BR batches") as described above were used. By this the pellet mass yield was increased and the screening range could be expanded. Lysozyme concentrations of 0.5 mg/mL and 0.8 mg/mL were screened against 4, 8, 12, 16 and 20 sonication cycles (Figure 3.3C and D). All cell extracts had a similar mean protein content around 15 mg/ml. In contrast to the shaking flask extracts, the deviations between the biological replicates were higher than the deviations between the different lysis settings within a single replicate (Figure A.3). We again performed TXTL tests and compared the corresponding fluorescence levels. For both lysozyme concentrations an increase in the number of sonication cycles resulted in an increase in the YFP end level where an optimum was reached in the range of 12-16 cycles with end level concentrations again up to 22 µM. The optimum was again proven to be significant using a t-test (Table A.6).
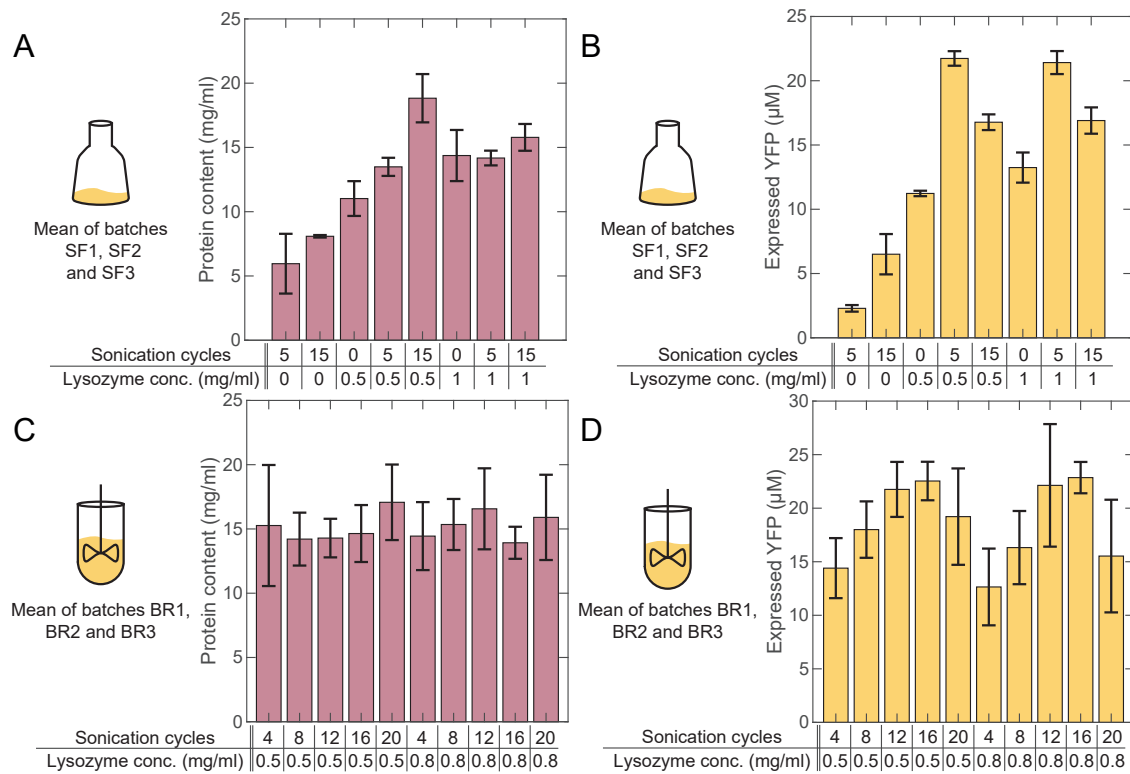
Figure 3.3: The mean total protein content was determined for three biological shaking flask replicates, which were lysed using a lysozyme concentration of 0, 0.5 or 1 mg/ml in combination with 0, 5 or 15 sonication cycles. (B) TXTL-test for the shaking flask replicates. The mean end levels are presented for a YFP expressed in the shaking flask cell extracts. (C) Mean total protein content for three bioreactor replicates, which were lysed using 0.5 or 0.8 mg/ml lysozyme in combination with 4, 8, 12, 16, or 20 sonication cycles. (D) TXTL-test for the bioreactor replicates. The mean end levels are shown for a YFP reporter expressed in the bioreactor cell extracts.

3.1.4 Batch-to-batch variations and previous studies

Batch-to-batch variations Despite the many advantages of cell-free protein expression studies, they potentially suffer from considerable batch-to-batch variations, which depend on details of the cell extract preparation procedure. Variability can result from variations in culture conditions, the growth state of the cells during harvest, cell suspension viscosity, energy input or heat production during lysis, lysis efficiency and activity of the extracted proteins [138][84]. In our shaking flask extracts we ob-

served standard deviations in the range of 2% and 24% of the signal intensity for three biological replicates, with a mean at 8%. For the bioreactor-prepared extracts we observed slightly higher deviations between the biological replicates in the range of 6% to 34%, but the mean was still at 19%. This might be a result of the differences in growth curves monitored for the bioreactor batches (extended lag phase for two of the three batches, cf. Figure 3.2). In our hands, common procedures such as careful monitoring of cell growth and harvesting point of the cells, cooling of the cell suspension over the full preparation time and mixing during cell lysis turned out to be sufficient for acceptably small batch-to-batch variations.

Comparison to other extracts and preparation methods Our best shaking flask replicates (S5/L0.5) and bioreactor replicates (S16/L0.5) were tested in comparison with a self-prepared batch using a bead beating protocol [59] and a commercially available kit (myTXTL Sigma 70 Master Mix Kit, Arbor Biosciences) (Figure 3.4A). In addition to our YFP plasmid we tested the pTXTL-p70a(2)-deGFP HP control plasmid shipped with the commercial kit, which codes for a GFP expressed from a lambda promoter. Our standard mVenus reporter has a higher quantum yield, brightness and codon adaptation index than the GFP from the commercial control plasmid. On the other hand, the computationally predicted translation rate [139] for the commercial GFP is higher by a factor of 10, which is mainly ascribed to the stronger secondary structure at the 5´UTR of our reporter mRNA transcripts or standby sites (see Table A.5 for a detailed comparison). The fluorescence signals were normalized to the maximum signal measured for each plasmid (see Figure A.5B for raw data). In the bead beating batch, we only observed 22% of the maximum signal for our YFP plasmid and 71% of the maximum for the p70-GFP plasmid. Furthermore, in the commercial kit we obtained a mere 1% of the maximum signal for our self-prepared YFP plasmid, while a 94% signal was measured for the kit's control plasmid. We surmised that residues from our plasmid preparation (following a standard phenol-chloroform extraction protocol) might have a detrimental effect on the commercial cell extract, as we observed also low signals for our RFP, GFP and CFP reporter plasmids in the commercial extract (Figure A.6A). We thus purified the p70-GFP control plasmid using a simple phenol-chloroform extraction protocol and repeated the TXTL test in the commercial kit, which indeed resulted in lower signal levels (Figure A.6C).

Comparison to the protocol by Kwon and Jewett [50] and the protocol by Fujiwara and Doi.[66] Whereas in the study of Kwon and Jewett the required

sonication energy input depended just on the cell suspension volume, we observed different optima for different culture methods, e.g., S5L0.5 for shaking flask and S16L0.5 for bioreactor cultivation. Further, we only used 404 J energy input for 5 cycles and 1.3 kJ for 16 sonication cycles compared to the 2.2 kJ expected from their study for a 4 ml cell suspension volume, but we supported the lysis step using lysozyme. As a result, our total extracted protein content was smaller (about 15 mg/ml compared to 40 mg/ml). Potentially, a fraction of non-lysed *E. coli* might be present after the lysis step in our protocol, but these cells would be removed during the centrifugation steps (in total 3) in the following extract processing steps. In addition, lysozyme is neither deactivated nor removed from the extract which makes survival of any remaining *E. coli* appear unlikely. On the other hand, the lower energy input used in our protocol potentially benefits the activity of the enzymes contained in the extract. In fact, the gene expression capability of our extract was similar as for the extract by Kwon and Jewett, as we reached a maximum YFP end level of 0.6 mg/ml expressed from a constitutive *E. coli* promoter at a plasmid concentration of 3 nM (5.7 µg/ml) (compared to 1 mg/ml for expression from a much stronger T7 promoter at a plasmid concentration of 13.3 µg/ml [50]).

In contrast, Fujiwara and Doi presented a protocol based on a combination of osmotic shock, lysozyme incubation and freeze thaw cycles. They reached a cell extract protein content in the range of 20-30 mg/ml and a protein expression yield of 10-20 µM (0.25-0.5 mg/mL) using either a template concentration of 1.5 nM with a T7 promoter or 10 nM with a OR2OR1 promoter. The details of both studies and also to the study by Sun et al. can be found in Table 3.1.

	This study	**Sun et al.**	**Kwon and Jewett**	**Fujiwara and Doi**
Strain	Rosetta 2 (DE3)	Rosetta 2 (DE3)	BL21 star (DE3)	BL21(DE3) codon plus (RIL)
Culture method	Shaking flask or bioreactor	Shaking flask	Shaking flask or bioreactor	Shaking flask
Culture medium	2xYTP shaking flask; 2xYTGP bioreactor	2xYTP	2xYTPG	LB
IPTG induction	No	No	Tested, but usually not used	Yes
Lysis method	Lysozyme incubation + sonication	Bead beating	Sonication	Lysozyme incubation + osmotic shock + freeze thaw cycles
Cell extract processing steps (after lysis)	In total 3 centrifugation steps; run off reaction; dialysis	In total 3 centrifugation steps; run off reaction; dialysis	1 centrifugation step; (run-off reaction and second centrifugation step was also tested)	1 centrifugation step (buffer exchange was also tested)
Total protein content (mg/ml)	15	30	40	20-30
Expressed protein (mg/ml)	0.6	0.75	1	0.25-0.5
TXTL test conditions	J23106 promoter; 9 nM corresponds to 17 µg/ml plasmid concentration	Lamda promoter or T7 promoter	Addition of T7 polymerase; T7 promoter (13.3 µg/ml plasmid concentration)	T7 promoter (1,5 nM template concentration) or OR2OR1 (10nM plasmid concentration)

Table 3.1: Comparison of different cell extract preparation protocols published by Falgenhauer et al. [1], Sun et al. [59], Kwon and Jewett [50], and Fujiwara and Doi. [66]

3.1.5 Assembly of bacteriophages

As an alternative to the synthesis of fluorescent proteins, we also assessed the quality of our cell extract via *in vitro* expression and assembly of T7 bacteriophages [40]. Phage assembly is a considerably more complex process than the expression of just a single protein [40][140] and can thus serve as a benchmark for the capability of the cell extract to support more complex biochemical processes. For quantitation of phage assembly, we performed plaque assays and determined the phage titers. In general, the infection titer depends on the concentration of phage particles, the ratio of phage particles compared to host cells, the physiological state of the host cell (competence) and the activation state (stress versus hunger/ feast state). These parameters are difficult to control over

several experiments, and we therefore performed a single experiment to compare the best lysis setting of our shaking flask (S5/L0.5) and bioreactor replicates (S16/L0.5) with the phage titers reached for the self-made bead beating batch and the commercial expression kit. All LAS extracts showed high phage titers up to 10^9 PFU/ml (with a mean of 10^8 PFU/ml for three biological replicates; Figure 3.4B and Figure A.7). This has to be compared to a mean titer of close to 10^7 PFU/ml for the commercial kit and merely 250 PFU/ml for the bead beating batch.
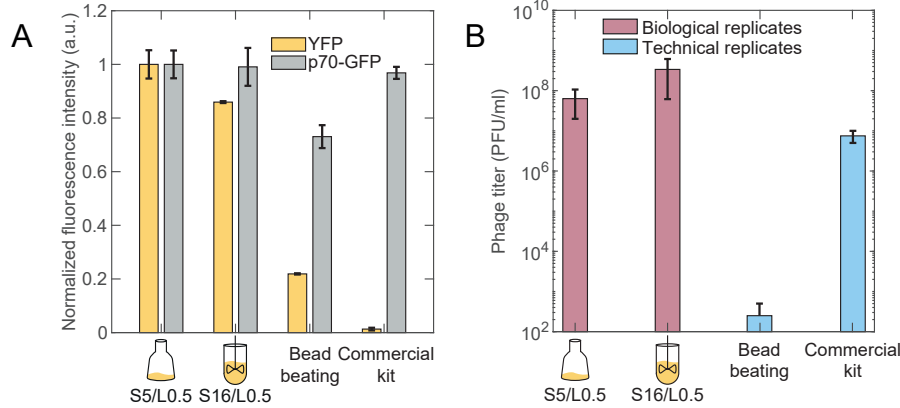
Figure 3.4: (A) The shaking flask replicates (S5/L0.5) and bioreactor replicates (S16/L0.5) were tested in comparison to a bead beating batch and a commercial kit. In addition to our self-prepared YFP plasmid we tested the p70-GFP control plasmid, which codes for GFP under the control of a lambda promoter. The fluorescence signals were normalized to the maximum measured signal for each plasmid. (B) Plaque assays for the best shaking flask and bioreactor replicates, a commercial extract and a beat beating batch.

Overall, our results show that there is an optimum number of sonication cycles for both cultivation methods (5 for shaking flask and 12-16 for bioreactor samples). Bacteria grown in the bioreactor – where they are subjected to larger shear forces than in shaking flask culture – have been previously found to change their morphology (resulting, e.g., in an increase in cell length [141]), which potentially allows them to sustain a larger number of sonication cycles. When expressing fluorescent proteins in extract prepared by the LAS protocol, a higher YFP end level was reached compared to a bead beating cell extract, but similar end levels were reached as in a commercial extract and in the study of Kwon and Jewett. Notably, our cell extract had significantly higher gene expression yields for plasmids which were purified with an inexpensive phenol chloroform precipitation (compared to the commercial kit and the bead beating batch). In the plaque assays our LAS extracts performed marginally better than the commercial extract and the bead beating batch again

showed the worst results.

3.1.6 Proteomics
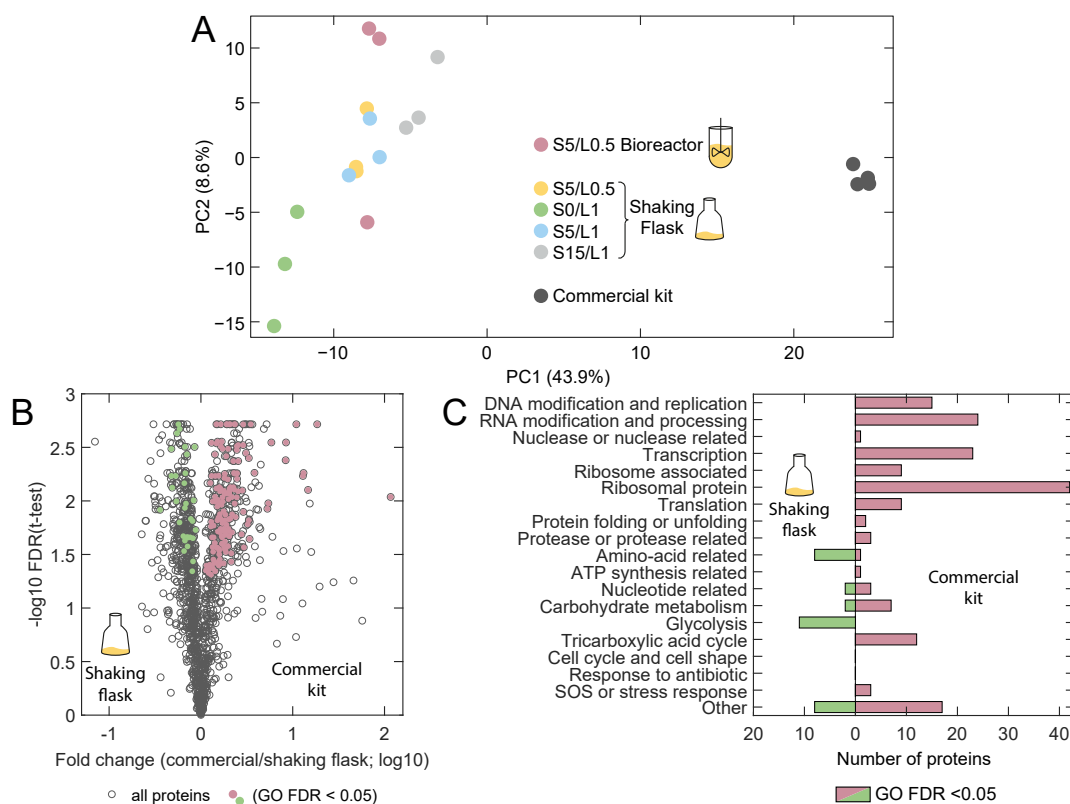
Figure 3.5: Comparison of the proteomes of extracts prepared from bioreactor and shaking flask cultures and a commercial kit. (A) The principal component analysis (PCA) shows that the proteome of the commercial extract is distinct from those of self-made extracts (separated on PC1, accounting for 44% of the total variance). (B) Volcano plot of shaking flask samples S5/L0.5 against the commercial kit. Proteins with a FDR below 0.05 and proteins exclusively found only in shaking flask preparations or in commercial extract were subjected to an enrichment analysis. Proteins which were assigned to GO terms with an FDR below 0.05 are highlighted in the plot. (C) Enrichment analysis derived from the comparison shown in (B). The significant proteins were assigned to keywords related to gene expression and energy regeneration.

In order to elucidate the molecular basis for the observed differences in performance, we analyzed selected cell extract samples and technical replicates of the commercial

extract using state-of-the-art quantitative proteomics similar to the study of Garenne et al. [67]. We chose the best performing shaking flask extracts (biological replicates SF1, SF2, SF3 with lysis setting S5/L0.5) and prepared and analyzed bioreactor samples with the same lysis settings (three biological replicates with setting S5/L0.5) to identify differences in protein composition that are correlated with the different culture conditions. We also analyzed shaking flask samples S0/L1, S5/L1 and S15/L1 to find differences in protein composition resulting from different lysis settings. We quantified around 1500 proteins in each preparation. A principal component analysis (PCA) showed a systematic difference between the proteomes of self-made extracts (biological replicates) and the commercial kit (technical replicates), which were separated on principal component 1 (PC1), explaining 44% of the total variance (Figure 3.5A). Self-made extracts scattered on PC2 (explaining 8.6% of the total variance). Compared to extracts prepared with sonication, the S0/L1 extract resulted in more negative PC1 values, indicating an influence of sonication on the proteome content of the extracts. This is in agreement with the results of the TXTL tests, as the samples prepared without sonication also had a reduced expression yield. Samples S5/L0.5 and S5/L1 overlapped in the PCA plot, which demonstrates a similar proteome composition and is consistent with the comparable TXTL test results.

Motivated by these results, we decided to further investigate the proteomic differences between different extracts using t-tests and further analyzed proteins with a false discovery rate (FDR) in the t-test below 0.05 and 'unique proteins'. Unique proteins are present in all three replicates of one extract but absent in all of the three replicates of the other extract (these proteins cannot be represented in a volcano plot). The distinctive proteins were then subjected to a gene ontology enrichment (GO) analysis to identify whether specific GO terms were statistically enriched. Proteins found in GO terms with an FDR <0.05 were subsequently roughly classified to keywords using the UniProt database [18] to simplify the representation.

In the first step we analyzed the differences between the extracts produced from shaking flask and bioreactor cultivation (Figure 3.6). The protein content of the two preparations was largely similar, except for three unique proteins in the shaking flask, which were found in GO terms with GO FDR < 0.05 that could be assigned to anaerobic growth conditions. This is a somewhat expected result as in contrast to bioreactor cultivation we did not provide additional oxygen in shaking flask cultivation.
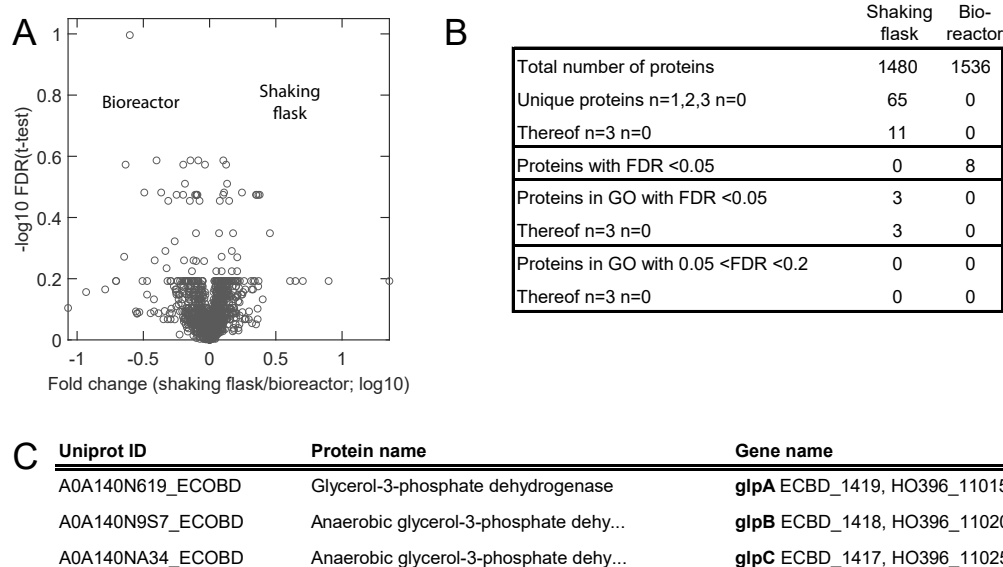
Figure 3.6: Comparing proteomes of extracts from shaking flask and bioreactor. (A) Volcano plot of shaking flask samples S5/L0.5 against bioreactor samples S5/L0.5. No protein had a FDR smaller than 0.05, but some unique proteins were found and used for enrichment analysis. (B) Summary of protein numbers. Compared to the bioreactor samples (S5/L0.5) the shaking flask extracts (S5/L0.5) contain 11 unique proteins, which were subjected to an enrichment analysis. (C) Result of enrichment analysis. For the shaking flask samples just 3 proteins were found in GO terms with GO FDR < 0.05. These proteins can be assigned to anaerobic growth conditions.

We next compared batches, which were lysed with different sonication cycles at the same lysozyme concentration of 1 mg/ml (S0/L1 vs. S5/L1 and S0/L1 vs. S15/L1, Figure A.8A and A.8B). In total we subjected 2, 27, 4 and 53 proteins to GO enrichment analysis and found just 7 (S5/L1) and 21 (S15/L1) proteins in significant GO terms (see Table A.7-A.9 for protein names and Table A.10 for the number of analyzed proteins in each comparison). For both comparisons these could be assigned to keywords related to DNA replication, relaxation, repair or recombination or were transcriptional regulators. For the gene expression yield this may play just a minor role, but it indicates that increased sonication energy input results not just in increased lysis efficiency but also in a more efficient release of proteins bound to the genome, which would be otherwise lost together with the genome during the first centrifugation step.

In the next step we compared our extracts against the commercial kit (Figure A.8C-F and Figure 3.5B and C). Here one has to keep in mind that we have no detailed

information about the preparation of the commercial extract, so that any observed deviation could be the result of differences during each single step of the preparation workflow. However, as the YFP expression capability of our extracts and the commercial system are very similar, it is still of interest to compare the differences between the proteomes of the extracts.

As our analysis of the bioreactor replicates S5/L0.5 and the shaking flask replicates S5/L0.5 revealed only minor differences, one would expect similar results for the comparison of either BR or SF samples against the commercial extract, respectively. In fact, however, the results are found to be quite different, because the bioreactor replicates show a higher variance among each other (as noted above), which results in higher p-values and thus an exclusion of the corresponding proteins in the enrichment analysis. As a consequence, the commercial and bioreactor extracts differ in abundance of just 8 and 5 proteins with FDR < 0.05 and showed 30 and 121 unique proteins, respectively. After enrichment analysis we further assigned 11 proteins to keywords, which were related to amino-acid biosynthesis, a nuclease or not relevant for cell-free gene expression (Figure A.8C and D).

In contrast, the commercial and shaking flask extracts differed in the abundance of 309 versus 356 proteins with FDR < 0.05 (corresponding to $-\log_{10}(\text{FDR}) > 1.3$) and had 116 and 53 unique proteins, respectively (Figure 3.5B and C). After enrichment analysis in total 172 and 31 proteins were further analyzed and assigned to keywords, respectively (40% and 8% of the proteins which had an FDR < 0.05 or were unique proteins). According to this analysis, in shaking flask batches proteins taking part in carbohydrate metabolism and glycolysis, amino-acid and nucleotide related pathways were found to be more abundant than in the commercial kit. In contrast, proteins which are related to RNA modification and processing, DNA modification and replication, transcription regulation, initiation and termination and the TCA cycle were found enriched in the commercial extract. The potential role of the single proteins found in these keywords is discussed in the following in detail.

Energy metabolism   Compared to the commercial extract, in our home-made batches 9 out of 10 enzymes from the glycolysis pathway were enriched, including glucokinase, which converts cytosolic glucose into glucose-6-phosphate (triosephosphate isomerase was the only enzyme that was not present at a higher abundance). Further, we found enzymes such as adenylate kinase or guanylate kinase enriched in our self-made batches, which potentially play a role in ATP and GTP regeneration *in vitro*. On the other hand, we detected higher abundance of the subunits IIB and

IIC of *E. coli's* major transmembrane carbohydrate transport system (the Pst system) in the commercial extract. This indicates that the lysis conditions used for the commercial extract might cause a higher fragmentation of the membrane and thus a more efficient release of trans-membrane proteins. In contrast to glycolysis, enzymes of the TCA cycle were enriched in the commercial extract. In shaking flask batches, we also detected an increased amount of dehydrogenases encoded by the glpABC operon, which belongs to the glycerol kinase pathway and is responsible for anaerobic energy generation. No differences were detected in the pentose phosphate pathway.

Transcription/translation   Together with core RNA polymerase, sigma factor RpoD (σ70) dominates the transcription in exponentially growing cells. We found RpoD more highly abundant in the commercial extract. Furthermore, during envelope stress several sigma factors are upregulated in *E. coli*, including sigma factor RpoE (σ24), which was also found more abundant in the commercial extract. Other stress factors such as translational regulator CsrA (envelope/periplasmic stress) and nitrogen-limitation factor RpoN (σ54) were enriched in the commercial cell-free system as well as the ribosomal subunit S22, which is associated with stationary bacterial growth. Apart from these stress indicators, a variety of other transcription regulators such as LacI, MarR, GntR, DeoR or LysR were enriched compared to our home-made shaking flask batches.

In addition, also translational capacity appeared to be enriched in the commercial system. In particular, we found 20 out of the 22 ribosomal proteins of the 30S subunit at higher abundance in the commercial extract, including S22 (see above) and the essential ribosomal protein S12, which takes part in both tRNA and ribosomal subunit interactions. In case of the 50S subunit, we found 24 out of 33 ribosomal proteins more abundant in the commercial extract, including the small ribosomal protein L34 (5.3 kDa). Also other translation-related proteins showed higher abundance in the commercial extract such as initiation (IF-2, IF-3) and elongation (EF-4, EF-Tu, SelB) factors, but also the ribosomal silencing factor (RsfS) which inhibits ribosome association and prevents translation.

Degradation of nucleic acids and proteins   We also found notable differences between the cell extracts in degradation pathways. In the commercial extract ribonucleases 2 and E were enriched (p-value 0.05), which are mainly involved in mRNA degradation. Other ribonucleases participating in RNA maturation and processing (RNAse 3, G, PH and R) were also more abundant in the commercial extract. On

the other hand, endoribonuclease L-PSP, also acting on mRNA, was found more highly concentrated in self-made batches. We also investigated the presence of proteases in the cell extracts. We found both subunits HslV and HslU (annotation at the transcript level) of the proteasome-like degradation complex HslVU (ClpQY) enriched in the commercial extract, which unfolds proteins under ATP consumption.

Biosynthesis Interestingly, many proteins involved in amino acid biosynthesis were more abundant in our self-made batches compared to the commercial extract, suggesting their potential use for amino acid production inside the extract starting from inexpensive precursors. Finally, the chaperone cofactor GroES was more highly expressed in the commercial batch, but not its chaperone complex GroEL, even though it is encoded by the same operon.

In short one can summarize, that ATP regeneration might be upregulated in our extract, if intermediates of the glycolysis pathway are used as an energy source. In addition, one can potentially use inexpensive precursors to produce amino acids inside the extract by taking advantage of the enriched proteins of the amino acid biosynthesis pathways. On the other hand, lysis efficiency (higher degree of fragmentation and release of DNA bound proteins) potentially is higher for the commercial extract, as some transmembrane proteins and DNA binding proteins were found to be enriched in this extract. In addition, also the translational capacity appeared to be enhanced in the commercial system as we found a higher abundance of ribosomal proteins and elongation factors. However also some stress factors, ribonucleases and proteases were found to be enriched in the commercial extract, which might have a negative effect on gene expression yield. In general, the results of this analysis have to be interpreted with care. Notably, 67% of the compared proteins only had a fold change of less than 2 (see Figure A.8F). In combination with the missing information about the activity of the single enzymes, the effect on gene expression yield is difficult to assess.

3.2 Conclusion

In conclusion, we have studied a cell extract preparation protocol, which uses a combination of lysozyme incubation and multiple sonication cycles. In contrast to earlier work, we observed a synergistic effect of lysozyme incubation and sonication on the

expression efficiency of the cell extract. Expression of YFP from a self-prepared reporter plasmid in our cell extract resulted in a 100-fold higher fluorescence signal than when using a commercial CFS. Using a commercial control plasmid, by contrast, resulted in similar signals in both types of extract, suggesting a sensitivity of the commercial product towards residues from the plasmid purification protocol used. Our lysozyme incubation/sonication-based extract performed better in the in vitro assembly of T7 bacteriophages, which we used as an example of a more complex assembly process. Our cell extract reached a phage titer, which was at least one order of magnitude higher than what could be obtained in a commercial system. By contrast, phage assembly using a bead beating protocol (instead of sonication) resulted in a very low titer. We attempted to rationalize the observed differences using state-of-the-art quantitative proteomics. This approach revealed rather small differences among our self-made batches. We found that higher sonication energy inputs might not just increase the lysis efficiency but potentially also promotes the dissociation of DNA-binding enzymes from DNA. In addition, we compared our extracts to a commercial system. Even though both our self-made and the commercial cell extract were prepared from *E. coli* Rosetta 2 (DE3) cells [67] we expected to observe differences in the abundance of proteins due to the different culture and lysis methods. Indeed, our proteomic analysis showed clear differences between the commercial cell extract and our batches. However, lack of information about the activity of the enzymes (rather than their abundance) limits the interpretability of these results. A range of other factors might come into play that influence expression yield. For instance, it has been reported that *in vitro* protein expression is limited by low ribosomal activities and in particular by the lack of ternary complexes formed by EF-Tu, tRNA and GTP [142]. Further, energy metabolism has been shown to play a crucial role in protein synthesis yield [30][143].

3.3 Genetic circuits

The produced cell extract can be used to test genetic circuits while taking advantage of the modularity of the system. In detail the system can be supplemented with purified proteins and small molecules. Also the concentrations of the DNA templates can be adjusted to optimize the circuit performance. To demonstrate this briefly a toehold switch based circuit with two positive feedback loops and mutual inhibition (Figure 3.7) designed and cloned during my Masters thesis [144] is tested in the LAS cell extract. TS1 regulating the expression of a T7 mutant (T3R5) is transcribed upon aTc induction. Leaky

translation of TS1 results in the expression of T3R5 at a low level. The polymerase co-transcribes T1, AT2 and the mRNA for a GFP reporter. The produced Tr1 activates TS1 and starts a positive feedback loop, whose functionality was already demonstrated by myself during my Masters thesis. An second subsystem with pLac, TS2, Tr2, AT1 and a RFP reporter is cloned to complete the circuit.
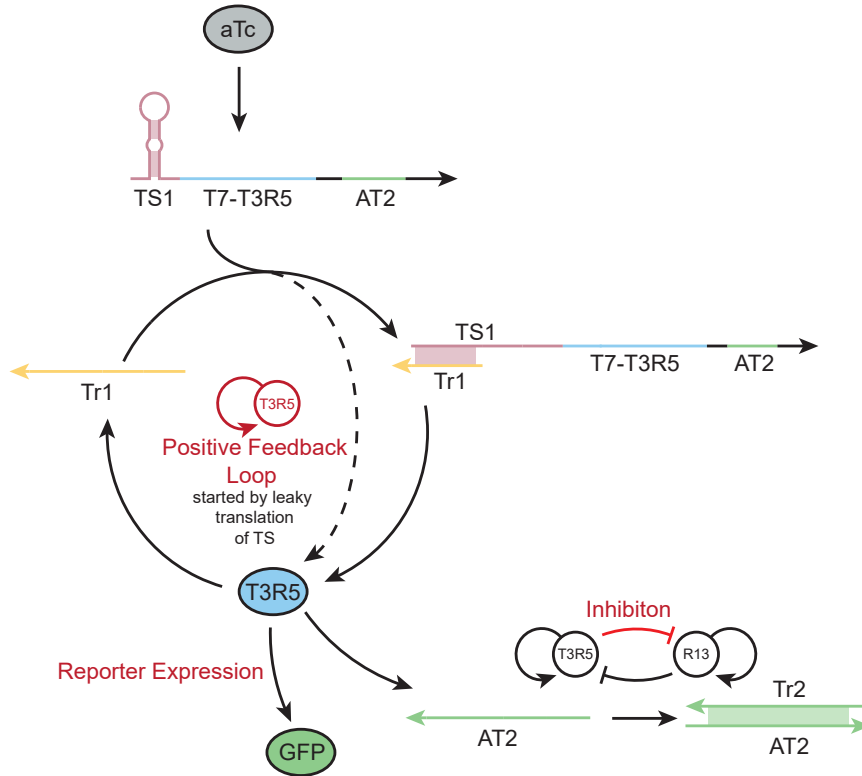
Figure 3.7: T3R5 subsystem. TS1 is transcribed after aTc induction. Leaky expression of TS1 starts the positive feedback loop. Inhibition of an analog R13 subsystem is implemented by transcribing AT2 and the system status is read out using a GFP reporter. Figure was adopted with modifications from my Master thesis [144].

In a first plate reader bulk measurement (Figure 3.8) the functionality of the positive feedback loops is again demonstrated (in absence of repressor proteins) and the mutual inhibition is tested. Unfortunately the inhibition of the T3R5 subsystem by the R13 subsystem is less effective than vice versa, which results in a high GFP level (instead of low levels for both reporters) for a fully active circuit. The cell-free system is then supplemented with purified repressor proteins (LacI or TetR) to successfully deactivate either one or both subsystems. Also the activation of the circuit with inducer molecules (IPTG or aTc) is shown.
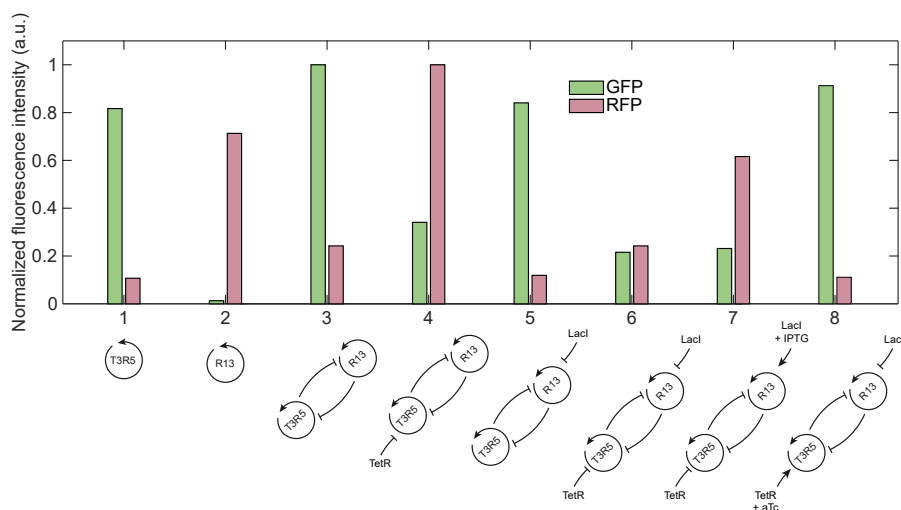
Figure 3.8: Plate reader bulk measurement. T3R5 and R13 subsystem measured separately show high signal levels of their own and low crosstalk levels for the antagonsitic reporter. The repressive effect of R13 on T3R5 is to low. Both subsystems show low signal levels in presence of their corresponding repressor proteins (LacI or TetR) and can be induced again in presence of inducer (IPTG or aTc).

In a second experiment (Figure 3.9), IPTG was screened against aTc and the end levels of the reporter proteins were measured. As expected two regions with one dominant and one repressed subsystems could be detected. Also a coexistence region exists, in which the R13 reporter shows an intermediate expression level. The system (R13 repression) has to be further optimized to lower also the T3R5 level in this region and to finally implement bistability.
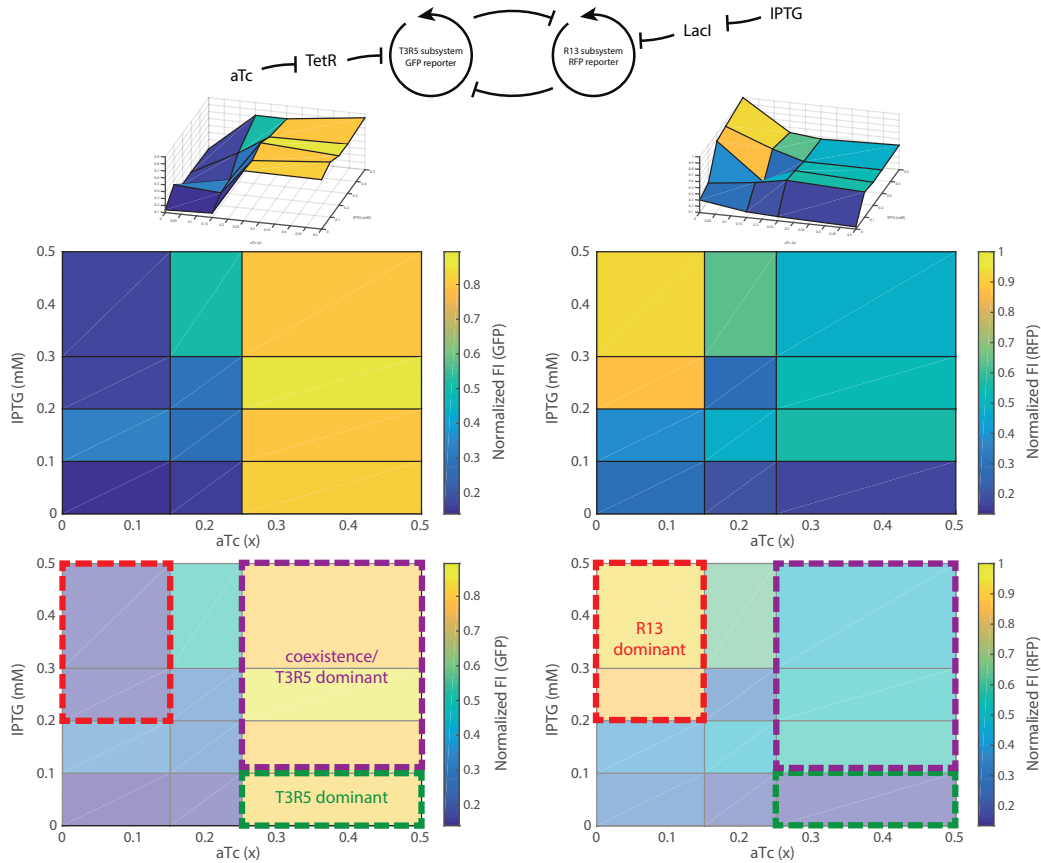
Figure 3.9: Plate reader bulk measurement. IPTG concentration was screened against aTc concentration and the endlevels of the T3R5 and R13 reporters are measured. The landscape shows two dominant and a coexistence region.

3.4 Materials and methods

3.4.1 Chemicals

Unless otherwise noted all chemicals were ordered from Sigma Aldrich (exceptions are listed in Table A.1). The composition of 2x YTP medium, S30A and S30B buffer was adopted from Sun et al. [59]. The TXTL buffer containing amino acids, nucleotides, tRNAs and other ingredients was also prepared according to Sun et al. and screening experiments for Mg-glutamate, K-glutamate, PEG and DTT concentrations were performed (see Figure A.2).

3.4.2 Cell extract preparation

Bacterial strains and culture conditions   Figure 3.1 gives a quick and Figure A.1 a detailed overview over the cell extract preparation workflow. Bacterial cell extracts were prepared from *E. coli* Rosetta 2 (DE3) cells. Cells from glycerol stocks were grown overnight in an incubator shaker (Innova44, New Brunswick) in 2x YTP medium containing selective antibiotic (Chloramphenicol, Cm) at 37 °C and 250 rpm. On the following day, cells were diluted 1:100 in 2x YTP+Cm medium and culti-vated either in eight shaking flasks (666 mL each, 37 °C, and 250 rpm) or in a 2 L bioreactor.

Cultivation of bacteria in a bioreactor   Cells were cultured in a 2 L lab scale bioreactor (Minifors 2, Infors) with pO_2 monitoring and pH control. Initially, the culture was agitated at 500 rpm and aerated with pressurized air at a rate of 2 L/min. To keep the oxygen saturation over 14%, we regulated the aeration rate stepwise up to 4 L/min and increased the stirrer speed up to 1000 rpm. We used a constant feeding with a glucose solution at a rate of 0.85 g/(L h) during the entire cultivation time. As we grow the bacteria in 2x YT medium supplemented with potassium phosphate (monobasic and dibasic), the medium is buffered and the external pH control during the cultivation is not obligatory.

Cell harvest and washing   Cells were harvested at OD 1.8-2.0 when cultivated in shaking flasks, or at OD 5-6 when grown in a bioreactor. After harvesting, we dis-tributed the cell suspension to four bottles (750 mL each) for centrifugation (15 min,

4 °C, max. speed 4600 rcf, Rotanta 460R, Hettich). In the case of shaking flask cultivation, a second round of cell harvesting was necessary. The supernatant was decanted and cells were resuspended in S30A buffer (300 mL per centrifuge bottle). Centrifugation and washing (2 bottles were pooled, washing with 2x 350 ml) was repeated. Afterwards, we resuspended the cell pellets in 2x 40 mL S30A buffer and transferred the suspension into two 50 mL falcon tubes. Cells were centrifuged at 3000 rcf at 4 °C for 10 min. The supernatant was decanted and the pellet was again centrifuged at 3000 rcf at 4 °C for 3 min. The supernatant was then removed using a pipette. After determination of the wet pellet mass (typically 20 g for bioreactor cultivation), the pellets were flash frozen in liquid nitrogen and stored at -80 °C.

Preparation of cell extract   Cell pellets obtained in the previous step were thawed on ice and resuspended in S30A buffer (1 mL buffer per gram pellet mass) by vortexing. The cell suspension was then split into 4 mL aliquots and up to 1 mg/mL lysozyme was added. After mixing by pipetting up and down the cell suspension was incubated on ice for 20 min. Cells were sonicated on ice using a SONOPULS mini20 (Bandelin) with a working frequency of 30 kHz at 10% amplitude (corresponds to 5 W power input). We applied 0-20 pulse cycles with durations of 10 s each. Tubes were sonicated in series, so the cooling time between different rounds is 10 sec times the number of tubes. Thus, our typical pausing time between sonication pulses was 80-100 sec per tube (for 8-10 samples). Since throughput with sonication is limited, we do not recommend to prepare larger volumes than 40 mL cell suspension at a time (at a constant sample volume of 4 mL per tube). After every second cycle, samples were mixed by pipetting up and down using a 5 mL pipette with the tip cut off. After lysis, samples were transferred into 2 mL tubes and centrifuged at 20,000 rcf for 30 min - 60 min at 4 °C until a sufficiently stable pellet had formed. Pellet-free supernatant was transferred into 2 mL screw cap tubes (1-1.5 mL volume per tube), leaving the caps unscrewed. The open tube containing the cell extract was inserted in a 15 mL Falcon tube as described by Sun et al. [59] and incubated at 37 °C and 250 rpm for 80 min in an Innvova44 shaker for a run-off reaction. Afterwards, samples were transferred into individual reaction tubes and centrifuged at 12,000 rcf for 10 min at 4 °C. Pellet-free supernatant was transferred into 10 kDa MWCO dialysis tubing and dialyzed against S30B buffer at 4 °C for 3 h. Cell extract was then extracted from the tubing, distributed into 1.5 mL centrifugation tubes and centrifuged at 12,000 rcf for 10 min at 4 °C. Cell extract was finally aliquoted into the desired aliquot size (usually 30 µl), flash frozen in liquid nitrogen and stored at

-80 °C.

3.4.3 Bicinchoninic acid (BCA) assay

Each cell extract batch was subjected to a Bicinchoninic acid (BCA) assay (Pierce BCA Protein Assay Kit, Reducing Agent Compatible Thermo Fisher Scientific) according to the manufacturer's protocol to determine the total protein content of the prepared cell extracts. The results displayed in Figure 3.3A and C show the mean and standard error of the mean of three biological replicates.

3.4.4 Transcription-translation of fluorescent proteins (TXTL test)

High copy number plasmids (iGEM part pSB1A3) containing a constitutive promoter (iGEM part J23106), an RBS (iGEM part B0034), the coding sequence for mScarlet-I (RFP), mVenus (YFP), GFPmut3 (iGEM part E0040) or mTurquoise-2 (CFP) and a Terminator (iGEM part B0015) were purified using a Qiagen Plasmid Midi Kit and afterwards phenol chloroform precipitated (see Figure A.11 for plasmid maps and Tables A.12-A.15 for plasmid sequences). Whereas Takahashi et al. [84] recommended to use a protein content of 10 mg/mL in the final cell-free protein expression reaction, we used the same dilution factor for all tests independently of the protein content. Our samples contained 33.3% cell extract, 41.7% buffer solution and 25% plasmid mix or water for blank samples. The final plasmid concentration in each sample was 3 nM and the TXTL buffer composition resulted in sample concentrations of 4 mM Mg-glutamate, 60 mM K-glutamate, 1.5 mM each amino acid except leucine, 1.25 mM leucine, 50 mM HEPES, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM CoA, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA and 2.5% (w/v) PEG-8000.

3.4.5 Fluorescence acquisition

Transcription-translation of fluorescent proteins was monitored with a plate reader (FLUOstar Omega, BMG Lab Tech) at a temperature of 29 °C. Fluorescence measurements were performed every 3-6 min using the corresponding filter sets for RFP,

YFP, GFP or CFP. Time traces were background corrected with blank values and molar concentrations were calculated using calibration curves for each of the four fluorescent proteins. To compare the fluorescence time traces, their maximum slopes (maximum protein expression rate) and end levels were determined. Figures display the mean of 3 biological replicates and the corresponding standard error of the mean.

3.4.6 Phage assembly

Phage assembly was performed according to the protocol of Rustad et al. [145] with the following adjustments: Phage DNA was mixed with cell extract, an energy solution and an amino acid solution as described in Sun et al. [59] using the same TXTL buffer composition as described in the TXTL test section. For 6 reactions à 13 µL, 2.5 µL PEG-8000 (36% w/v), 4 µL dNTPs (25 mM), 0.8 µL ATP (500 mM), 37.5 µL TXTL buffer, 2 µL GamS (150 µM), 28.5 µL cell extract and 1.6 µL DNA (10 nM) were mixed with nuclease-free water to a final volume of 80 µL. All constituents were mixed (except DNA) on ice and incubated for 5 min, followed by the addition of DNA. This 13 µL assembly mix was incubated for 4 h at 29 °C to express the bacteriophages.

3.4.7 Plaque assay

Plaque assays were performed using the top-agar method, with 0.5% agarose in NZCYM (Carl Roth), a standard medium for *E. coli* cultures and bacteriophages. The agar was melted and stored before use in a water bath at 48 °C. Separately, phage dilutions of 10^2-10^8-fold in phage buffer (1x PBS, 1mM $MgCl_2$, 1mM $MgSO_4$) were prepared. 100 µL of each dilution was mixed with an equal volume of an overnight culture of the corresponding host bacterium. This mixture was added to the 0.5% agarose NZCYM medium aliquots and poured on a 1% NZCYM agar plate. After solidified at room temperature, the plates were incubated at 37 °C until plaques became visible.

3.4.8 Sample preparation for mass spectrometry

All cell extracts were dried to completeness using a centrifugal evaporator (Centrivap Cold Trap -50, Labconco, US). The resulting pellets were dissolved in lysis buffer (8 M Urea, 5 mM EDTA, 100 mM NH_4HCO_3) to a final concentration of 1 mg/mL. Next, 45 µg of each sample were reduced with 10 mM dithiothreitol (DTT) (30 min at 30 °C) and alkylated with 55 mM 2-chloroacetamide (CAA) (30 min in the dark at 25 °C). The samples were diluted 1:4 in 50 mM NH_4HCO_3 and double-digested with trypsin (1 h and 13 h at 30 °C, Trypsin gold Mass Spectrometry Grade, Promega), which was added twice at a ratio of trypsin:protein = 1:100 (by mass). The reaction was stopped with 1% formic acid (FA) and the resulting peptides were purified. For that in-house built C18 tips (5 disks of Sep-Pak Vac C18 material, Waters, US) were equilibrated with 250 µl 100% acetonitrile (ACN), 250 µl elution solution (40% ACN, 0.1% FA) and 250 µl washing solution (2% ACN, 0.1% FA) at 1500 g. The samples were loaded into the tips (centrifugation for 2 min at 500 g) and washed three times with washing solution for 2 min at 1500 rcf. Finally, the peptides were eluted with 100 µl elution solution for 2 min at 500g. The samples were dried to completeness and resuspended in washing solution 45 µl right before the MS measurement.

3.4.9 Proteomics

Proteomics data acquisition   Generated peptides were analyzed on an Dionex Ultimate 3000 RSLCnano system coupled to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific, Bremen, GER). For each analysis an injection amount of about 0.1 µg of peptides was delivered to a trap column (ReproSil-pur C18-AQ, 5 µm, Dr. Maisch, 20 mm x 75 µm, self-packed) at a flow rate of 5 µL/min in 100% solvent A (0.1% formic acid in HPLC grade water). After 10 min of loading, peptides were transferred to an analytical column (ReproSil Gold C18-AQ, 3 µm, Dr. Maisch, 400 mm x 75 µm, self-packed) and separated using a 50 min gradient from 4% to 32% of solvent B (0.1% formic acid in acetonitrile and 5% (v/v) DMSO) at 300 nL/min flow rate. Both nanoLC solvents contained 5% (v/v) DMSO. The Fusion Lumos Tribrid Mass Spectrometer was operated in data dependent acquisition and positive ionization mode. MS1 spectra (360–1300 m/z) were recorded at a resolution of 60,000 using an automatic gain control (AGC) target value of 4×10^5 and maximum injection time (maxIT) of 50 ms. After peptide fragmentation using

higher energy collision induced dissociation (HCD), MS2 spectra of up to 20 precursor peptides were acquired at a resolution of 15,000 with an automatic gain control (AGC) target value of 5×10^4 and maximum injection time (maxIT) of 22 ms. The precursor isolation window width was set to 1.3 m/z and normalized collision energy to 30%. Dynamic exclusion was enabled with 20 s exclusion time (mass tolerance +/-10 ppm). MS/MS spectra of species that were singly-charged, unassigned or with charge states > 6+ were excluded.

Proteomics data analysis   Peptide identification and quantification was performed using the software MaxQuant (version 1.6.3.4) with its built-in search engine Andromeda [146][147]. MS2 spectra were searched against the *E. coli* (strain B / BL21-DE3) reference proteome from Uniprot (UP000002032, 4156 protein entries), supplemented with common contaminants (built-in option in MaxQuant). Trypsin/P was specified as proteolytic enzyme. Precursor tolerance was set to 4.5 ppm, and fragment ion tolerance to 20 ppm. Results were adjusted to 1% false discovery rate (FDR) on peptide spectrum match (PSM) level and protein level employing a target-decoy approach using reversed protein sequences. The minimal peptide length was defined as 7 amino acids, the match-between-run function was disabled. For full proteome analyses carbamidomethylated cysteine was set as fixed modification and oxidation of methionine and N-terminal protein acetylation as variable modifications. Proteins were quantified using Label Free Quantification (maxLFQ28). The maxLFQ intensity was log transformed before downstream analysis. T-tests were used in the differential analysis (using R version 3.6.3). The false discovery rates (FDRs) were calculated from the p-value using the Benjamini-Hochberg method [148]. Proteins with FDR < 0.05 and unique proteins (proteins which were present in all three replicates of one sample, but not present in one of the three replicates of the other sample) were selected as significantly differentially expressed proteins and passed to the DAVID functional annotation [149] for enrichment analysis. To simplify the presentation, proteins in GO terms with FDR below 0.05 were roughly classified according to keywords using the UniProt database [150] related to protein expression and energy regeneration and their influence on gene expression was interpreted accordingly.

3.4.10 Plate reader bulk measurements for genetic circuit

High copy number plasmids (iGEM part pSB1A3) containing the genetic circuit were purified using a Qiagen Plasmid Midi Kit and afterwards phenol chloroform

precipitated (plasmid concentrations are listed in Table 3.2). We kept the TXTL sample composition the same as described in section 3.4.4.

Signal of fluorescent proteins was monitored with a plate reader (FLUOstar Omega, BMG Lab Tech) at a temperature of 29 °C. Fluorescence measurements were performed every 3-6 min using the corresponding filter sets for RFP and GFP. Time traces were background corrected with blank values and end levels were determined.

Plasmid codes for...	Concentration used in TXTL test (nM)
pT7 T3R5-GFP & pT7 T3R5 - Tr1 - ribozyme - AT2	2
pT7 R13 - mKate2 & pT7 R13 - Tr2 ribozyme - AT1	2
pT7 T3R5 - AT2	3
pT7 R13 - AT1	4
pTet - TS1 - T3R5	1.3
pLac - TS2 - R13	2

Table 3.2: Circuit plasmid concentrations used for TXTL tests.

3.4.11 Protein purification

pSB1A3 plasmids coding for the fluorescent proteins mScarlet-I-6His, mVenus-6His, GFPmut3-6His and mTurquoise-2-6His under the control of a T7 promoter were transformed to BL21 (DE3) pLys. TetR was overexpressed in the same strain after transformation of a pSB1K3-pT7-TetR-6His plasmid, which was a kind gift from Dr. Katharina Häußerman. A plasmid coding for LacI including a C-terminal His-tag controlled by a pBAD promoter (Addgene Plasmid #46394) was transformed to Top10 cells.

Overnight cultures were diluted 1:100 (culture volume 500 mL), grown to OD 0.4 and induced with 1 mM IPTG or 0.2% L(+)-Arabinose. Cells were harvested on the next day by centrifugation (15 min, 4 °C, max. speed 4600 rcf, Rotanta 460R, Hettich) and resuspended in 20 ml buffer A (100 mM Tris, 300 mM NaCl, 20 mM imidazole, pH 8.0) including 1mM benzamidine and 1 mg/ml lysozyme. Cells were lysed using a sonicator (6x: 50% amplitude, 30 s pulse and 15 s waiting time) after 20 min incubation on ice. The cell lysate was centrifuged at 10,000 rcf for 30 min and loaded on a pre-equilibrated Ni-NTA column. Protein elution was performed on a FLPC (ÄKTA pure, GE Healthcare) using a standard elution protocol (5 CV washing with buffer A, 15 CV gradient, 0-100% buffer B (100 mM Tris, 300 mM NaCl, 500 mM imidazole, pH 8.0), 5 CV of 100% buffer B). Fractions containing the desired protein were collected and buffer was exchanged to storage buffer (TetR: 20 mM sodium phosphate, 50 mM NaCl, 30% glycerol, pH 7.2, 1 mM benzamidine;

fluorescent proteins and LacI: 200 mM Tris, 200 mM KCl, 10 mM EDTA, 3 mM DTT, 30% glycerol, pH 7.4). Proteins were stored at -80 °C. Protein concentration was determined using a spectrophotometric measurement (Nanophotometer, Implen) and purity was confirmed by performing a SDS PAGE. Dilution series of fluorescent proteins served during plate reader measurements as calibration standards to convert units from a.u. to mg/mL.

Chapter 4

Transcriptional interference in toehold switch-based RNA circuits

E. Falgenhauer, A. Mückl, M. Schwarz-Schilling, and F. C. Simmel

E.F., M.S.-S., and F.C.S. initialized the project. E.F. and F.C.S. planned the project. E.F. and A.M. performed the experiments. E.F., A.M., and F.C.S. wrote the manuscript, and all authors discussed the results and commented on the manuscript.

In the following, we focus on genetic constructs with toehold switches (TS) cis-encoded with their triggers (Tr) and anti-triggers (AT). In detail we activate the expression of a fluorescent protein from a toehold switch by inducing the production of trigger RNA from an inducible promoter. Counteracting this activation process, the production of an anti-trigger RNA with a sequence complementary to the trigger is used to buffer away trigger molecules and correspondingly reduce protein production. We explore the influence of promoter orientation on the performance of our gene circuits in the context of transcriptional interference. Based on our results, we combine the best performing designs, and use them to implement toehold switch-based logic gates in *E. coli*.

4.1 Results and discussion

4.1.1 Modulation of toehold switching via antisense trigger molecules

The fundamental circuit motif utilized in our work comprises a toehold switch to-gether with its cognate trigger and anti-trigger RNA molecules. As shown in Figure 4.1A, Tr can disrupt the stem of the TS via toehold-mediated strand invasion, which exposes the RBS and start codon initially sequestered in the hairpin and thus activates translation of the mRNA sequence (which, in this case, codes for a fluo-rescent protein). AT can counteract the Tr (and thus reduce reporter expression) by sequestering excess Tr via direct hybridization, or by removing Tr from activated trigger-toehold switch complex via toehold-mediated strand displacement.

For an initial characterization of this circuit motif, we tested linear fragments cod-ing for TS regulating a YFP reporter (mVenus), Tr, AT or two ATs separated by a ribozyme under the control of a T7 promoter in a commercial cell-free expression system (PURExpress, Figure 4.2). TS can be activated by co-transcription of Tr and this activation can in turn be repressed by the presence of AT. The signal level reaches the leaky expression level of the TS, when the AT template is provided in double excess compared to the Tr template concentration. The same result can be achieved with a template that codes for two concatenated ATs separated by a ri-bozyme, in which case the AT template does not have to be added in excess.
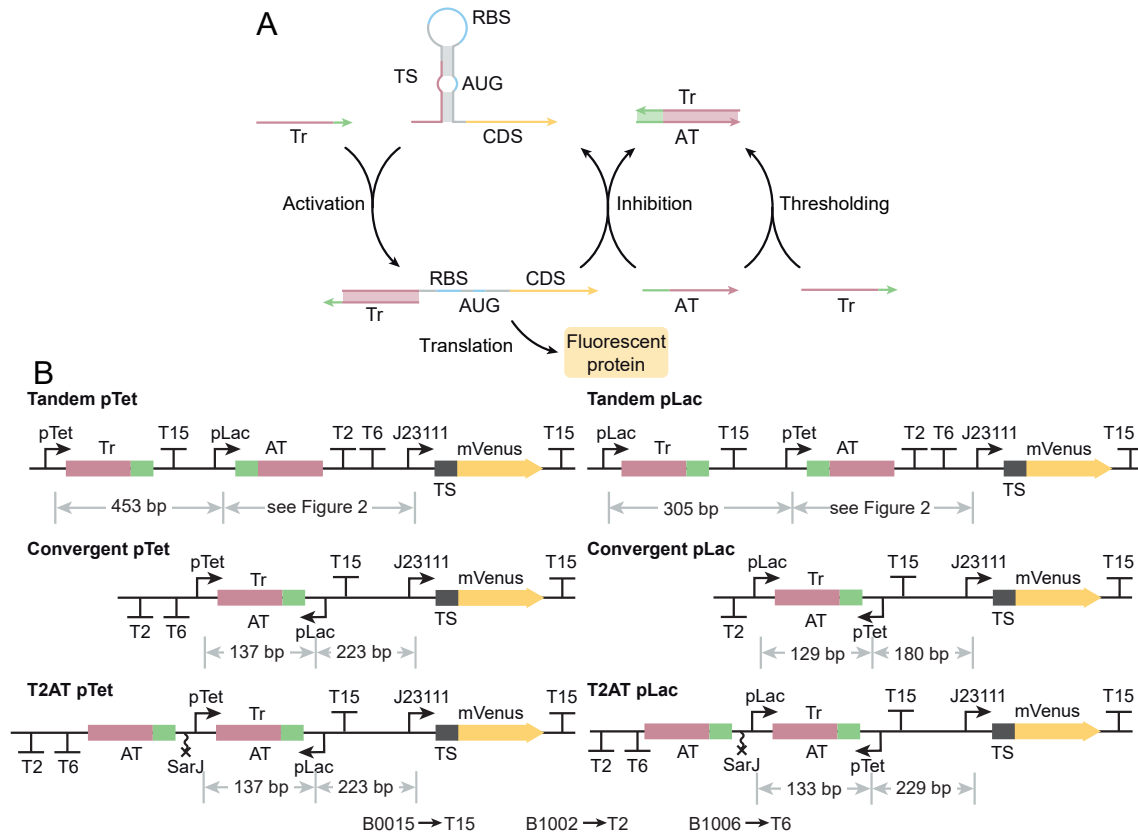
Figure 4.1: (A) In the inactive state of a TS, the ribosome-binding site is hidden in an RNA hairpin loop and is thus inaccessible for the ribosome. Tr binds the toehold, opens the secondary structure via a strand displacement process and therefore activates gene expression. AT can counter-act Tr by sequestering excess Tr via direct hybridization (thresholding), or by removing Tr from an activated trigger-toehold switch complex via toehold-mediated strand displacement. (B) Promoter arrangements of the pTet (left column) and pLac designs (right column). Several tandem designs without overlapping transcription cassettes were tested as reference systems. The versions differ in promoter distance and terminators upstream of the constitutive J23111 promoter. In the convergent design the trigger/anti-trigger sequence is embedded between the two promoters, the transcription terminators are downstream of the converging promoters. A second convergent promoter design was created (designs in the bottom line), which is extended in one direction by a self-cleaving ribozyme and a second anti-trigger sequence. The promoter distances are indicated (the distances are measured between the centers of the promoter sequences to be independent of promoter orientation). Compared to the pTet designs (left), in the pLac designs (tandem pLac, convergent pLac, and T2AT pLac) shown on the right the pTet promoters were exchanged by pLac and vice versa.
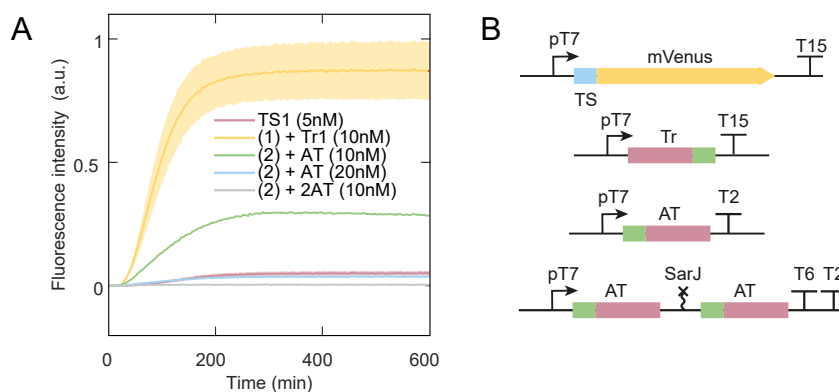
Figure 4.2: (A) TXTL test in PURExpress with linear templates coding for toehold switch, trigger or anti-trigger under the control of a T7 promoter. TS can be activated, if Tr is cotranscribed and deactivated if also AT is cotranscribed. The signal level reaches only the leaky expression level of the TS for a double excess of AT template compared to Tr template concentration. The same signal level is reached for a template coding for two ATs in a row which is added in the same concentration as the trigger template. The data were obtained from n=3 biological replicates, the colored shadow represents the standard error of the mean. (B) Linear templates used for the experiment shown in (A).
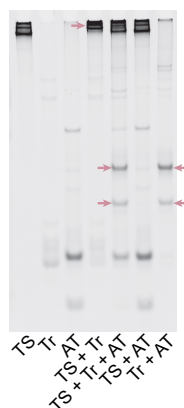
Figure 4.3: Purified toehold switch (TS), trigger (Tr) and anti-trigger (AT) RNA in an electrophoretic mobility shift assay (EMSA). TS RNA shows three bands possibly due to different secondary structures. Tr RNA shows one band and AT RNA shows tow bands corresponding to transcripts terminated after either the terminator or read throughs, terminated by the end of the linear transcript. Hybridization of TS and Tr cause a shift of the TS band (red arrow). Hybridization of Tr and AT form two new bands in presence and absence of TS. Interactions between TS and AT were not observed. Figure was adapted from [144].

We further confirmed the functionality of TS, Tr and AT in an electrophoretic mobility shift assay (Figure 4.3). TS, Tr and AT RNA were *in vitro* transcribed and purified after degradation of the template. TS, Tr and AT RNA were loaded in separate or after performing a hybridization reaction on a native PAGE. TS shows three separate bands, which might be a result of different secondary structures. Tr shows as expected one band. Two bands observed for AT might be a result of termination at the terminator or at the end of the linear template. Hybridization products can be observed for Tr and TS and also for Tr and AT. Undesired hybridization between TS and AT was not observed.

4.1.2 Promoter arrangements for trigger and anti-trigger RNA

We next tested the effect of triggers and anti-triggers on the toehold switch in *E. coli* by cloning the corresponding genetic elements onto a single plasmid ("in *cis*"). As our standard designs we chose tandem promoters where we terminated each transcript upstream of the following promoter (Figure 4.1B). In such designs different types of TI effects are plausible, which are either based on an enhanced local RNAP concentration close to active promoter sites or on collision or passage effects caused by RNAPs, which read through a terminator and elongate over a downstream promoter. We tested these effects in several versions with altered promoter distances and terminator arrangements. As alternative designs we encoded the sequence complementary Tr and AT molecules using two convergent promoters flanking the trigger sequence similar as in naturally occurring cis-encoded antisense regulation. As our cell-free characterization experiments indicated a beneficial effect of excessive AT concentrations, we also cloned a convergent promoter design (T2AT), in which antisense transcription would result in the generation of two anti-trigger molecules, while sense transcription would only produce a single trigger. The two anti-triggers on the antisense transcript were separated from each other post-transcriptionally using a self-cleaving ribozyme as an insulator [74]. The functionality of this 2AT design was also confirmed in cell-free expression experiments using the PURExpress system as described above (Figure 4.2).

While the toehold switch controlling the expression of a YFP reporter protein was transcribed constitutively (iGEM part J23111, Figure 4.1B), we regulated the transcription of trigger RNA using an aTc inducible Tet promoter (pLTet-O1, referred to as pTet) and the transcription of anti-trigger using an IPTG inducible Lac promoter (pLLac-O1 referred to as pLac) adapted from Ref. [151]. We also tested the same

promoter arrangements as in the "pTet designs" with exchanged promoters (referred to as "pLac designs"). We expected a difference in reporter expression due to the differences in promoter strength and the resulting TI effects. The promoter strength of J23111 is about 300% higher compared to pTet (Figure B.1) and pTet is about 30% stronger than pLac (Figure B.2).

The corresponding circuit plasmids were each transformed into *E. coli* strain DH5a Z1, which endogenously codes for LacI and TetR, and the cells were subjected to varying inducer concentrations from a dilution series of the Tr inducer in the absence of the AT inducer. The fluorescence intensity normalized by the absorbance signal (FI/Abs) after 240 min of induction was fitted by the Hill curve Eq. 4.1, where α denotes the scaling factor (fold change), β the offset of the curve (baseline expression), k the induction threshold and n the Hill coefficient. Additionally the repression curves of the designs were measured (at maximum trigger induction and increasing AT inducer concentration) and the data were fitted with the Hill function given in Eq. 4.2.

$$\left(\frac{\text{FI}}{\text{Abs}}\right)_{240\,\text{min}} (c_{\text{Tr inducer}}) = \alpha \cdot \frac{1}{1 + \left(\frac{k}{c_{\text{Tr inducer}}}\right)^{\text{n}}} + \beta \qquad (4.1)$$

$$\left(\frac{\text{FI}}{\text{Abs}}\right)_{240\,\text{min}} (c_{\text{AT inducer}}) = \alpha_2 \cdot \frac{1}{1 + \left(\frac{c_{\text{AT inducer}}}{k_2}\right)^{\text{n}_2}} + \beta_2 \qquad (4.2)$$

4.1.3 Comparison of the tandem reference systems

In the tandem designs all promoters point in the same direction and Tr and AT are transcribed from either a pLac or a pTet promoter, while the toehold switch controls the expression of an mVenus reporter and is transcribed from a constitutive J23111 promoter. As shown in the schemes in Figure 4.4A, the inducible promoters are separated by a long distance of over 400 bp for the pTet designs and of over 300 bp for the pLac designs including a strong double terminator (iGEM part B0015, termination strength >98%). The constitutive promoter and its upstream promoter are separated by either a shorter or longer distance in combination with either a single (iGEM part B1002, termination strength 98%) or a double terminator (B1002 in combination with iGEM part B1006, termination strength 99%). The induction curves for the long DT, long ST and short DT versions overlap for the pTet and pLac designs, whereas in both cases the short ST versions show an upregulated signal (Figure 4.4B and D). The overall signal obtained from the pTet designs is reduced by

more than 50% compared to the pLac designs, even though in this case transcription of the trigger is regulated by the 36% stronger pTet promoter. In Figure 4.4C, the sum of the offset and scaling factor, as a measure for the maximum signal of the designs, is plotted against the distance between the active Tr promoter and J23111. Both the pLac and pTet versions show a relatively constant expression level over a large distance range (700 to >800 bp for pTet and about 550 to 750 bp for pLac). Remarkably, however, an increase in the maximum signal can be observed, when comparing the short DT and short ST designs, which are associated with only relatively small changes in promoter distance (35 bp for the pTet designs and 60 bp for the pLac designs), suggesting a "threshold distance" below which the transcription units start to influence each other. Note that this effect cannot be explained by terminator read-through and subsequent occlusion, collision or sitting duck interference effects as the AT transcription cassette is actually inactive in the experiments shown in Fig. 4.4 (no AT inducer is present).

We hypothesize that the observed effect is the result of an increased local RNAP concentration around active promoter sites, which can 'overlap' for short distances and thus cause an upregulation of the stronger promoter (in this case, J23111). The length of the threshold or 'overlap' distance is determined by the promoter strength and is expected to be longer for a stronger promoter. To further support our hypothesis, we measured the mRNA level of TS-mVenus transcribed by J23111 for the long ST vs. short ST pLac versions using RT-qPCR (Figure 4.5). Indeed, the short ST version shows an about 30% higher RNA level confirming the upregulation of J23111. The overall lower signal of the pTet designs might be explained by a resource-sharing effect as pTet recruits more RNAPs than pLac, and thus less RNAPs are available for TS-mVenus transcription.
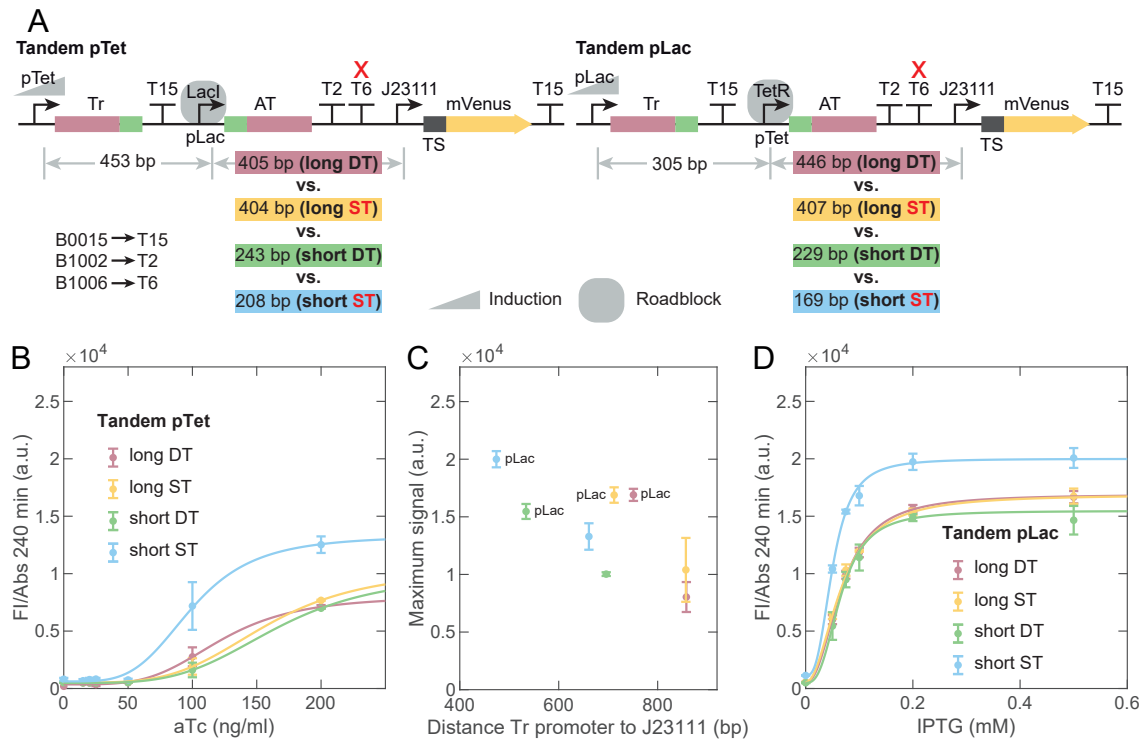
Figure 4.4: Comparison of the tandem reference systems. (A) We tested different circuit designs with either a short or a long distance between the constitutive promoter J23111 and the promoter immediately upstream in combination with a single (ST) or a double terminator (DT). In each case, the inducible promoters (pTet and pLac) were separated by a large distance (>400 bp in the pTet designs (left) and >300 in the pLac designs (right)) and a double terminator. (B) Induction curves of the pTet designs in absence of AT inducer, showing an upregulation of the short ST design. (C) Maximum signal of the designs (sum of the Hill fit parameters α and β) as a function of the distance of the two active promoters. Data points belonging to the pLac designs are indicated. (D) Induction curves for the tandem pLac designs, again showing upregulation of the short ST design. The data were obtained from n=3 biological replicates, error bars represent the standard error of the mean.
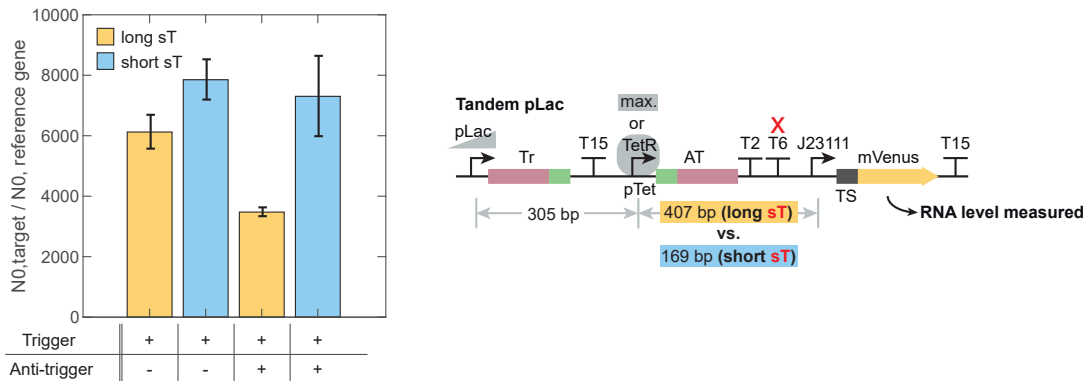
Figure 4.5: The TS-mVenus-RNA level transcribed by the constitutive J23111 promoter is shown for Tandem pLac long ST and Tandem pLac short ST after induction with 0.5 mM IPTG and 0 ng/ml or 200 ng/ml aTc. The RNA levels are upregulated for both inducer combinations for the short ST design. Lower RNA levels were detected for long ST in presence of both inducers.

We next studied the repression curves of the different designs, for which we varied the AT inducer concentration, while trigger transcription was maximally induced (Figure 4.6). We observed a constant absolute reduction in signal in the range of 4000-6000 a.u. (repression capability α_2, Figure 4.6C) for all designs except for Tandem pTet short ST, and constant ON/OFF ratios $((\alpha_2 + \beta_2)/\beta_2)$ of around 3 for all pTet designs and of around 1.4 for the pLac designs (Table B.1). The higher ON/OFF ratios of the pTet designs are a consequence of the comparatively low ON signals. In contrast to the pLac designs, the promoter distance of tandem pTet short ST seems to again fall below a threshold distance for which the repression capability shifts to a significantly higher level.

We again measured the TS-mVenus mRNA level for pLac long ST and short ST (Figure 4.5). Compared to the mRNA level in the presence of Tr inducer alone, the level decreases by over 50% in the presence of both Tr and AT inducer for the long ST design. By contrast, the level for the short ST design stayed approximately constant. Despite the strong effects on the TS-mVenus mRNA level, the fluorescence readout seems to mask these for the pLac designs, which might be related to the higher promoter strength of pTet and the resulting excess of AT compared to Tr. Multiple effects contribute to the observed expression levels and the degree of repression by the anti-triggers. Next to effective differences in the ratios between triggers, anti-triggers and the toehold switches caused by the different promoter

strengths, additional local RNAP concentration effects resulting from an interplay of the pLac and pTet promoters may play a role. Furthermore, resource sharing effects between three (instead of two) active promoters might have influence on the repression capability. Unfortunately, we were not able to study these effects in greater detail as the short transcript length of Tr and AT (< 100 nt without terminator) did not allow a quantitative comparison via qPCR, and the analysis of the TS-mVenus-RNA level alone does not provide sufficient information.
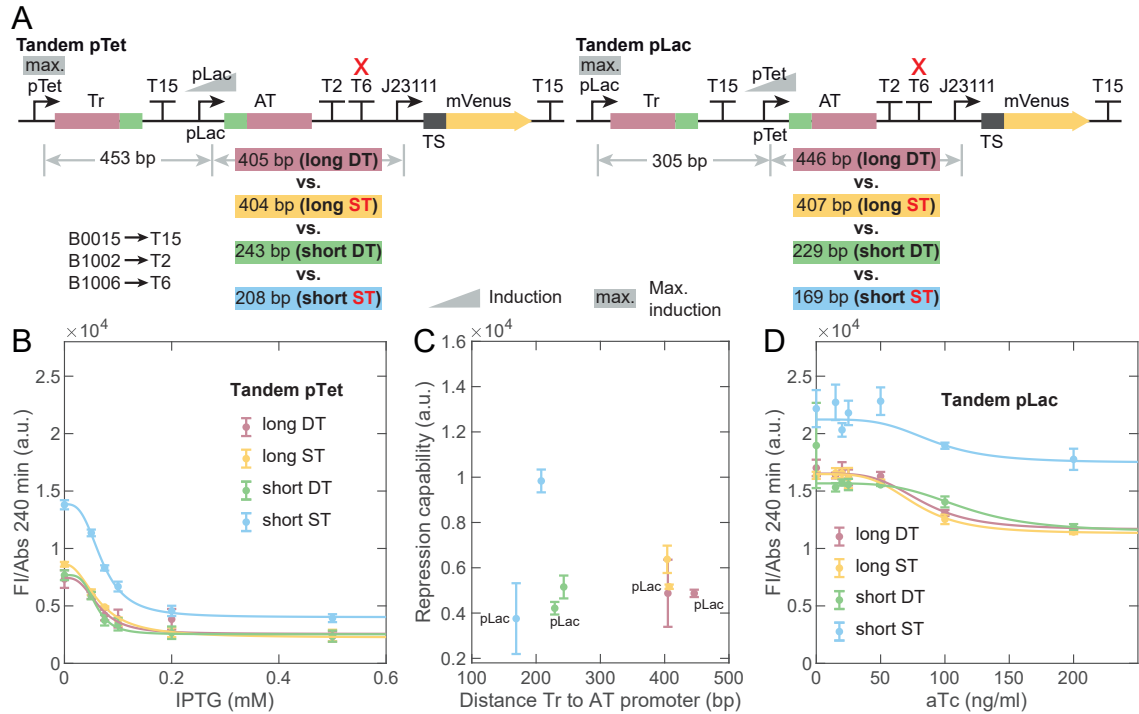
Figure 4.6: : Repression curves for all Tandem designs. (A) Schematic representation of the experiment. AT transcription was induced in presence of a fully induced Tr promoter. (B) Repression curves for tandem pTet designs. (C) All designs except tandem pTet short ST show the same repression capability (ON-OFF, fit parameter α_2). (D) Repression curves for the pLac designs.
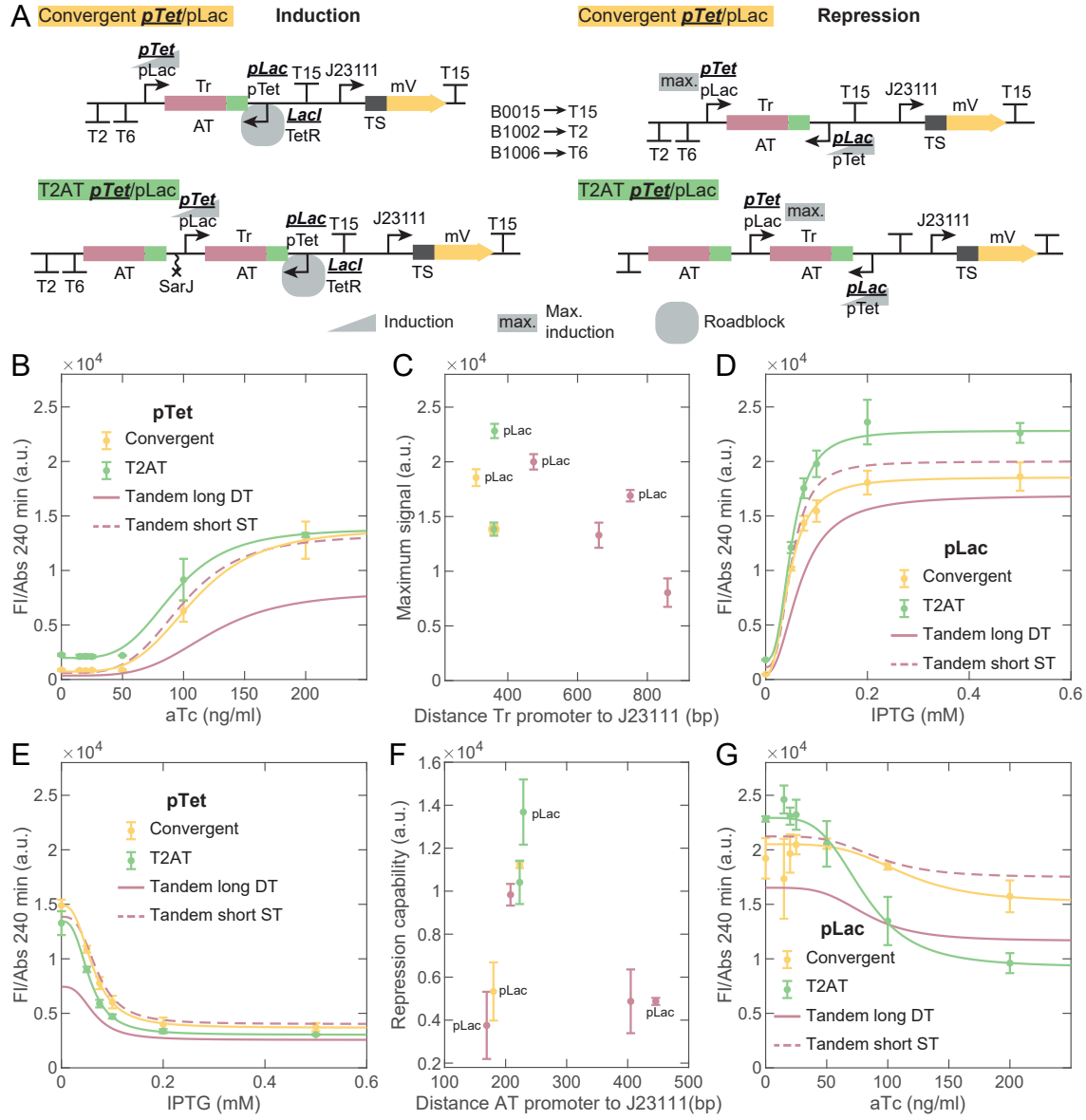
Figure 4.7: Characterization of convergent designs. (A) Schematic representation of gene constructs and experiments performed. Left: Induction curves for all convergent designs are measured in the presence of a roadblock (repressor) bound to the inactive AT promoter. Right: The repression capability of the designs is tested by induction of AT transcription at maximum trigger transcription. (B) The induction curves for the Tet promoter in the presence of the LacI roadblock are compared for convergent pTet and T2AT pTet. As a reference, fit curves for tandem pTet long DT and short ST are plotted in the same graph. The convergent pTet and T2AT pTet design are upregulated compared to the tandem pTet long DT design, but reach a comparable signal level as tandem pTet short ST. (C) Distance dependence of the maximum signal ($\alpha + \beta$) of the pTet and pLac designs shown in B and D. (D) Induction curves for the Lac promoter in the presence of the TetR roadblock. Both convergent pLac designs are upregulated compared to tandem pLac long DT. T2AT pTet even exceeds the tandem pLac short ST signal. (E) Repression curves for the pTet designs. The absolute signal repression capability (α_2, ON-OFF) for convergent pTet and T2AT pTet is comparable to tandem pTet short ST. (F) Overview of repression capability as a function of promoter distance for the pTet and pLac designs. (G) Repression curves for the pLac designs. Convergent pLac shows a similar signal decrease as tandem pLac short ST. T2AT shows the steepest decrease in signal of all designs. The data were obtained from n=3 biological replicates; error bars represent the standard error of the mean.

4.1.4 Induction and repression curves of the convergent designs

We next investigated circuit designs with convergent promoter arrangements, which enable to shorten the distance between Tr promoter and J23111 as the AT transcription cassette is now located on the antisense strand. The distance between J23111 and the closest upstream promoter varies between 180 and 229 bp including a DT (Figure 4.1), and with a separation of 129 -137 bp also the distance between the convergent promoters was kept very short. We again first focused on the trigger induction curves in the absence of an inducer for anti-trigger transcription. Repressor proteins bound to a converging promoter can cause potential roadblock effects. This effect is not expected for the tandem designs as in their case the first transcript is terminated upstream of the second promoter. Previous studies have shown that roadblocks can either cause premature transcriptional termination or can be cooperatively dislodged by several polymerases, if the RNAP flux is high enough (at high promoter strength). In contrast to a study by Bordoy et al. [128], whose transcripts overlapped just for a part of the sequence, for our designs (Figure 4.1B) we would expect to obtain fully functional trigger RNA sequences even in the case of transcriptional termination at the roadblock as the full trigger sequence is up-

stream of the roadblock. We therefore did not expect a direct effect of the roadblock on the system performance. Indeed, convergent pTet and T2AT pTet both reach a signal, which is comparable to the Tandem short ST design (Figure 4.7B). T2AT pLac shows a higher signal compared to convergent pLac and Tandem pLac short ST (Figure 4.7D). Whereas the local RNAP concentration effect seems to saturate (no further increase in maximum signal for lower distances) for the pTet designs, the results are ambiguous for the pLac designs (Figure 4.7C). The repression curves of convergent pTet and pLac overlap with their corresponding tandem short ST designs. T2AT shows a different trend for pTet and pLac. Whereas T2AT pTet has the same repression capability as tandem pTet short ST, T2AT pLac reaches the highest signal decrease of all tested designs. Again the situation is more difficult than for the induction curves as three active promoters are present. In addition to the abovementioned resource sharing and local RNAP concentration effects also RNAP collision and passage effects are expected in these designs. As we observed the same distance dependence of convergent pTet and convergent pLac as for the corresponding tandem short ST designs (Figure 4.7F), however, these additional effects are either negligible or are masked by the system readout.

In summary, our results on the trigger induction curves and the anti-trigger repression capability suggest that the dominant transcriptional interference effect in our circuit designs is caused by local changes in RNAP concentration, which is influenced by the promoter distance. For short promoter distances, we observe a significant increase in maximum expression, and the "critical distance" below which such an enhancement is observed appears to depend on the promoter characteristics (between 660 and 700 bp for the combination of pTet and J23111 and 530 to 470 bp for the combination of pLac and J23111). We were not able to detect effects, which could be caused by premature termination of transcripts based on roadblock effects, which was an expected result as our convergent designs always enclosed the full trigger sequence. The pLac designs benefitted most from the TI effects and showed up to two-fold higher maximum signals (induction curves in the absence of AT) compared to the pTet designs even though Tr transcription was regulated by the weaker pLac promoter. The highest ON/OFF ratio (repression) among all pLac designs was obtained for T2AT pLac (≈ 2.5), which benefited from the double AT transcript. For the pTet designs the results of tandem short ST, convergent and T2AT are comparable with ON/OFF ratios around 4.

4.1.5 Implementation of a two-input two-output logic gate
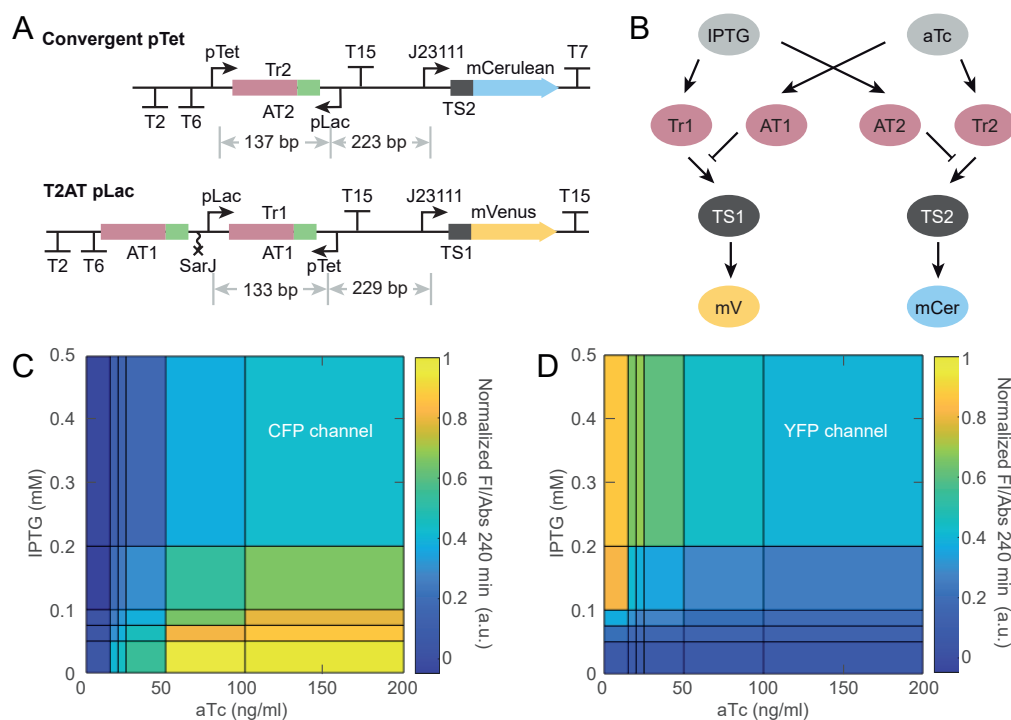
Figure 4.8: Two-input two-output logic gate. (A) A convergent pTet design based on TS2, Tr2, AT2 and a CFP reporter is combined with a T2AT pLac design, resulting in a two-input two-output logic gate. Each scheme represents a DNA fragment on a plasmid. (B) Regulation scheme of the circuit. Tr1 and AT2 are transcribed under IPTG induction, Tr2 and AT1 under aTc induction, and together regulate translation from TS1 and TS2. YFP is expressed if TS1 and CFP is expressed if TS2 is active. (C) Experimentally determined heat map representation of CFP expression from the gate after 240 min of induction. For high aTc but low IPTG concentrations the CFP reporter is expressed to a high level. (D) Heat map of YFP expression from the gate after 240 min of induction. YFP is highly expressed for high IPTG but low aTc concentrations.

Based on these results, we designed a logic gene circuit, for which we utilized two separate toehold switches together with their triggers and anti-triggers (Figure 4.8A and B). One trigger/anti-trigger pair was based on the T2AT pLac design, while the other utilized a convergent pTet design. The toehold switch controlling the expression of YFP was identical to that used to investigate the TI effects above, while the second switch regulating a CFP reporter was designed with an altered sequence (version 2[111]). As a result we obtained a gene circuit, in which Tr1 and

AT2 are expressed under IPTG induction, while Tr2 and AT1 are expressed by aTc induction, i.e., each inducer activates one of the toehold switches and simultaneously represses the other. Taken together, the circuit acts as a two-input two-output logic gate implementing the function:

$$\text{mVenus} = \text{IPTG NIMPLY aTc} \qquad (4.3)$$

$$\text{mCerulean} = \text{aTc NIMPLY IPTG} \qquad (4.4)$$

where NIMPLY is the negation of the IMPLY function. The experimentally determined gene regulation function of the two-input two-output gate has two regions, where one of the reporters is dominant, while the other is expressed at a very low level, and a third region, where both reporters are co-expressed at a half-maximal level (Figure 4.8C and D). Even though the signal levels of the two subsystems (convergent pTet and T2AT pLac) were lower when operated in combination compared to the signals obtained for the individual circuits (most likely due to resource sharing effects), the two linked toehold switch systems display almost the same ON/OFF ratios (2.2 ± 0.2 for the YFP subsystem and 3.3 ± 0.2 for the CFP subsystem) as when studied in isolation (Table B.1). Due to the different reporter proteins, we normalized the fluorescence signals to the maximum measured signal in each channel. We note that the circuit design shown in Figure 4.8B could, in principle, be used to implement an XOR gate. Identifying both reporters with each other (e.g., interpreting "fluorescence" as a single output), the output of the gate can be reformulated to:

$$\text{mV OR mCer} = (\text{IPTG NIMPLY aTc}) \text{ OR } (\text{aTc NIMPLY IPTG}) = \text{IPTG XOR aTc} \qquad (4.5)$$

As the signal in the high induction region (high IPTG and high aTc) is about half-maximal (rather than zero), however, the sum of the signals in the co-expression region would reach the signal level of the dominant region, and thus the experimentally realized gate would rather behave like a logic OR gate. We finally set out to demonstrate dynamic switching between the different states of the logic gate. To this end, we chose four distinct points of the gene regulation landscape (corresponding to the logical input combinations 00, 01, 10, 11) as the starting or end points of a switching process and characterized all 16 possible combinations of the initial and final states using plate reader measurements (Figure 4.9). In four of the possible combinations (no change in inducer input, corresponding to the diagonal elements in Figure 4.9) the system stays idle at its respective starting point, as expected. The measured signal levels show small disturbances over time, but the

states are clearly distinguishable for all time points. For all other transitions except the transition from 11 to 00 (which shows an initial increase in the CFP signal, followed by a convergence of both outputs to a low fluorescence state), the system performs as expected and correct switching can be observed within the first 250 min after change of the inducer concentrations. However, the end levels are lower than expected for the low to high transitions and higher than expected for the high to low transitions (signals do not converge to 1, 0.5, or 0, respectively). Transitions from a high to a lower expression state can be potentially improved by adding degradation tags to the reporter proteins or by optimizing the growth conditions to speed up the dilution of the reporters by cell division. This would be particularly beneficial for the transition 10 (IPTG = 0.5 mM, aTc = 0 ng/mL) to 11 (IPTG = 0.5 mM, aTc = 200 ng/mL) as the final state (with the YFP level too high) is ambiguous. Transitions from low to high expression states can also potentially be improved by optimizing growth conditions and oxygen supply to support protein expression and maturation.
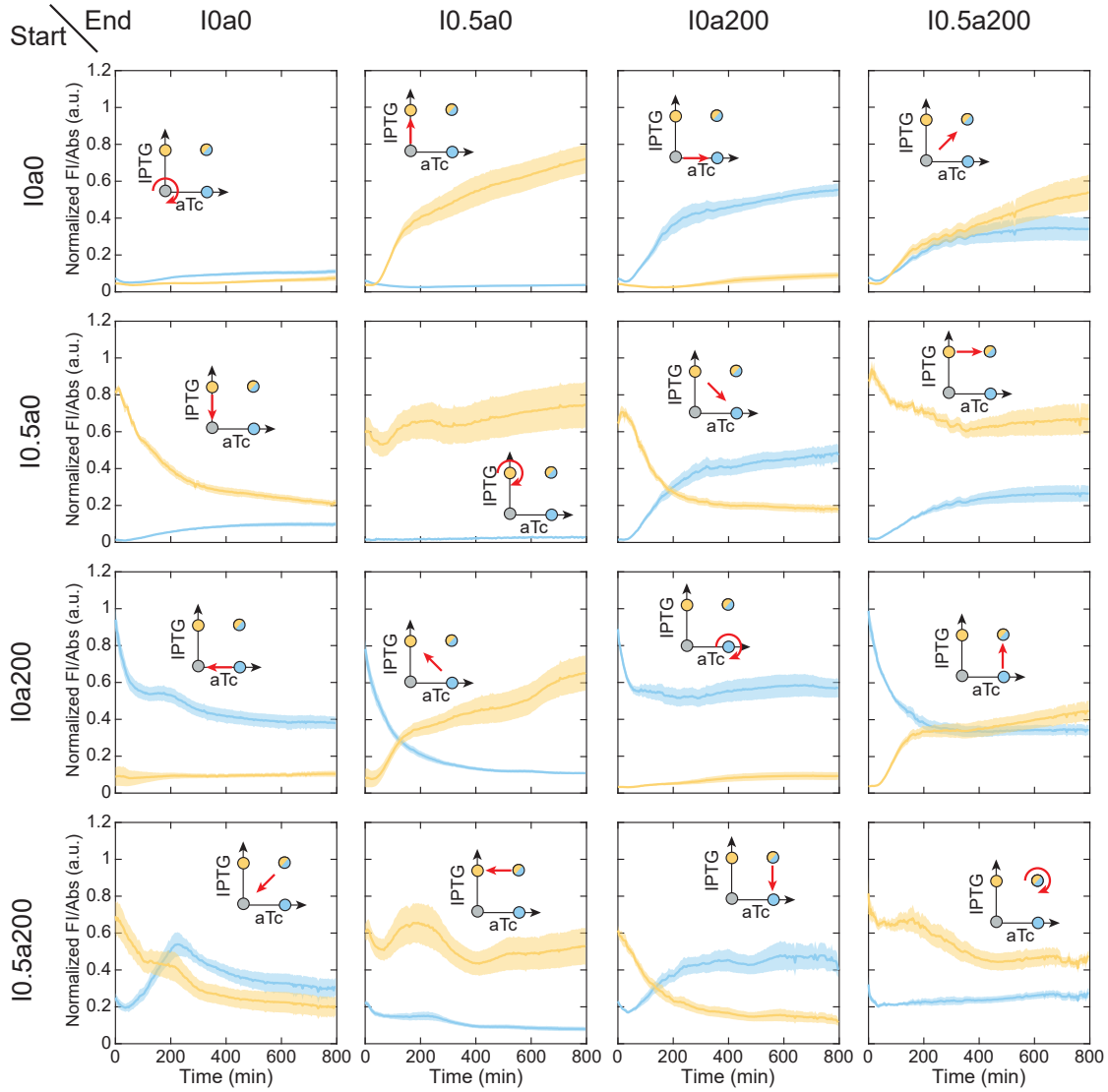
Figure 4.9: Switching experiment characterizing transitions between the different states of the logic gate. We use a shorthand notation for the inducer combinations, where Ixay denotes concentrations of x mM IPTG and y ng/ml aTc. Inducer concentrations corresponding to the logical input combinations 00 (I0a0), 10 (I0.5a0), 01 (I0a200) and 11 (I0.5a200) were chosen as start or end point of switching experiments and all possible transitions were tested. Mean fluorescence time traces of three biological replicates for the mVenus reporter are displayed in yellow, traces for the mCerulean reporter are displayed in blue. The shaded region gives the standard error of the mean. The small icons show the transition path.

4.2 Conclusion

In the present study, we investigated the influence of promoter arrangement on the performance of synthetic riboregulators – toehold switches – that are activated by short trigger RNAs and de-activated by the corresponding complementary anti-trigger RNAs. We particularly focused on potential transcriptional interference effects, which are expected to sensitively depend on the distance and orientation of the involved promoters. Induction curves obtained with an active Tr and TS promoter suggest a locally enhanced RNAP concentration as the main transcriptional interference effect, which appeared to play a role only for distances below a certain threshold distance. In our experiments, this distance was between 660-700 bp for the combination of the pTet and J23111 promoters, and between 470 and 539 bp for the combination of the pLac and J23111 promoters, and therefore depended on the characteristics of the involved promoters. In the presence of a third active promoter controlling the transcription of anti-trigger RNA, TI effects were more difficult to assess. While we measured differences in mRNA levels consistent with a TI effect, these differences were masked at the protein expression (fluorescent reporter) level, preventing a more detailed study of the effects in the presence of the third promoter. The use of RNA trigger molecules and their exactly sequence-complementary anti-triggers within toehold switch circuits naturally suggests their transcription from the same DNA template using convergent promoters. Such a promoter arrangement leads to particularly compact circuit designs, in which the promoter distances can be shortened even further, which may be used to utilize a local RNAP concentration effect. We avoided the production of non-functional Tr and AT molecules resulting from premature termination at roadblocks (repressor proteins) by placing potential roadblock binding sites outside of the coding regions. In general, we observed that when trigger transcription is regulated by the (weaker!) pLac promoter, while anti-trigger transcription is controlled by the (stronger) pTet promoter, up to twofold higher maximum signals can be obtained compared to the opposite promoter arrangement. The best designs were assembled to create a two-input two-output logic gate, which shows fast switching kinetics. In summary, our work demonstrates that TI effects can be usefully implemented in the context of RNA-based toehold switch circuits, and a careful choice of promoter strengths, arrangements and distances can result in a strongly enhanced performance.

4.3 Materials and methods

4.3.1 Chemicals

Unless otherwise noted all chemicals were ordered from Sigma Aldrich, all enzymes including restriction enzymes, ligases, PCR and Gibson Assembly master mixes were ordered from New England Biolabs (NEB), all primers were ordered from Eurofins Genomics, linear gene fragments were ordered from Integrated DNA Technologies (IDT) and all plasmids were sequenced by GATC Services/Eurofins Genomics.

4.3.2 Plasmids

Plasmids were cloned in Max efficiency DH5α cells (original stock from Thermo Fisher Scientific) using a standard restriction-ligation protocol with commercially available restriction enzymes (mainly AatII and EcoRI). Linear gene fragments were ordered from IDT and cloned into the target vectors (iGEM part pSB1A3, iGEM part pSB3C5). The vector pSB1A3 has an ampicillin resistance marker and a pMB1 origin of replication. pSB3C5 is a vector with p15A as origin of replication and has a chloramphenicol antibiotic resistance marker. Design versions with elongated promoter distances were cloned using overhang PCR and a subsequent Gibson Assembly according to the NEB standard protocol. Versions with shorter promoter distances (deletion of terminators) were cloned using round the horn cloning with primers flanking the region to be deleted and amplifying the full plasmid. Final plasmid sequences are listed in Tables B.2-B.14.

4.3.3 Cloning SOPs

4.3.4 Cloning of plasmids
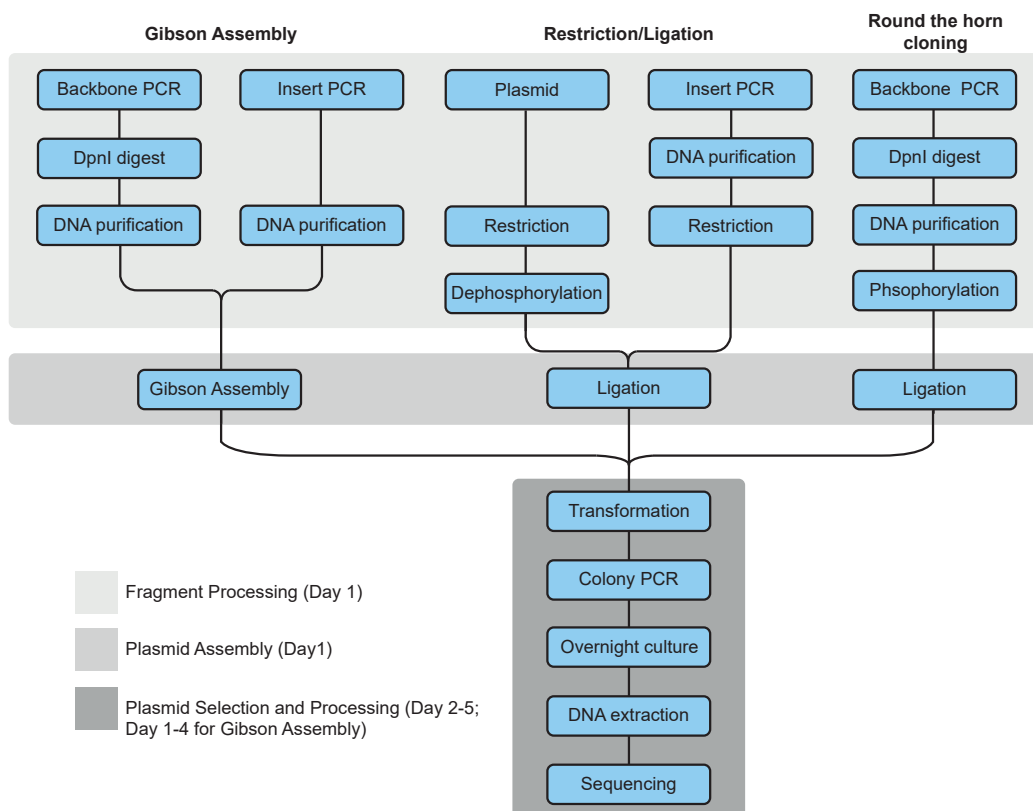
Figure 4.10: Standard procedures for plasmid construction. (a) For Gibson Assembly, insert and backbone require only a PCR step to get homologous regions which act as sticky ends for the plasmid assembly. (b) In the restriction/ligation protocol the plasmid and the insert are restricted separately, then the fragments are mixed and ligated. (c) In order to remove specific regions form a plasmid everything but this part is PCR amplified. The purified product can be ligated and transformed. All presented cloning methods share five steps of plasmid selection and processing.

In Figure 4.10 the standard procedures of Gibsons Assembly, Restriction/Ligation and round the horn cloning are summarized.

Gibson Assembly

For Gibson Assembly backbone and insert are PCR amplified using a master mix (Phusion High-Fidelity (HF) PCR Master Mix with HF buffer, NEB) to add 20-30 basepair overlaps to the sequences to be assembled. The backbone template is DpnI digested and both reaction mixtures (backbone and insert) are purified using a PCR and DNA cleanup kit (QIAquick PCR Purification Kit, Qiagen). A Gibson Assembly master mix provides a mixture of a exonuclease, a polymerase and a ligase. The reaction mix is incubated at 37 °C for 3 min followed by a second incubation step at 50 °C for 30 min.

Restriction/Ligation

The PCR amplified and purified insert is restricted in parallel with the backbone with the same pair of restriction enzymes. The plasmid backbone is dephosphory-lated using Antarctic Phosphatase. After heat inactivation and over night ligation at 16 °C the sample can be transformed.

Round the horn cloning

In order to remove specific regions from a plasmid or in order to include or exchange small DNA fragments (e.g. a promoter) round the horn cloning can be used. The plasmid is PCR amplified using primers, which can have overhangs containing the new insert, point in divergent directions and flank the region, which has to be exchanged or deleted. The plasmid template is digested with DpnI. The purified PCR product is ligated over night.

Plasmid Selection and Processing

The cloning methods share the following five steps of plasmid selection and processing, which take in total four days. The plasmid sample is transformed to DH5α by heat shock transformation or electroporation. Then a colony PCR is performed, the desired colonies are cultured over night (5ml) and the DNA is extracted (QIAprep Spin Miniprep Kit, Qiagen) on the next day. Purified plasmids were sequenced by GATC Biotech (Germany) using mainly the sequencing primers VF2 and VR.

4.3.5 Bacterial strains and culture conditions

Plasmids were transformed to DH5alpha Z1. Cells from glycerol stocks were grown in 5 ml LB medium (Luria/Miller, Carl Roth) containing selective antibiotics and 20 mM glucose at 37 °C overnight and 250 rpm. The following day, cells were diluted 1:100 in M9 medium containing the selective antibiotics and were grown again at 37 °C and 250 rpm to an OD of 0.3, keeping the antibiotic conditions the same.

4.3.6 Plate reader experiments

Cells (14 µl) were transferred to 384 well plates (ibidi) and induced with in total 1µl (to avoid inhomogenous dilution) of IPTG (final concentration: 0mM, 50 nM, 75 nM, 0.1 mM, 0.2 mM and 0.5 mM) in presence of either 0 ng/ml (pLac designs induction) or 200 ng/ml aTc (pTet designs repression) or with aTc (biomol, final concentration: 0 ng/ml, 15 ng/ml, 20 ng/ml, 25 ng/ml, 50 ng/ml, 100 ng/ml and 200 ng/ml) in presence of either 0 mM (pTet designs induction) or 0.5 mM IPTG (pLac designs repression) using a liquid dispenser (I.DOT, Dispendix). Fluorescence and absorbance time traces were recorded using a plate reader (Fluostar Omega, BMG Labtech). mVenus fluorescence was recorded using a 500/10 nm excitation and a 540/10 nm emission filter, mCerulean fluorescence was recorded using a 440/10 nm excitation and a 480/10 nm emission filter, absorbance was measured at 600 nm. Blank values of pure M9 medium (bacterial autofluorescence was neglected) were subtracted, the fluorescence levels at 240 min were normalized to the corresponding absorbance values (FI/Abs) and the mean of 3 biological replicates and the standard error of the mean was plotted against the inducer concentration. Hill curves (Equation 1 for induction curves and 2 for repression curves) were fitted to the data and the maximum signal (calculated from $\alpha + \beta$), the repression capability (ON-OFF, α_2), and the ON/OFF ratio (ON: max induction, OFF: max. repression, $\alpha_2 + \beta_2/\beta_2$) was evaluated. Errors were propagated using a quadratic addition (Equation B.1-5).

4.3.7 Logic gate switching experiments

DH5alpha Z1 cells containing the 2I/2O gate plasmid were cultured overnight in four separate culture tubes with LB + glucose medium and the selective antibiotic

(Cm). The cultures were induced with either of four inducer combinations of 0 mM or 0.5 mM IPTG against 0 ng/ml or 200 ng/ml aTc. Cells were diluted 1:100 in M9 medium keeping the inducer conditions and growth conditions the same. At OD 0.3 the cultures were centrifuged (3000 g, 3 min, 4 °C) and resuspended in 1/10 of the starting volume. Inducer conditions were kept constant or changed to one of the three other combinations by diluting 30 µl of bacteria in 270 µl M9 medium containing the corresponding inducers. Fluorescence and absorbance time traces were recorded with a plate reader (Fluostar Omega – BMG Labtech) using 96 well plates (ibidi). mVenus and mCerulean fluorescence was measured using the same settings as mentioned above. FI/Abs time traces are shown in Figure 4.9 (mean of 3 biological replicates, shaded area: standard error of the mean).

4.3.8 Cell-free expression tests

Functionality of TS, Tr, AT and the modified anti-trigger (2AT = AT-Ribozyme-AT) was proven in a cell-free test using linear DNA fragments coding for a T7 promoter and the desired RNA sequences (TS1-YFP, trigger or modified and unmodified anti-trigger). Linear DNA fragments (ordered from IDT) were PCR amplified (Phusion High Fidelity PCR Master Mix, NEB) and purified (Monarch PCR & DNA Cleanup Kit, NEB). Concentrations were measured at 260 nm using a spectrophotometer (NanoPhotometer, Implen). The reaction was set up using a commercial kit (PURExpress, NEB) according to user guidelines, as linear fragments are just stable in absence of RecBCD and DNases. Fluorescence time traces were recorded using a plate reader (Fluostar Omega, BMG Labtec) using the filter set for mVenus as mentioned above.

4.3.9 Electrophoretic mobility shift assay (EMSA)

PCR amplified templates coding for TS1-YFP, Tr1 or AT1 under the control of a T7 promoter were purified using a commercial kit (Monarch PCR & DNA Cleanup Kit, NEB). The concentration was measured (Nanophotometer, Implen) and an in vitro transcription overnight at 37 °C was set up (in total 16 mM rNTPs (4 mM each; Thermo Fisher Scientific), 40 mM Tris (pH 7.9), 18 mM $MgCl_2$, 1 mM DTT, 2 mM spermidine, 15-40 nM DNA template, 250 nM T7 polymerase (self-purified). On the next day DNA was digested with DNase I (NEB) and the RNA was phenol chloroform precipitated. RNA concentration was determined using a

denaturing Urea-PAGE (8 M Urea, 8% polyacrylamide, 37.5:1 acrylamide/bis solution, Biorad) and a commercial RNA ladder as reference (RiboRuler Low Range RNA Ladder, Thermo Fisher Scientific) in combination with the image analysis software ImageJ and its built-in gel analysis toolbox. RNAs were mixed in equal concentrations (TS1+Tr1, TS1+Tr1+AT1, TS1+AT1, Tr1+AT1) in presence of 5 mM MgCl$_2$ and incubated at room temperature for 15 min. Hybridization products and unmixed RNAs were loaded on a native PAGE (8% polyacrylamide), stained with Sybr-Green II and visualized by a gel documentation system (Quantum CX5, Vilber).

4.3.10 RT-qPCR quantitation

For the relative quantification of RNA levels between different constructs with short and long promoter separation (tandem pLac short ST and Tandem pLac long ST versions), an overnight culture was diluted 1:100 in M9 medium containing Carbenicillin and 20 mM glucose. After 3 h at 37 °C in a shaking incubator, an OD value of about 0.3 was reached and cells were induced either with IPTG (0.5 mM) alone which induces Tr1 or a combination of 0.5 mM IPTG and 200 ng/ml aTc which induces Tr1 and AT1. After additional 3h of growth in the shaking incubator at 37 °C, cells were collected at 1 ml at OD 0.812 adjusted for each sample. All cells were centrifuged at 4 °C for 2 min at 5000 rcf. RNA isolation was performed according to the RNA isolation II protocol (Bioline) with the following exception: after the addition of buffer RLY, the samples were flash frozen in liquid nitrogen and stored at -80 °C until the following day and the continuation of RNA isolation procedure. The RNA yield and quality was analyzed with a spectrophotometer (NanoPhotometer, Implen). The values obtained for A260/280 were >2 for all samples. Isolated RNA was flash frozen in LN2 and stored at -80°C until reverse transcription was performed according to the RevertUpII Kit from biotechrabbit with 300 ng RNA for each sample. The cDNA samples were aliquoted and flash frozen in LN2 and kept at -80 °C until RT-qPCR was performed.

The RT-qPCR reactions were performed on a BioRad IQ5 instrument with the following settings: dynamic well factor method by addition of 10 nM Fluorescein to each sample and detection in the FAM channel by cycling 1x 1 min 95 °C, 45x 30 s 95 °C → 15 s 60 °C, as well as melt curve recording between 55–95 °C. The reactions were performed in 10 µl total volume with 4 µl of a 1:100 dilution of the cDNA and LunaScript Universal MasterMix 2x (New England Biolabs) in white

PCR stripes with flat lid (AB-1191, ThermoFisher). Three technical replicates were recorded for each sample (as well as minus reverse transcriptase control and non-template control). The gene specific primers for mVenus and the reference genes can be found in the supporting information file of the original publication [2]. The raw data were extracted and converted to RDML file. The RDML file was analyzed with LinRegPCR tool 35 to determine PCR efficiency and corrected target quantity N0. Amplification bias due to artefacts revealed by melt curves was not considered because of the non-saturating dye used in the master mix. The best stable reference gene was chosen with Normfinder 36 and N0,mVenus was normalized to N0,cysG.

Chapter 5

Discussion and Outlook

5.1 Cell extract project

Summary   In summary we presented the production of active cell extract based on a lysozyme assisted sonication protocol. We determined the optimal lysis settings for *E. coli* cultivated in either shaking flasks or a bioreactor. Our extracts performed in cell-free gene expression tests of reporter proteins comparable to commercial extracts and showed a better performance for PCP purified plasmids used as templates for the TXTL reactions. As a more complex system, we also assembled phages in our extracts and obtained higher phage titers compared to the commercial extract. Finally, we also analyzed the protein composition of selected extracts and found enzymes related to energy regeneration and biosynthesis of small molecules like AA and nucleotides enriched in our extract, while protein related to transcription, translation and processing of nucleic acids were found to be enriched in the commercial extract.

Our produced cell extract can be directly used as a cell-free gene expression system and future projects can take advantage of the high activity e.g. by producing therapeutic phages or testing genetic circuits. One could also adjust the cell extract preparation protocol for larger batch volumes (upscaling). Costs of cell-free gene expression could be even further decreased by fueling the reaction with low cost AA precursors and taking advantage of the highly abundant enzymes needed for their biosynthesis. In the following some of these ideas are briefly presented:

Production of phages for phage therapy   Bacteriophages or simply phages are the natural enemies of bacteria as they are able to infect and mostly also kill bacteria. A substantial fraction of phages is not inherently toxic to other cells than their host. Therefore purified phages can be applied directly to treat bacterial infections

in patients without causing a collateral damage to the patients organs and tissues or the normal microbiota [152]. In face of globally rising antimicrobial resistance, this approach, also known as phage therapy, is of increasing clinical relevance as alternatives to antibiotics. Phage therapy dates back to the codiscovery by Twort [153] and d'Hérelle [154] in the early 1900s. Traditional methods to produce and purify phages include the cultivation of the pathogen bacterial strain, which is potentially harmul to the manufacturer, who might get infected, and also the patient, as pathogens of the host strain, prophages fragments or endotoxins might be coisolated. Therefore manufacturers have to fulfill high regulatory requirements and the *in vitro* assembly of phages is very attractive. Purified phage DNA or synthetic DNA can be added directly to the cell extract, which can be produced of engineered strains with low endotoxin levels (e.g. mutation in myristoyltransferase msbB can reduce lipid A level [155]). Cofactors, needed for the cell-free production, the stabilization of the phage genome (e.g. GamS to stabilize linear DNA [156]) or the GMO-free modification of the phages (e.g. engineered capsid proteins) can be easily added to the reaction, and therefore *E. coli* cell extract might be suited for the production of a wide range of phages with engineerd functions, making the cultivation of a variety of pathogenic bacteria unnecessary. This concept was already presented in 2018 by the iGEM team 'Phactory' [157] hosted in our lab and was the basis for founding the spin-off 'Invitris' [158] by Dr. Kilian Vogele and Dr. Parick Grossmann, who benefited already from our new cell extract preparation protocol.
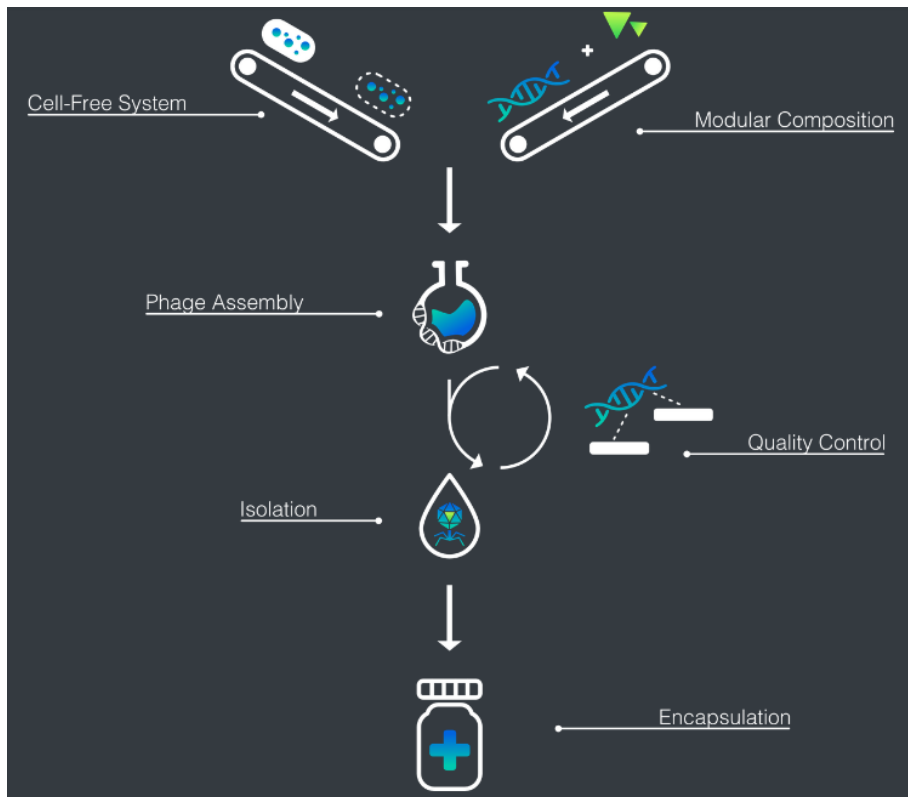
Figure 5.1: Potential workflow for the cell-free prodution of therapeutic bacteriophages suggested by the iGEM team Phactory hosted in our lab in 2018 [157]. Cell-free system is incubated with the phage genome and additional factors needed for the assembly of functional phages. For a therapeutic application on patients, the quality of the isolated phages has to be monitored continuously. In the last step, phages are encapsulated to protect them from gastric acid after oral administration. The figure was adapted from [157].

Upscaling For the production of therapeutic phages in commercial scales, the cell extract batch volume has to be increased significantly and the production process has to be automatized to further reduce batch-to-batch variations. To increase the batch volume several bottlenecks have to be overcome. In the first place, the pellet mass after cell harvest and washing can be increased by optimizing the fermentation process. E.g. Zawada et al. [159] used a defined medium and a glucose feeding strategy designed to meet metabolic demand but avoid acetate production. They reached an OD of up to 50 and reported the production of active cell extract using a high-pressure homogenizer.

The adaption of the fermentation process to typical industrial scales of about 1000 L requires specialized equipment like pre-fermenters and disc-stack centrifuges and the guidance of bioprocess technologists to run the devices. The lysozyme incubation step can be upscaled easily, as this process is only dependent on the concentration of the comparably inexpensive lysozyme. For the sonication step one has to switch to a flow-through sonicator, which solves simultaneously the problem of constant mixing and does not require a splitting of the cell suspension in several aliquots. For the subsequent centrifugation step one has to check, whether a disc-stack centrifuge is suited to separate the cell extract and the cell debris or whether centrifuge bottles, which would limit the process throughput, have to be used. The run-off reaction might be performed in flow through mode, where the pipes are embedded in temperature regulated water baths (37 °C and 4 °C for cooling the extract down again). For the dialyis step dialysis tubings with larger diameters are available. Obviously, tests with intermediate batch volumes would be necessary to evaluate the feasibility of the suggested workflow.

5.2 Transcriptional interference and toehold switch project

Summary   We tested toehold switches with their cognate triggers and anti-triggers in different promoter arrangements and evaluated the effect of transcriptional interference on the system performance *in vivo*. In detail we transcribed TS constitutively and chose inducible promoters for Tr and AT (pLac and pTet). We observed for our induction curves with an active toehold switch and trigger promoter a rise in reporter signal as soon as a certain promoter distance was undercut. This observation was mainly caused by enhanced local RNAP concentrations around active promoters, which started to interfere below the threshold distance. We supported this thesis by measuring mRNA levels.
The sequence complementarity of trigger and anti-trigger suggested a more compact circuit design with convergent promoters and shorter promoter distances below the threshold distance. Simultaneously production of non-functional Tr and AT caused by premature transcription termination at roadblocks could be avoided by placing the roadblock binding sites outside the coding regions. However the local RNAP concentration effect seemed to saturate below the threshold distance as we could not observe a significant increase in reporter signal for most of the designs.
In addition we observed up to twofold higher maximum signals for promoter ar-

rangements, in which trigger transcription is regulated by the actual weaker pLac promoter, while anti-trigger transcription is controlled by the stronger pTet promoter compared to the opposite promoter arrangement with exchanged pLac and pTet promoters. The best designs were assembled to create a two-input two-output logic gate, which shows fast switching kinetics.

TI effects for three active promoters   However, in presence of a third active promoter regulating AT transcription, TI effects measured on the transcription (mRNA) level were mask by the system readout preventing a more detailed study of the system. As our readout depends on the concentration of active TS, the levels of the expressed RNAs (Tr, AT, TS) are not accessible directly. Therefore simultaneously the readout mechanism has to be simplified while keeping the promoter distances (in the tandem design) short. Expressing aptamers (e.g. (Baby) Spinach [160][161] and Malachite green aptamer [162]) instead of Tr and AT would meet both criteria. In such a design also TS has to be deleted to preserve the fluorescent readout of J23111. Based on the results one might be able to implement a model and simulate the TI effects also for the more complicated TS circuits taking advantage of transfer functions for Tr and AT measured with purified RNAs in cell extract during my Masters thesis [144].

Alternative designs and XOR gate   Whereas we mainly observed local RNAP concentration effects, also other TI effects arise especially for the convergent design and can change the system performance significantly. E.g. collision effects dominate for greater promoter distances and will have asymmetric effects on both promoters as it was reported by Bordoy et al. [128]. Alternative convergent designs can be constructed with a second Tr and AT between the face-to-face promoters or spacer sequences can be placed strategically. If we assume that transcription is terminated close to the middle region, it makes a great difference whether the spacer is placed upstream or downstream the trigger region in particular because the antisense transcript will show automatically the opposite placing. Some design ideas presented in Figure 5.2 might show improved repression characteristics compared to our tested designs and might enable the construction of a XOR gate with a single reporter as discussed in section 4.1.5.
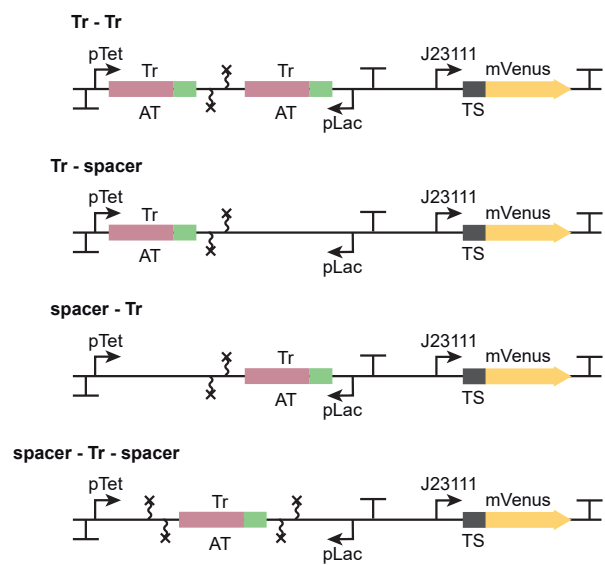
Figure 5.2: Alternative convergent designs with elongated promoter distances. Designs with two triggers and anti-triggers or one Tr and AT in combination with a spacer sequence are conceivable. The spacer sequence can be placed upstream or downstream of Tr or can be splitted. The antisense transcripts shows automatically the opposite placement for a decentralized position of Tr. Versions with exchanged pLac and pTet promoters might also show different results.

Appendix A

Supplementary figures and information for chapter 3

A.1 Supplementary figures
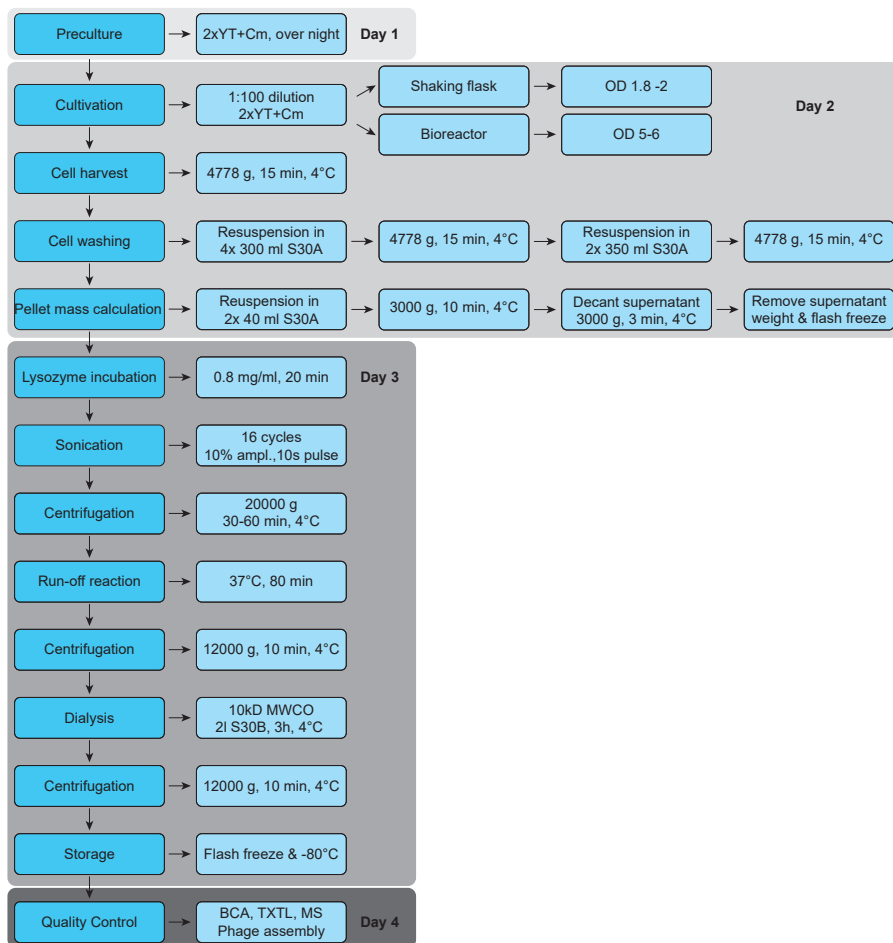
Figure A.1: Detailed flow chart of the cell extract preparation protocol. See methods part in main text for more information.
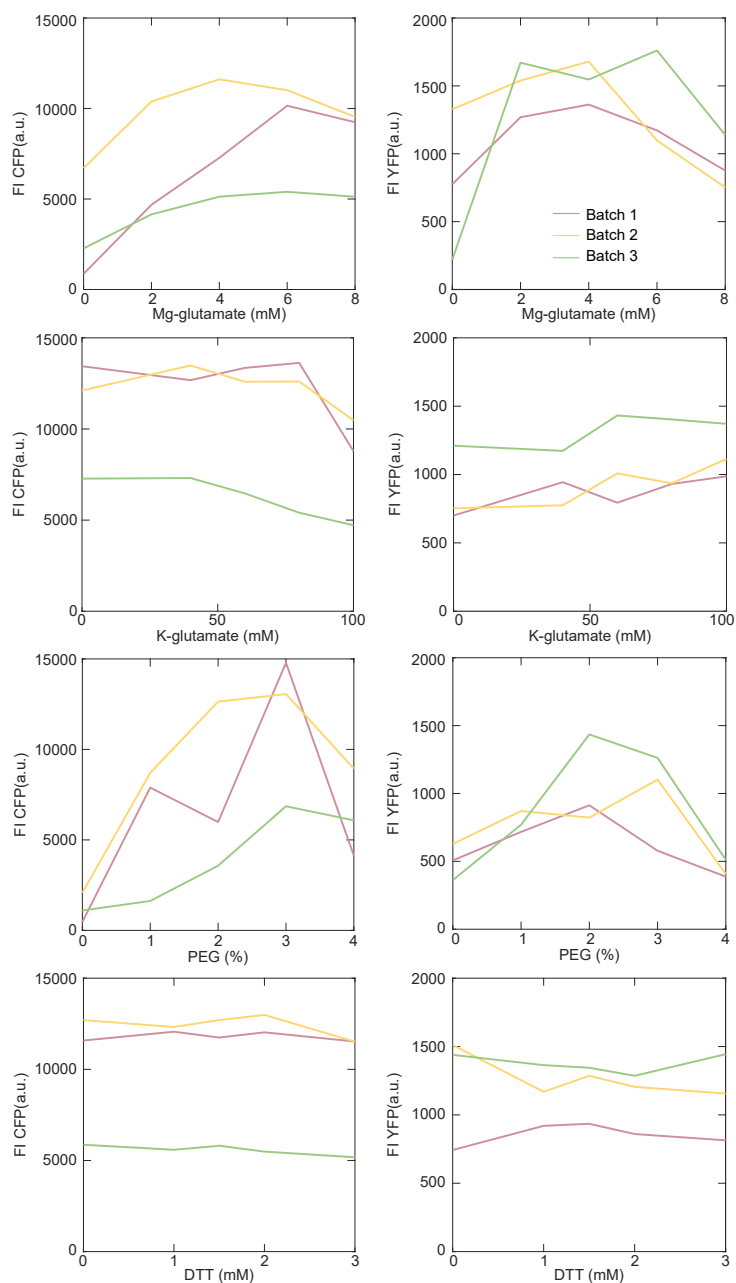
Figure A.2: Cell extract buffer screening experiments. Mg-glutamate, K-glutamate, PEG and DTT concentration was screened for 3 different cell extract batches prepared using the S16/L0.8 lysis setting. The fluorescence end levels for 2 different reporter plasmids (CFP and YFP) are shown. 4 mM Mg-glutamate, 60 mm K-glutamate, 2.5% PEG and 0 mM DTT were chosen.
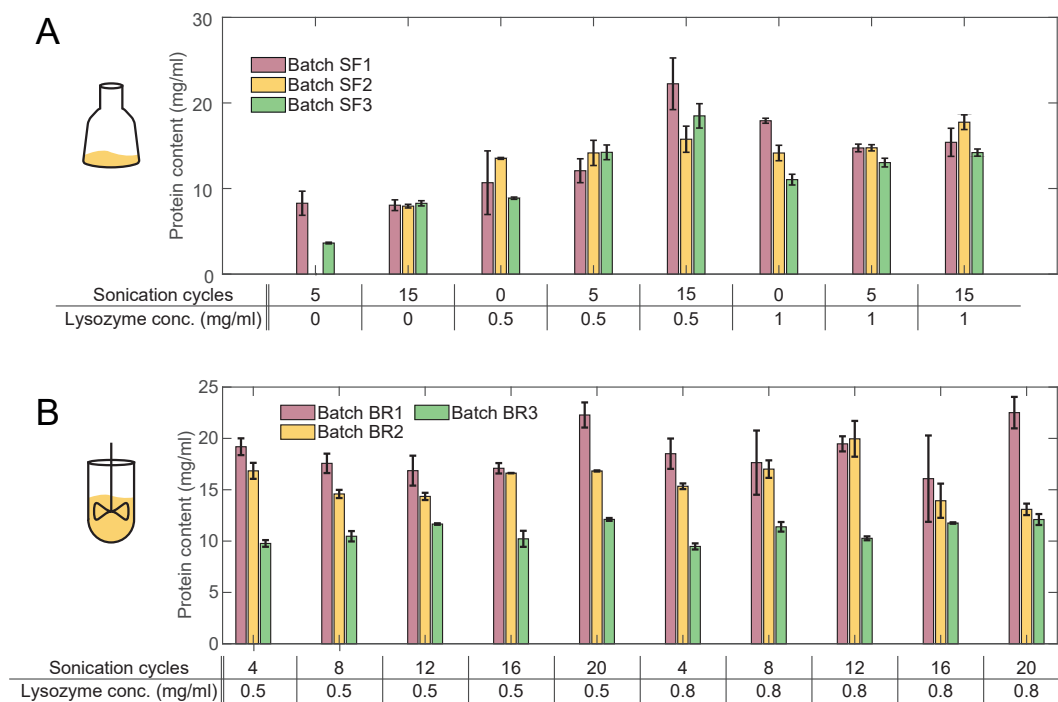
Figure A.3: (A) For the shaking flask replicates lysozyme incubation had a higher impact on cell lysis than sonication cycles. The protein content was generally higher in samples without sonication cycles (S0/L0.5 and S0/L1) than in samples not incubated with lysozyme (S5/L0 and S15/L0). Samples which were treated with a combination of lysozyme incubation and sonication cycles showed a high protein content. (B) In contrast to the shaking flask replicates, the bioreactor replicates showed a higher deviation in protein content among the biological replicates than among the different lysis settings.
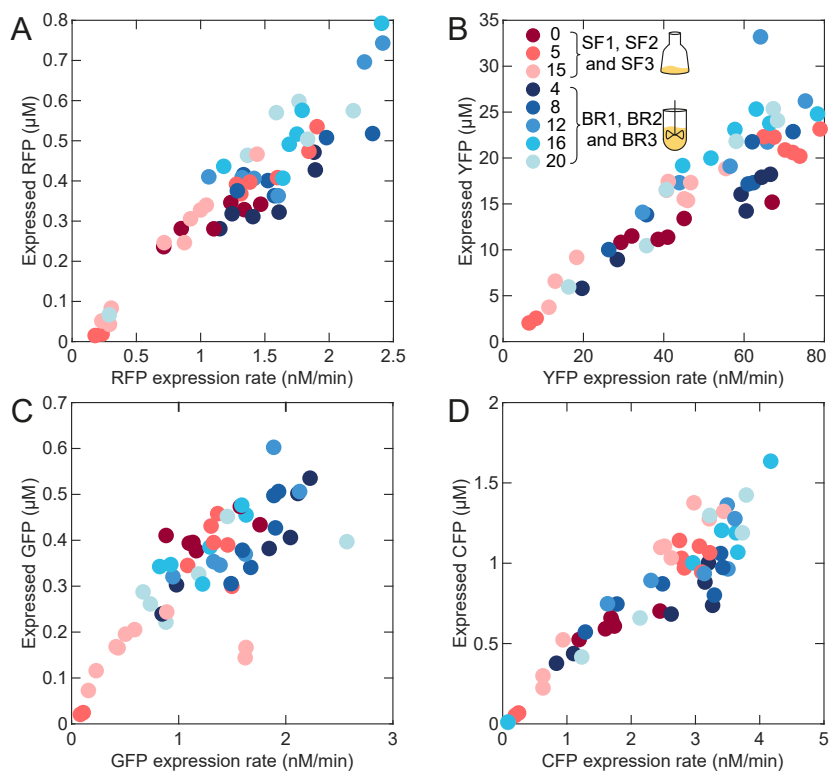
Figure A.4: Expressed protein end level against maximum protein expression rates. The maximum expression rates correlate well with the expressed protein end levels.
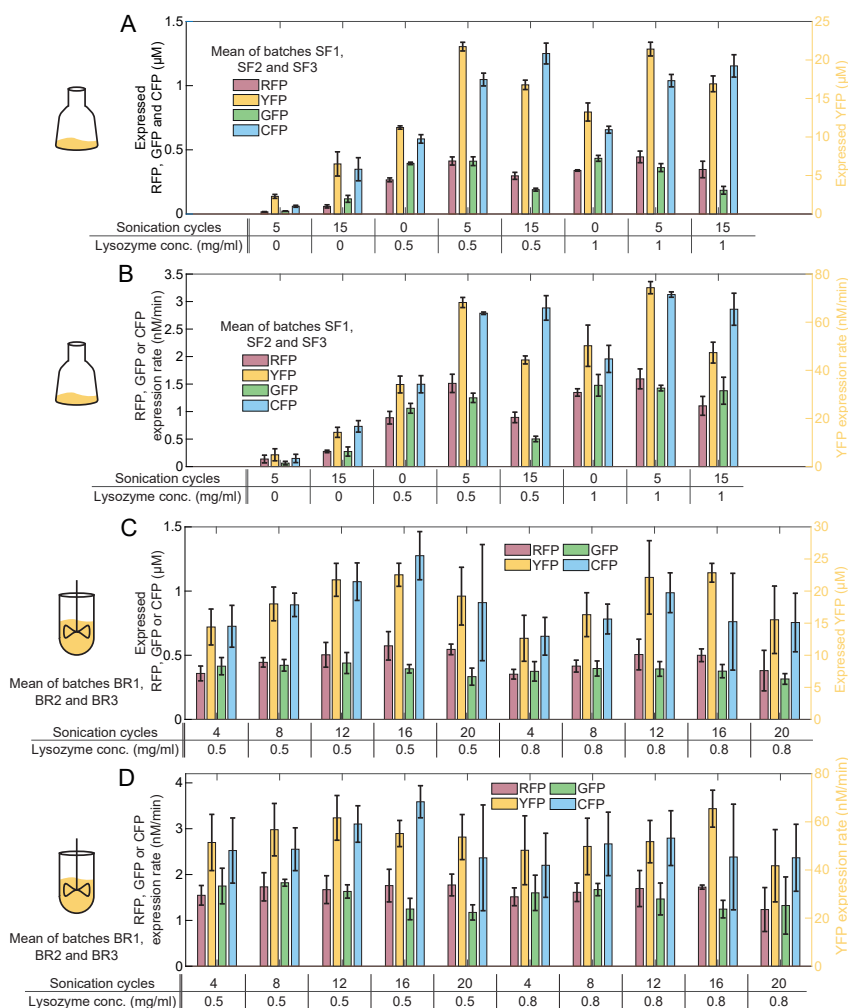
Figure A.5: End levels of expressed proteins and maximum protein expression rates. (A) Shaking flask samples, which were not treated with lysozyme but with 5 or 15 sonication cycles, show the lowest fluorescence end levels in a cell-free test for all tested reporter proteins. Samples which were lysed without sonication cycles show higher signals, so lysozyme has no negative effect on the protein synthesis. The fluorescence intensities of samples with nonzero lysis conditions are the highest, a signal decrease can be observed for 15 sonication cycles. (B) The maximum protein expression rates show the same trends. (C) For both lysozyme concentrations an increase of sonication cycles results in an increased fluorescence signal in the bioreactor samples. Independent of the lysozyme concentration, 12-16 sonication cycles appear optimal, while 20 cycles result in reduced signals. (D) Mean of bioreactor replicates BR1, BR2 and BR3. The maximum expression rate correlates with the expressed protein end levels.
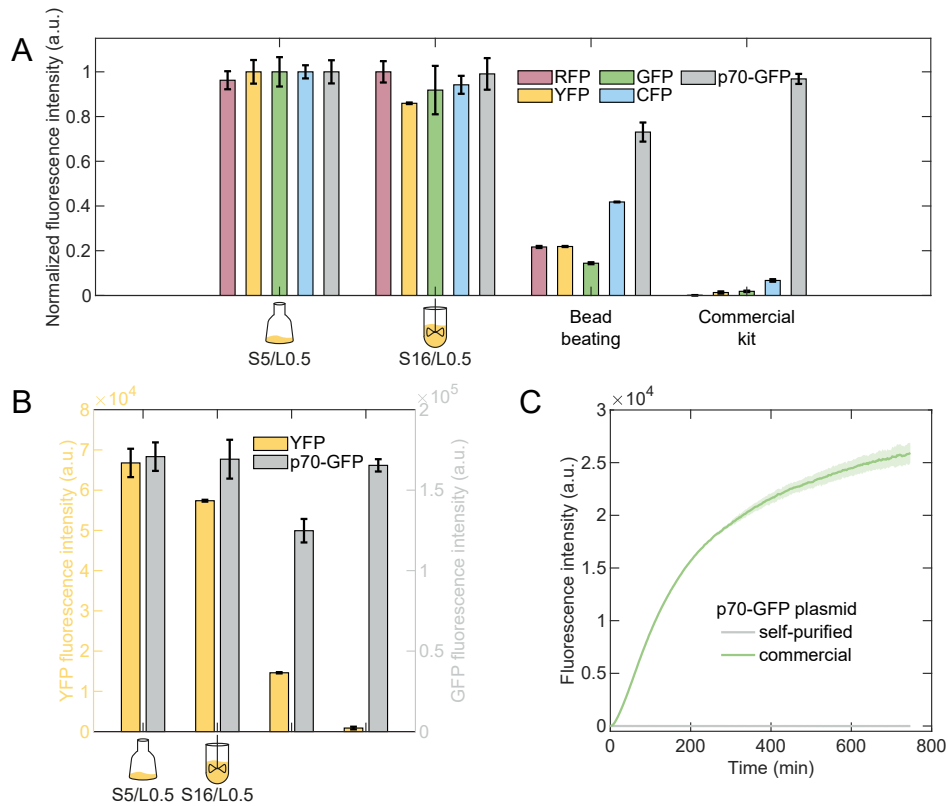
Figure A.6: Comparison of the best shaking flask replicates with the best bioreactor replicates, a bead beating batch and a commercial cell extract. (A) Normalized fluorescence data for all tested reporters. The commercial extract shows low levels for all reporter plasmids except the p70 control plasmid. (B) Not normalized fluorescence data for the YFP plasmid and the GFP control plasmid. The levels are very different, but these levels depend not just on protein concentration but also on quantum yield, brightness and other parameters. (C) p70-GFP control plasmid in commercial cell extract. We purified the p70-GFP control plasmid using our standard technique and performed a TXTL test. The commercial plasmid shows a high signal whereas our self-purified one shows almost no signal indicating sensitivity to our purification method.
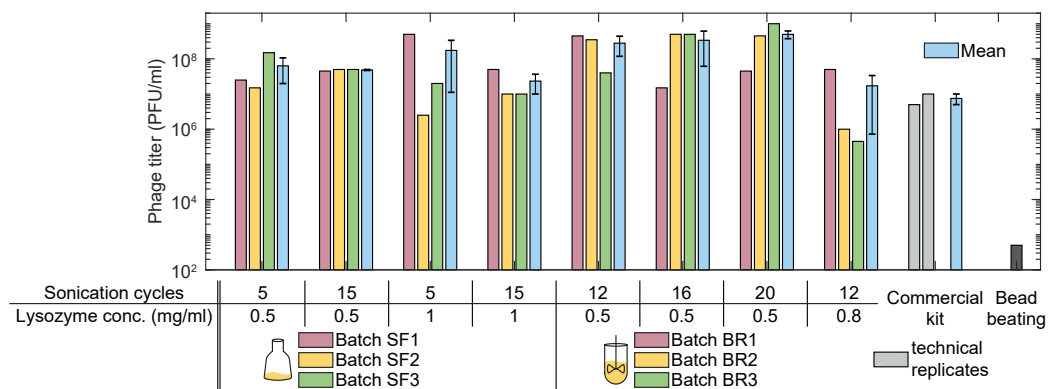
Figure A.7: Plaque assays for the shaking flask and bioreactor replicates. For all self-made cell extracts except for the S12/L0.8 samples comparable phage titers were measured. For the commercial kit one order of magnitude less phages could be assembled and the bead beating batch performed much worse, just 500 phages could be counted in one replicate, no plaques could be counted for the second replicate.
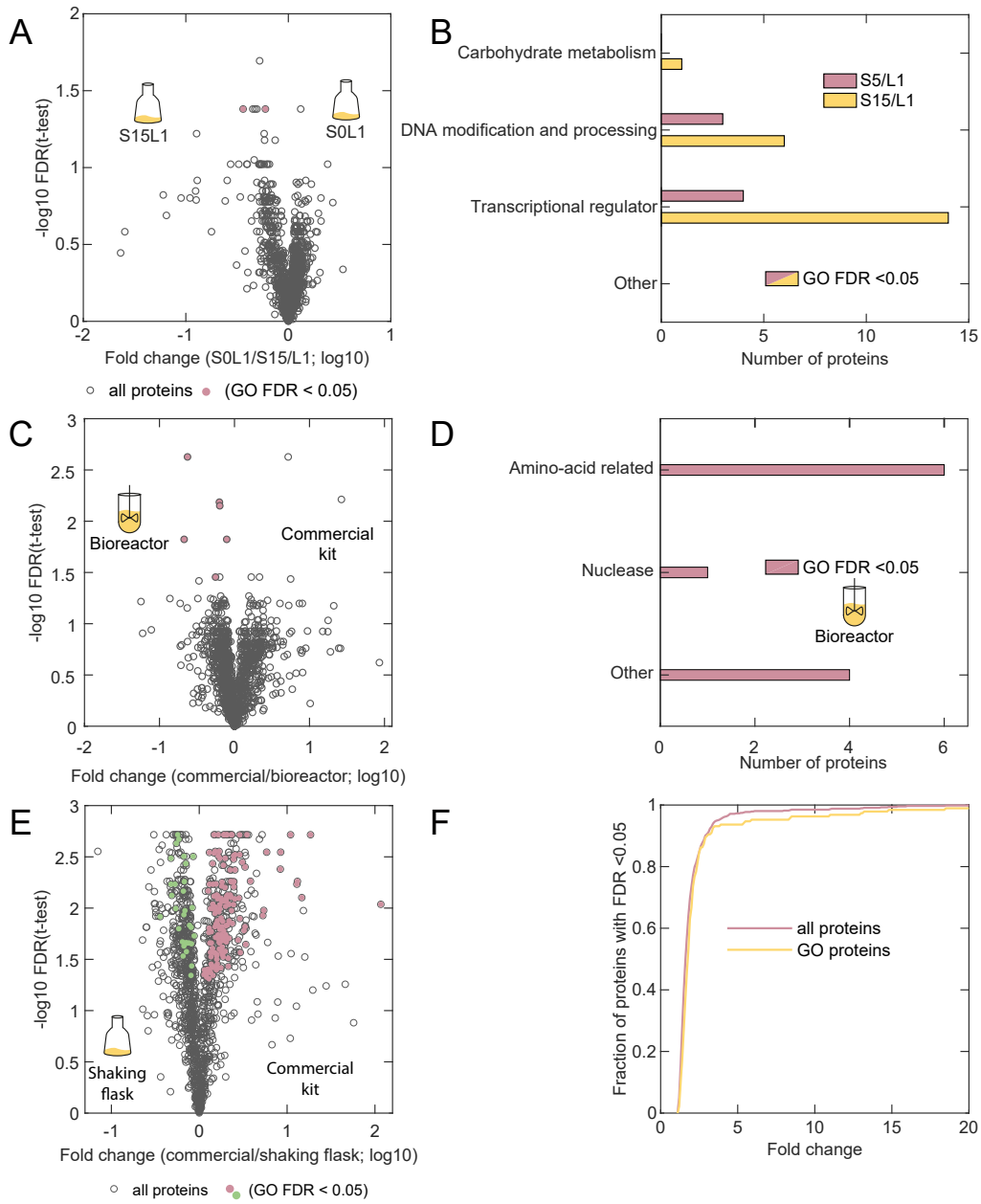
Figure A.8: Comparison of the proteomes of extracts prepared from bioreactor and shaking flask cultures and a commercial kit. (A) Volcano plot of shaking flask samples S0/L1 against S15/L1. The t-test of S0/L1 against S15/L1 extracts gives 1 and 6 proteins with FDR < 0.05 (i.e., –log10(FDR) > 1.3). In addition, 3 (S0/L1) and 47 (S15/L1) unique proteins could be found and were also subjected to an enrichment analysis. (B) Enrichment analysis derived from the comparison of S0/L1 against S0/L5 and from the comparison shown in (A). No GO terms with FDR < 0.05 could be found for the S0/L1 extracts in both comparisons but in total 7 (S0/L1 against S5/L1) and 21 (S0/L1 against S15/L1) proteins were found in GO terms with FDR < 0.05 for the S5/L1 and S15/L1 samples respectively. For both cases these could be assigned to keywords related to DNA replication, relaxation, repair or recombination or were transcriptional regulators. (C) Volcano plot of bioreactor samples S5/L0.5 against the commercial kit. The commercial and bioreactor extracts differ in abundance of just 8 and 5 proteins with FDR < 0.05 (i.e., –log10(FDR) > 1.3), but had 30 and 121 unique proteins, which were subjected to an enrichment analysis. (D) Enrichment analysis derived from the comparison shown in (C). GO terms with FDR < 0.05 included in total 11 proteins in the bioreactor samples, which were related in amino-acid biosynthesis, a nuclease or not relevant for cell-free gene expression. (E) Volcano plot of shaking flask samples S5/L0.5 against the commercial kit. The commercial and shaking flask extracts differ in abundance of 309 versus 356 proteins with FDR < 0.05 (i.e., –log10(FDR) > 1.3). Proteins with an FDR below 0.05 and proteins exclusively found only in shaking flask preparations or in commercial extract were subjected to an enrichment analysis. Proteins which were assigned to GO terms with an FDR below 0.05 are highlighted in the plot. (F) Fraction of proteins with FDR < 0.05 against fold change and fraction of proteins found in GO terms with FDR < 0.05 against fold change (unique proteins are not considered, as no fold change can be calculated). 27% of the proteins are less than 1.5 fold enriched, 67% less than 2 fold and 90% less than 3 fold.

A.2 Supplementary tables

2-Chloroacetamide (CAA)	Merck (GER)
Acetonitrile (ACN)	Merck (GER)
CTP	Carl Roth (GER)
GamS	Arbor Bioscience (US)
Glycerol	Carl Roth (GER)
GTP	Carl Roth (GER)
IPTG	Carl Roth (GER)
Nuclease-free water	Carl Roth (GER)
Phosphate buffered saline (PBS)	VWR Life Science (GER)
RTS Amino Acid Sampler	Biozym Scientific (GER)
Tris	Carl Roth (GER)
Trypsin	Roche (CH)
UTP	Carl Roth (GER)

Table A.1: List of chemicals which were not ordered from Sigma Aldrich.

All buffer and media compositions were adapted from Sun et al. [59]. Anyways the composition of the growth medium and the buffers needed for cell washing and cell extract dialysis are listed below. A detailed protocol for the TXTL buffer preparation is shown in the protocol of Sun et al. [59].

2xYT	31 g/l
K_2HPO_4	40 mM
KH_2PO_4	22 mM

Table A.2: 2xYTP medium

Potassium glutamate	60 mM
Magnesium glutamate	14 mM
Tris	50 mM

Table A.3: S30A (cell washing, 2l are needed). To reach pH 7.7, titrate with acetic acid. Add DTT to 2mM final concentration just before use. Store at 4 °C.

Potassium glutamate	60 mM
Magnesium glutamate	14 mM
Tris	5 mM

Table A.4: S30B (cell extract dialysis, 2l are needed). To reach pH 8.2, titrate with 2 M Tris. Add DTT to 1 mM final concentration just before use. Store at 4 °C.

Protein	mScarlet I	mVenus	GFP mut3	mTurquoise 2	p70-GFP (deGFP3)
Ex. Max (nm)	569	515	500	434	508
Em. Max (nm)	593	527	513	474	518
QY	0.54	0.64	0.39	0.93	0.19
Brightness	56.16	66.56	34.87	27.9	5.6
maturation time (min)	36	17.6	4.1	33.5	
pKa	5.4	5.5		3.1	6.9
additional mutations	none	2V--> QK; 70Q-->M; 81Q-->R; 232L-->H	2S-->R	147I-->F	2SKGE-->---; 64F-->L; 203C-->T
CAI E. coli B	0.636	0.878	0.588	0.782	0.646
CAI Shigella flex. 2a	0.765	0.934	0.733	0.853	0.791
GC content full mRNA	49.31	47.6	40.9	50.11	59.13
GC content CDS	49.5	47.3	39.05	50.56	61.06
Predicted translation rate	3052	905	746	357	9373
ΔG total	-2.01	0.69	1.12	2.76	-4.51
ΔG mRNA-rRNA	-7.53	-7.53	-9.42	-10.08	-10.1
ΔG spacing	0.67	0.67	0.67	0.67	0.29
ΔG stacking	0	0	0	0	0
ΔG standby	4.9	4.9	4.9	4.9	0.01
ΔG start	-2.76	-2.76	-2.76	-2.76	-2.76
ΔG mRNA	-2.79	-5.49	-7.81	-10.11	-8.33

Table A.5: Characteristics of reporter proteins mScarlet I [163], mVenus [164], GFP mut3 [165], mTurquoise2 [166] and p70-GFP (deGFP3) [167]. Codon adaption indices for mScarlet, mVenus, GFP and mTurquoise calculated by CAIcal [168]. The codon usage table of *Escherichia coli B* and of *Shigella flex. 2a* were used for the calculation. The second one is more related to MRE600, which is the origin strain of the purified tRNA used for the TXTL buffer. The translation rate was predicted using an RBS calculator [139].

Culture method	Lysozyme (mg/ml)	Data set	p-value
	0.5	1	$1.32 \cdot 10^{-5}$
Shaking flask	1	2	$1.20 \cdot 10^{-3}$
	0.5 and 1 combined	3	$1.60 \cdot 10^{-8}$
	0.5	4	$6.43 \cdot 10^{-2}$
Bioreactor	0.8	5	$4.46 \cdot 10^{-2}$
	0.5 and 0.8 combined	6	$8.30 \cdot 10^{-3}$

Table A.6: : T-test for TXTL data. We aimed to proof our hypothesis, that our data show an optimum in the number of sonication cycles. The data shown in Figure 2 B and D were split in two sets with fixed lysozyme concentration and different sonication cycles: For shaking flask cell extracts these were 0, 5, and 15 sonication cycles in combination with a lysozyme concentration of 0.5 or 1 mg/ml respectively (data set 1 and 2; samples S5/L0 and S15/L0 were excluded from the t-test). For the bioreactor extracts these were 4, 8, 12, 16, and 20 sonication cycles in combination with a lysozyme concentration of 0.5 or 0.8 mg/ml respectively (data set 4 and 5). As we observe the same trend in the data independent on the lysozyme concentration, we also introduced combined data sets (data set 3 and 6). A parabola $y=ax^2+bx+c$ with x being the number of sonication cycles and y being the TXTL end level of the YFP reporter was fitted to the data sets. The fit parameter a was tested against the hypothesis 'Fit parameter a is zero' and the according p-value was calculated. We could reject the null hypothesis for all data sets (p-value < 0.05) except for data set 4, which has a p-value of 0.06. So the observed optima in the TXTL data are statistically significant.

A.3 Enriched proteins and summary of protein numbers

Uniprot ID	Fold change	-log10(FDR)	Protein name	Gene name	Keyword
A0A140NFX5_ECOBD	2,715709578	2,715709578	D-aminoacyl-tRNA deacylase	**dtd** ECBD_4140, HO396_19315	amino-acid related
A0A140N9P3_ECOBD	2,254748493	2,254748493	NADH-quinone oxidoreductase	**nuoG** ECBD_1378, HO396_11215	ATP synthesis related
A0A140N3Q3_ECOBD	NaN	NaN	L-lactate dehydrogenase	**lldD** ECBD_0120, HO396_17700	carbohydrate metabolism and respiration
A0A140NA88_ECOBD	NaN	NaN	Cytochrome bo(3) ubiquinol oxidase ...	**cyoB** ECBD_3227, HO396_01970	carbohydrate metabolism and respiration
A0A140NAN1_ECOBD	1,535175058	1,535175058	Quinone-dependent D-lactate dehydro...	**dld** ECBD_1525, HO396_10490	carbohydrate metabolism and respiration
A0A140NCG6_ECOBD	2,400717199	2,400717199	Cytochrome bd ubiquinol oxidase sub...	**cydA** ECBD_2928, HO396_03500	carbohydrate metabolism and respiration
A0A140NCT2_ECOBD	2,127606292	2,127606292	Fumarate reductase flavoprotein sub...	**frdA** ECBD_3875, HO396_20675	carbohydrate metabolism and respiration
A0A140NDR8_ECOBD	2,715709578	2,715709578	Ubiquinone/menaquinone biosynthesis...	**ubiE** ECBD_4190, HO396_19055	carbohydrate metabolism and respiration
A0A140N599_ECOBD	1,769795266	1,769795266	DNA topoisomerase 4 subunit B	**parE** ECBD_0709, HO396_14780	DNA modification and replication
A0A140N626_ECOBD	2,372032264	2,372032264	DNA gyrase subunit A	**gyrA** ECBD_1429, HO396_10965	DNA modification and replication
A0A140N6C1_ECOBD	1,368085361	1,368085361	DNA gyrase subunit B	**gyrB** ECBD_0004, HO396_18280	DNA modification and replication
A0A140NB69_ECOBD	2,379969605	2,379969605	Integration host factor subunit alp...	**ihfA** himA, ECBD_1933, HO396_08640	DNA modification and replication
A0A140NBC7_ECOBD	1,436044176	1,436044176	Chromosome partition protein MukE	**mukE** ECBD_2672, HO396_04835	DNA modification and replication
A0A140NBR4_ECOBD	2,116329269	2,116329269	Chromosome partition protein MukB	**mukB** ECBD_2671, HO396_04840	DNA modification and replication
A0A140NCD3_ECOBD	1,92703912	1,92703912	Recombination-associated protein Rd...	**rdgC** ECBD_3268	DNA modification and replication
A0A140NCN5_ECOBD	1,524964845	1,524964845	DNA polymerase III subunit gamma/ta...	**dnaX** ECBD_3186, HO396_02170	DNA modification and replication
A0A140NCX5_ECOBD	1,522980698	1,522980698	DNA topoisomerase 1	**topA** ECBD_2348, HO396_06510	DNA modification and replication
A0A140NDN9_ECOBD	1,818400105	1,818400105	DNA protection during starvation pr...	**dps** pexB, ECBD_2811, HO396_04120	DNA modification and replication
A0A140NDT9_ECOBD	1,452719489	1,452719489	Chromosome partition protein MukF	**mukF** ECBD_2673, HO396_04830	DNA modification and replication
A0A140NDV2_ECOBD	2,545724335	2,545724335	Integration host factor subunit bet...	**ihfB** himD, ECBD_2683, HO396_04780	DNA modification and replication
A0A140NHJ4_ECOBD	2,036802479	2,036802479	DNA helicase	**uvrD** mutU, recL, ECBD_4228, HO396_18870	DNA modification and replication
A0A140N210_ECOBD	1,973292106	1,973292106	SelB translation factor	**selB** ECBD_0140, HO396_17605	other
A0A140N500_ECOBD	1,682729453	1,682729453	Multidrug efflux transporter EmrAB ...	**emrR** ECBD_1036, HO396_12955	other
A0A140N5K0_ECOBD	1,739020298	1,739020298	2-octaprenyl-6-methoxyphenyl hydrox...	**ubiH** visB, ECBD_0830, HO396_14020	other
A0A140N763_ECOBD	1,995921991	1,995921991	RNA binding S1 domain protein	ECBD_0338	other
A0A140N7J9_ECOBD	2,262035298	2,262035298	Biotin carboxylase	**accC** ECBD_0489, HO396_15875	other
A0A140N7L1_ECOBD	1,51110941	1,51110941	Amidohydrolase	**abgA** ECBD_2279, HO396_06865	other
A0A140N843_ECOBD	1,432249637	1,432249637	FAD-dependent 2-octaprenylphenol hy...	**ubiI** visC, ECBD_0831, HO396_14015	other
A0A140N8E8_ECOBD	2,127606292	2,127606292	(P)ppGpp synthetase I, SpoT/RelA	ECBD_0945, HO396_13430	other
A0A140N8K8_ECOBD	1,719258427	1,719258427	Sua5/YciO/YrdC/YwlC family protein	ECBD_2356, HO396_06465	other
A0A140N8W1_ECOBD	NaN	NaN	Signal peptidase I	**lepB** ECBD_1113, HO396_12540	other
A0A140N8Y0_ECOBD	2,372032264	2,372032264	Fumarate hydratase class I	**fumB** ECBD_2034, HO396_08130	other
A0A140N919_ECOBD	NaN	NaN	Ancillary SecYEG translocon subunit...	**yfgM** ECBD_1173, HO396_12250	other
A0A140N9K0_ECOBD	NaN	NaN	Amidohydrolase	**abgB** ECBD_2280, HO396_06860	other
A0A140N9T6_ECOBD	2,715709578	2,715709578	Ubiquinone biosynthesis O-methyltra...	**ubiG** ECBD_1428, HO396_10970	other
A0A140N9Z7_ECOBD	2,714499052	2,714499052	NADH-quinone oxidoreductase subunit...	**nuoF** ECBD_1377, HO396_11220	other
A0A140NA70_ECOBD	NaN	NaN	Cytochrome d ubiquinol oxidase subu...	**cydB** ECBD_2927, HO396_03505	other
A0A140NBI8_ECOBD	1,395578749	1,395578749	Mischarged aminoacyl-tRNA deacylase	**yeaK** ECBD_1857, HO396_09015	translation, initiation, regulation, termination
A0A140NBK6_ECOBD	NaN	NaN	Protease 4	**sppA** ECBD_1878, HO396_08910	Protease
A0A140NC78_ECOBD	NaN	NaN	PTS N-acetyl glucosamine transporte...	**nagE** ECBD_2983, HO396_03200	other

ID	Value 1	Value 2	Protein name	Gene/locus	Category
A0A140NCR0_ECOBD	1,826979921	1,826979921	Ribonuclease R	**rnr** ECBD_3855, HO396_20795	Nuclease
A0A140NDI9_ECOBD	1,575341538	1,575341538	3-methyl-2-oxobutanoate hydroxymeth...	**panB** ECBD_3485	other
A0A140NEB0_ECOBD	NaN	NaN	2-octaprenyl-3-methyl-6-methoxy-1,4...	**ubiF** ECBD_2989, HO396_03135	other
A0A140NEV1_ECOBD	2,036802479	2,036802479	Probable cytosol aminopeptidase	**pepA** ECBD_3776, HO396_21195	other
A0A140NF24_ECOBD	2,715709578	2,715709578	DNA-binding protein HU-alpha	**hupA** ECBD_4032, HO396_19880	other
A0A140NFA9_ECOBD	2,715709578	2,715709578	DNA-binding protein HU-beta	**hupB** ECBD_3215, HO396_02020	other
A0A140NFX6_ECOBD	NaN	NaN	Membrane protein insertase YidC	**yidC** ECBD_4327, HO396_18320	other
A0A140NHC7_ECOBD	1,399920367	1,399920367	GTP-binding protein TypA	**typA** ECBD_4156, HO396_19235	other
A0A140NHQ8_ECOBD	2,545724335	2,545724335	ATP-dependent protease subunit HslV	**hslV** ECBD_4092, HO396_19540	Protease
A0A140SS21_ECOBD	1,67112025	1,67112025	23S rRNA (guanosine-2'-O-)-methyltr...	**rlmB** ECBD_3854, HO396_20800	other
A0A140SSA6_ECOBD	1,975462238	1,975462238	Xaa-Pro dipeptidase	**pepQ** ECBD_4178, HO396_19115	Protease
TRMA_ECOBD	1,876494138	1,876494138	tRNA/tmRNA (uracil-C(5))-methyltran...	**trmA** ECBD_4059, ECD_03850, B21_03799	translation: initiation, regulation, termination
A0A140NEN6_ECOBD	1,54674934	1,54674934	10 kDa chaperonin	**groS** groES, ECBD_3889, HO396_20605	protein folding or unfolding
A0A140NF86_ECOBD	2,520931426	2,520931426	ATP-dependent protease ATPase subun...	**hslU** ECBD_4093, HO396_19535	protein folding or unfolding
A0A140N2S3_ECOBD	1,432249637	1,432249637	50S ribosomal protein L22	**rplV** ECBD_0436, HO396_16175	Ribosomal protein
A0A140N2T1_ECOBD	1,777900735	1,777900735	50S ribosomal protein L6	**rplF** ECBD_0446, HO396_16125	Ribosomal protein
A0A140N2Z9_ECOBD	1,739020298	1,739020298	30S ribosomal protein S9	**rpsI** ECBD_0517, HO396_15755	Ribosomal protein
A0A140N340_ECOBD	1,370183834	1,370183834	50S ribosomal protein L27	**rpmA** ECBD_0557, HO396_15555	Ribosomal protein
A0A140N3G7_ECOBD	1,706209039	1,706209039	50S ribosomal protein L3	**rplC** ECBD_0431, HO396_16200	Ribosomal protein
A0A140N3H4_ECOBD	1,721845921	1,721845921	50S ribosomal protein L14	**rplN** ECBD_0441, HO396_16150	Ribosomal protein
A0A140N3L9_ECOBD	1,871209857	1,871209857	50S ribosomal protein L28	**rpmB** ECBD_0089, HO396_17855	Ribosomal protein
A0A140N4K1_ECOBD	1,818400105	1,818400105	30S ribosomal protein S3	**rpsC** ECBD_0437, HO396_16170	Ribosomal protein
A0A140N4M0_ECOBD	2,511036592	2,511036592	50S ribosomal protein L17	**rplQ** ECBD_0457, HO396_16070	Ribosomal protein
A0A140N528_ECOBD	1,453504867	1,453504867	30S ribosomal protein S19	**rpsS** ECBD_0435, HO396_16180	Ribosomal protein
A0A140N537_ECOBD	1,874352239	1,874352239	30S ribosomal protein S8	**rpsH** ECBD_0445, HO396_16130	Ribosomal protein
A0A140N548_ECOBD	2,252064104	2,252064104	30S ribosomal protein S4	**rpsD** ECBD_0455, HO396_16080	Ribosomal protein
A0A140N5A3_ECOBD	2,09019138	2,09019138	50S ribosomal protein L5	**rplE** ECBD_0443, HO396_16140	Ribosomal protein
A0A140N5B4_ECOBD	2,116329269	2,116329269	30S ribosomal protein S13	**rpsM** ECBD_0453, HO396_16090	Ribosomal protein
A0A140N5D7_ECOBD	2,106162063	2,106162063	50S ribosomal protein L21	**rplU** ECBD_0556, HO396_15560	Ribosomal protein
A0A140N5K8_ECOBD	1,577896025	1,577896025	50S ribosomal protein L4	**rplD** ECBD_0432, HO396_16195	Ribosomal protein
A0A140N5L7_ECOBD	1,564050884	1,564050884	50S ribosomal protein L24	**rplX** ECBD_0442, HO396_16145	Ribosomal protein
A0A140N6T7_ECOBD	2,091658201	2,091658201	50S ribosomal protein L19	**rplS** ECBD_1080, HO396_12730	Ribosomal protein
A0A140N6W8_ECOBD	2,508907286	2,508907286	30S ribosomal protein S7	**rpsG** ECBD_0408, HO396_16305	Ribosomal protein
A0A140N6Y5_ECOBD	1,993382971	1,993382971	30S ribosomal protein S10	**rpsJ** ECBD_0430, HO396_16205	Ribosomal protein
A0A140N6Z2_ECOBD	2,114660618	2,114660618	50S ribosomal protein L16	**rplP** ECBD_0438, HO396_16165	Ribosomal protein
A0A140N6Z9_ECOBD	1,719258427	1,719258427	30S ribosomal protein S5	**rpsE** ECBD_0448, HO396_16115	Ribosomal protein
A0A140N711_ECOBD	2,033564102	2,033564102	50S ribosomal protein L15	**rplO** ECBD_0450, HO396_16105	Ribosomal protein
A0A140N7D8_ECOBD	2,127606292	2,127606292	30S ribosomal protein S16	**rpsP** ECBD_1077, HO396_12745	Ribosomal protein
A0A140N7G4_ECOBD	1,550848276	1,550848276	50S ribosomal protein L30	**rpmD** ECBD_0449, HO396_16110	Ribosomal protein
A0A140N7J1_ECOBD	1,673120024	1,673120024	50S ribosomal protein L2	**rplB** ECBD_0434, HO396_16185	Ribosomal protein
A0A140N7K8_ECOBD	1,714325265	1,714325265	30S ribosomal protein S14	**rpsN** ECBD_0444, HO396_16135	Ribosomal protein
A0A140N7L9_ECOBD	1,739020298	1,739020298	30S ribosomal protein S11	**rpsK** ECBD_0454, HO396_16085	Ribosomal protein
A0A140N811_ECOBD	1,555826867	1,555826867	30S ribosomal protein S15	**rpsO** ECBD_0575, HO396_15455	Ribosomal protein

A0A140N846_ECOBD	2,214717979	2,214717979	50S ribosomal protein L25	**rplY** ECBD_1472, HO396_10755	Ribosomal protein
A0A140N8B4_ECOBD	2,017062565	2,017062565	30S ribosomal protein S21	**rpsU** ECBD_0676, HO396_14955	Ribosomal protein
A0A140NBA5_ECOBD	1,818400105	1,818400105	30S ribosomal protein S1	**rpsA** ECBD_2684, HO396_04775	Ribosomal protein
A0A140NCE1_ECOBD	2,520931426	2,520931426	30S ribosomal protein S22	**sra** ECBD_2159, HO396_07470	Ribosomal protein
A0A140NDV1_ECOBD	2,201228736	2,201228736	50S ribosomal protein L9	**rplI** ECBD_3831, HO396_20915	Ribosomal protein
A0A140NF32_ECOBD	1,61938993	1,61938993	50S ribosomal protein L11	**rplK** ECBD_4050, HO396_19790	Ribosomal protein
A0A140NFK2_ECOBD	2,068401615	2,068401615	30S ribosomal protein S2	**rpsB** ECBD_3450, HO396_00845	Ribosomal protein
A0A140NFU3_ECOBD	2,092890112	2,092890112	30S ribosomal protein S20	**rpsT** ECBD_3593, HO396_00120	Ribosomal protein
A0A140NGG7_ECOBD	1,735318381	1,735318381	30S ribosomal protein S6	**rpsF** ECBD_3834, HO396_20900	Ribosomal protein
A0A140NGH1_ECOBD	1,975462238	1,975462238	30S ribosomal protein S18	**rpsR** ECBD_3832, HO396_20910	Ribosomal protein
A0A140NHV0_ECOBD	1,897705914	1,897705914	50S ribosomal protein L34	**rpmH** ECBD_4330, HO396_18305	Ribosomal protein
A0A140SS63_ECOBD	1,706209039	1,706209039	50S ribosomal protein L7/L12	**rplL** ECBD_4047, HO396_19805	Ribosomal protein
A0A140SS71_ECOBD	2,206783483	2,206783483	50S ribosomal protein L31	**rpmE** ECBD_4088, HO396_19560	Ribosomal protein
A0A140N4C9_ECOBD	2,167111388	2,167111388	Ribosomal RNA small subunit methylt...	ECBD_0794	Ribosome associated
A0A140N5Y7_ECOBD	1,311643666	1,311643666	Ribosome-binding factor A	**rbfA** ECBD_0573, HO396_15465	Ribosome associated
A0A140N989_ECOBD	1,844365916	1,844365916	ATP-dependent RNA helicase SrmB	**srmB** ECBD_1104, HO396_12585	Ribosome associated
A0A140NBQ1_ECOBD	2,038435631	2,038435631	ATP-dependent RNA helicase RhlE	**rhlE** ECBD_2826, HO396_04045	Ribosome associated
A0A140ND33_ECOBD	1,403819411	1,403819411	Ribosomal RNA large subunit methylt...	**rlmI** ECBD_2627	Ribosome associated
A0A140ND50_ECOBD	2,23285838	2,23285838	Ribosomal RNA large subunit methylt...	**rlmL** rlmKL, ECBD_2647, HO396_04960	Ribosome associated
A0A140NDB6_ECOBD	1,83519948	1,83519948	50S ribosomal protein L10	**rplJ** ECBD_4048, HO396_19800	Ribosome associated
A0A140NEN1_ECOBD	2,233174935	2,233174935	Ribosome-associated protein YbcJ	**ybcJ** ECBD_3130, HO396_02460	Ribosome associated
A0A140NHX7_ECOBD	1,849800897	1,849800897	Der GTPase-activating protein YihI	**yihI** ECBD_4162, HO396_19205	Ribosome associated
A0A140N2L7_ECOBD	1,804072352	1,804072352	Ribonuclease PH	**rph** ECBD_0083, HO396_17885	RNA modification and processing
A0A140N5F4_ECOBD	2,260642936	2,260642936	Polyribonucleotide nucleotidyltrans...	**pnp** ECBD_0576, HO396_15450	RNA modification and processing
A0A140N5P0_ECOBD	1,567846121	1,567846121	Ribosomal RNA small subunit methylt...	**rsmI** ECBD_0594, HO396_15355	RNA modification and processing
A0A140N6N0_ECOBD	2,254491907	2,254491907	Phenylalanine--tRNA ligase alpha su...	**pheS** ECBD_1931, HO396_08650	RNA modification and processing
A0A140N6U7_ECOBD	1,607545562	1,607545562	Pseudouridine synthase	**rluD** ECBD_1090, HO396_12675	RNA modification and processing
A0A140N719_ECOBD	1,903366827	1,903366827	Ribosome maturation factor RimM	**rimM** ECBD_1078, HO396_12740	RNA modification and processing
A0A140N7K6_ECOBD	2,486257014	2,486257014	Ribonuclease G	**rng** ECBD_0500, HO396_15835	RNA modification and processing
A0A140N7Y4_ECOBD	1,659661707	1,659661707	Pseudouridine synthase	**rsuA** ECBD_1474, HO396_10745	RNA modification and processing
A0A140N8N4_ECOBD	2,715709578	2,715709578	Phenylalanine--tRNA ligase beta sub...	**pheT** ECBD_1932, HO396_08645	RNA modification and processing
A0A140N998_ECOBD	2,068401615	2,068401615	Ribonuclease 3	**rnc** ECBD_1114, HO396_12535	RNA modification and processing
A0A140N9A7_ECOBD	1,416878204	1,416878204	Translational regulator CsrA	**csrA** ECBD_1029, HO396_13015	RNA modification and processing
A0A140N9H5_ECOBD	1,688346029	1,688346029	tRNA/rRNA methyltransferase	ECBD_1099, HO396_12610	RNA modification and processing
A0A140N9P4_ECOBD	2,553374934	2,553374934	Dual-specificity RNA methyltransfer...	**rlmN** trmG, ECBD_1169, HO396_12270	RNA modification and processing
A0A140N9R6_ECOBD	1,395578749	1,395578749	Exoribonuclease 2	**rnb** ECBD_2331, HO396_06595	RNA modification and processing
A0A140NAA1_ECOBD	2,447254088	2,447254088	Ribonuclease E	**rne** ECBD_2516, HO396_05625	RNA modification and processing
A0A140NB54_ECOBD	1,994193494	1,994193494	RNA chaperone ProQ	**proQ** ECBD_1809, HO396_09260	RNA modification and processing
A0A140NBF5_ECOBD	1,419769861	1,419769861	Serine--tRNA ligase	**serS** ECBD_2702, HO396_04685	RNA modification and processing
A0A140NCL7_ECOBD	2,715709578	2,715709578	Pseudouridine synthase	**rluB** ECBD_2353, HO396_06485	RNA modification and processing
A0A140NFF4_ECOBD	2,715709578	2,715709578	Queuine tRNA-ribosyltransferase	**tgt** ECBD_3255, HO396_01830	RNA modification and processing
A0A140NG24_ECOBD	1,659661707	1,659661707	Poly(A) polymerase I	**pcnB** ECBD_3476	RNA modification and processing
A0A140SSA5_ECOBD	1,977286864	1,977286864	ATP-dependent RNA helicase RhlB	**rhlB** ECBD_4260, HO396_18695	RNA modification and processing

A0A140N4Y7_ECOBD	2,235066468	2,235066468	ATP-dependent RNA helicase DeaD	**deaD** csdA, ECBD_0578, HO396_15435	SOS or stress response
A0A140NF66_ECOBD	2,097195799	2,097195799	Catalase-peroxidase	**katG** ECBD_4081, HO396_19595	SOS or stress response
A0A140NHF7_ECOBD	1,396921716	1,396921716	LexA repressor	**lexA** ECBD_3990, HO396_20100	SOS or stress response
A0A140N2U0_ECOBD	2,486257014	2,486257014	DNA-directed RNA polymerase subunit...	**rpoA** ECBD_0456, HO396_16075	transcription: regulation, initiation, termination
A0A140N3D6_ECOBD	1,566160693	1,566160693	Catabolite activator protein	**crp** ECBD_0391, HO396_16390	transcription: regulation, initiation, termination
A0A140N683_ECOBD	1,659661707	1,659661707	RNA polymerase sigma factor RpoD	**rpoD** ECBD_0674, HO396_14965	transcription: regulation, initiation, termination
A0A140N6I0_ECOBD	1,566006197	1,566006197	HTH-type transcriptional regulator ...	**galR** ECBD_0887, HO396_13725	transcription: regulation, initiation, termination
A0A140N6J0_ECOBD	2,068401615	2,068401615	DNA-directed RNA polymerase subunit...	**rpoZ** ECBD_0076, HO396_17915	transcription: regulation, initiation, termination
A0A140N6Q0_ECOBD	1,343326004	1,343326004	Transcription elongation factor Gre...	**greB** ECBD_0339	transcription: regulation, initiation, termination
A0A140N6R1_ECOBD	2,102483642	2,102483642	DNA-binding protein	**stpA** ECBD_1050, HO396_12880	transcription: regulation, initiation, termination
A0A140N749_ECOBD	NaN	NaN	RNA polymerase sigma factor	**rpoE** ECBD_1108, HO396_12565	transcription: regulation, initiation, termination
A0A140N785_ECOBD	1,688346029	1,688346029	HTH-type transcriptional repressor ...	**nanR** ECBD_0521, HO396_15735	transcription: regulation, initiation, termination
A0A140N7D2_ECOBD	NaN	NaN	Transcriptional regulator PhoB	**phoB** ECBD_0374, HO396_16475	transcription: regulation, initiation, termination
A0A140N7D6_ECOBD	2,537728387	2,537728387	Transcription termination/antitermi...	**nusA** ECBD_0571, HO396_15475	transcription: regulation, initiation, termination
A0A140N7P3_ECOBD	1,600500752	1,600500752	RNA polymerase sigma-54 factor	**rpoN** ECBD_0540, HO396_15640	transcription: regulation, initiation, termination
A0A140N937_ECOBD	1,413350845	1,413350845	DNA-binding transcriptional regulat...	**ydfH** ECBD_2099, HO396_07780	transcription: regulation, initiation, termination
A0A140NBI7_ECOBD	NaN	NaN	Transcriptional regulator, TetR fam...	ECBD_2027	transcription: regulation, initiation, termination
A0A140NCH9_ECOBD	2,260642936	2,260642936	DNA-binding protein	**hns** ECBD_2385, HO396_06315	transcription: regulation, initiation, termination
A0A140NCP4_ECOBD	2,417519113	2,417519113	Transcription-repair-coupling facto...	**mfd** ECBD_2487	transcription: regulation, initiation, termination
A0A140NF01_ECOBD	1,963684673	1,963684673	Transcription termination factor Rh...	**rho** ECBD_4257, HO396_18710	transcription: regulation, initiation, termination
A0A140NG87_ECOBD	NaN	NaN	Transcriptional regulator MraZ	**mraZ** ECBD_3536, HO396_00405	transcription: regulation, initiation, termination
A0A140NGM4_ECOBD	NaN	NaN	HTH-type transcriptional regulator ...	**ulaR** ECBD_3843, HO396_20855	transcription: regulation, initiation, termination
A0A140NH27_ECOBD	2,033564102	2,033564102	DNA-directed RNA polymerase subunit...	**rpoC** ECBD_4045, HO396_19815	transcription: regulation, initiation, termination
A0A140NHL8_ECOBD	2,09019138	2,09019138	Transcription termination/antitermi...	**nusG** ECBD_4051, HO396_19785	transcription: regulation, initiation, termination
A0A140NI00_ECOBD	1,671048798	1,671048798	Transcription antitermination prote...	**rfaH** ECBD_4183, HO396_19090	transcription: regulation, initiation, termination
A0A140SS80_ECOBD	2,476778848	2,476778848	DNA-directed RNA polymerase subunit...	**rpoB** ECBD_4046, HO396_19810	transcription: regulation, initiation, termination
A0A140N3T4_ECOBD	1,840087474	1,840087474	Translation initiation factor IF-2	**infB** ECBD_0572, HO396_15470	translation: regulation, initiation
A0A140N6C6_ECOBD	2,542539645	2,542539645	Elongation factor 4	**lepA** ECBD_1112, HO396_12545	translation: regulation, initiation
A0A140N6E7_ECOBD	2,010027151	2,010027151	Peptide chain release factor 2	**prfB** ECBD_0846, HO396_13935	translation: regulation, initiation
A0A140N7C8_ECOBD	1,714325265	1,714325265	SsrA-binding protein	**smpB** ECBD_1067, HO396_12795	translation: regulation, initiation
A0A140N9R4_ECOBD	1,827844667	1,827844667	Translation initiation factor IF-3	**infC** ECBD_1927, HO396_08670	translation: regulation, initiation
A0A140NCI6_ECOBD	1,352429122	1,352429122	Elongation factor Tu	**tuf** ECBD_4053, HO396_19775	translation: regulation, initiation
A0A140NCP8_ECOBD	1,500612867	1,500612867	Ribosomal silencing factor RsfS	**rsfS** rsfA, ECBD_3014, HO396_03015	translation: regulation, initiation
A0A140NFM7_ECOBD	2,553374934	2,553374934	50S ribosomal protein L1	**rplA** ECBD_4049, HO396_19795	translation: regulation, initiation
A0A140N775_ECOBD	1,32727143	1,32727143	Malate dehydrogenase	**mdh** ECBD_0511, HO396_15785	tricarboxylic acid cycle
A0A140N9G2_ECOBD	2,715709578	2,715709578	Succinate--CoA ligase [ADP-forming]...	**sucD** ECBD_2932, HO396_03475	tricarboxylic acid cycle
A0A140NA80_ECOBD	2,715709578	2,715709578	Succinate dehydrogenase flavoprotei...	**sdhA** ECBD_2937, HO396_03450	tricarboxylic acid cycle
A0A140NAN3_ECOBD	1,762171174	1,762171174	Isocitrate dehydrogenase [NADP]	**icd** ECBD_2463, HO396_05890	tricarboxylic acid cycle
A0A140NBF4_ECOBD	2,715709578	2,715709578	Succinate--CoA ligase [ADP-forming]...	**sucC** ECBD_2933, HO396_03470	tricarboxylic acid cycle
A0A140NC10_ECOBD	2,258292721	2,258292721	Citrate synthase	**gltA** ECBD_2941, HO396_03430	tricarboxylic acid cycle
A0A140NDX4_ECOBD	1,64536185	1,64536185	Dihydrolipoyllysine-residue succiny...	**odhB** ECBD_2934, HO396_03465	tricarboxylic acid cycle
A0A140NDZ9_ECOBD	2,233174935	2,233174935	Succinate dehydrogenase iron-sulfur...	**sdhB** ECBD_2936, HO396_03455	tricarboxylic acid cycle
A0A140NE66_ECOBD	1,781439302	1,781439302	Oxoglutarate dehydrogenase (succiny...	**sucA** ECBD_2935, HO396_03460	tricarboxylic acid cycle

A0A140NFP9_ECOBD	2,434120388	2,434120388	**Aconitate hydratase B**	**acnB** ECBD_3501, HO396_00580	tricarboxylic acid cycle
A0A140NGN0_ECOBD	1,948708918	1,948708918	**Aspartate ammonia-lyase**	**aspA** ECBD_3892, HO396_20590	tricarboxylic acid cycle
A0A140SS67_ECOBD	1,792683719	1,792683719	**Phosphoenolpyruvate carboxylase**	**ppc** ECBD_4068, HO396_19660	tricarboxylic acid cycle

Table A.7: Proteins enriched in the commercial extract (derived from the comparison of the commercial extract and the shaking flask batches S5/L0.5)

Uniprot ID	Fold change	-log10(FDR)	Protein name	Gene name	Keyword
A0A140N4Y5_ECOBD	1,629766172	1,629766172	Aspartate-semialdehyde dehydrogenas...	**asd** ECBD_0309	amino-acid related
A0A140N627_ECOBD	1,975462238	1,975462238	S-adenosylmethionine synthase	**metK** ECBD_0798, HO396_14190	amino-acid related
A0A140N770_ECOBD	2,503832349	2,503832349	4-hydroxy-tetrahydrodipicolinate sy...	**dapA** ECBD_1211, HO396_12065	amino-acid related
A0A140N7T1_ECOBD	1,651582735	1,651582735	Succinyl-diaminopimelate desuccinyl...	**dapE** ECBD_1218, HO396_12030	amino-acid related
A0A140NDW9_ECOBD	2,165393138	2,165393138	2,3,4,5-tetrahydropyridine-2,6-dica...	**dapD** ECBD_3453, HO396_00830	amino-acid related
A0A140NEB2_ECOBD	1,671048798	1,671048798	4-hydroxy-tetrahydrodipicolinate re...	**dapB** ECBD_3585, HO396_00160	amino-acid related
A0A140N783_ECOBD	2,505331739	2,505331739	Glyceraldehyde-3-phosphate dehydrog...	**gapA** ECBD_1865, HO396_08975	carbohydrate metabolism and respiration
A0A140NB59_ECOBD	2,004735299	2,004735299	Phosphofructokinase	**pfkB** ECBD_1922, HO396_08695	carbohydrate metabolism and respiration
A0A140N640_ECOBD	2,233174935	2,233174935	2,3-bisphosphoglycerate-independent...	**gpml** gpmM, pgml, ECBD_0113, HO396_17735	glycolytic process
A0A140N6G0_ECOBD	1,995921991	1,995921991	Enolase	**eno** ECBD_0950, HO396_13405	glycolytic process
A0A140N821_ECOBD	2,123646427	2,123646427	Fructose-bisphosphate aldolase	**fbaA** ECBD_0813, HO396_14115	glycolytic process
A0A140N8E1_ECOBD	2,672314268	2,672314268	Phosphoglycerate kinase	**pgk** ECBD_0812, HO396_14120	glycolytic process
A0A140N9C3_ECOBD	2,486257014	2,486257014	Glucokinase	**glk** ECBD_1284, HO396_11675	glycolytic process
A0A140N9D9_ECOBD	1,935051911	1,935051911	2,3-bisphosphoglycerate-dependent p...	**gpmA** ECBD_2912, HO396_03615	glycolytic process
A0A140N9V8_ECOBD	2,631339246	2,631339246	Pyruvate kinase	**pykF** ECBD_1969, HO396_08460	glycolytic process
A0A140NCD7_ECOBD	2,715709578	2,715709578	Glucose-6-phosphate isomerase	**pgi** ECBD_4012, HO396_20000	glycolytic process
A0A140NDL0_ECOBD	1,660343405	1,660343405	Pyruvate dehydrogenase E1 component	ECBD_3505	glycolytic process
A0A140NE27_ECOBD	1,831677521	1,831677521	Acetyltransferase component of pyru...	**aceF** ECBD_3504, HO396_00565	glycolytic process
A0A140NFX2_ECOBD	2,233174935	2,233174935	Probable phosphoglycerate mutase Gp...	**gpmB** ECBD_3625, HO396_21960	glycolytic process
A0A140N4P3_ECOBD	1,728448851	1,728448851	4-hydroxy-3-methylbut-2-en-1-yl dip...	**ispG** gcpE, ECBD_1171, HO396_12260	other
A0A140N5W3_ECOBD	1,91680254	1,91680254	Alanine transaminase AlaA	**alaA** ECBD_1369, HO396_11250	other
A0A140N6S0_ECOBD	1,799059234	1,799059234	2-C-methyl-D-erythritol 2,4-cyclodi...	**ispF** ECBD_0978, HO396_13270	other
A0A140N733_ECOBD	1,435570997	1,435570997	2-C-methyl-D-erythritol 4-phosphate...	**ispD** ECBD_0977, HO396_13275	other
A0A140N7Z8_ECOBD	1,812227448	1,812227448	Dihydropteroate synthase	**folP** ECBD_0565, HO396_15515	other
A0A140N8J7_ECOBD	1,343326004	1,343326004	dTDP-4-dehydrorhamnose reductase	**rfbD** ECBD_1615, HO396_10030	other
A0A140NAW9_ECOBD	1,65836961	1,65836961	Adenylate kinase	**adk** ECBD_3182, HO396_02190	other
A0A140NDD0_ECOBD	1,575341538	1,575341538	Thymidylate kinase	**tmk** ECBD_2503, HO396_05690	other
A0A140NDZ1_ECOBD	1,531912673	1,531912673	Thiamine pyrophosphate protein TPP ...	**poxB** ECBD_2723, HO396_04570	other
A0A140NFK0_ECOBD	2,258292721	2,258292721	Methionine synthase	**metH** ECBD_4018, HO396_19970	other
A0A140NGD5_ECOBD	1,663133441	1,663133441	4-hydroxy-3-methylbut-2-enyl diphos...	**ispH** lytB, ECBD_3587, HO396_00150	other
A0A140N457_ECOBD	2,434120388	2,434120388	Guanylate kinase	**gmk** ECBD_0077, HO396_17910	other

Table A.8: Proteins enriched in the shaking flask extracts S5/L0.5 (derived from the comparison of the commercial extract and the shaking flask batches S5/L0.5)

S5/L1 vs. S0/L1

S5/L1 proteins in GO terms with GO FDR <0.05

Entry name	Protein name	Gene name	GO (biological process)	Keyword

S5/L1 proteins in GO terms with GO FDR < 0.05

Entry name	Protein name	Gene name	GO (biological process)	Keyword
A0A140NAS5_ECOBD	Transcriptional regulator, AsnC fam...	ECBD_2 706		Transcriptional regulator
A0A140N479_ECOBD	**DNA mismatch repair protein MutS**	**mutS** ECBD_0 991	mismatch repair	DNA repair
A0A140NF57_ECOBD	Transcriptional regulator, LysR fam...	ECBD_4 063		Transcriptional regulator
A0A140NFS3_ECOBD	Transcriptional regulator, LacI fam...	ECBD_4 090	regulation of transcription, DNA templated	Transcriptional regulator
A0A140N8H2_ECOBD	**Nucleoid-associated protein YejK**	**yejK** ECBD_1 471		DNA relaxation
A0A140N6R1_ECOBD	**DNA-binding protein**	ECBD_1 050	regulation of transcription, DNA templated	Transcriptional regulator
A0A140NCD3_ECOBD	**Recombination-associated protein Rd...**	**rdgC** ECBD_3 268	DNA recombination	DNA recombination

S0/L1 proteins in GO terms with GO FDR <0.05

S0/L1 vs. S15L1

S15/L1 proteins in GO terms with GO FDR <0.05

Entry name	Protein name	Gene name	GO (biological process)	Keyword
A0A140N231_ECOBD	Transcriptional regulator, IclR fam...	ECBD_0 160	regulation of transcription, DNA-templated	Transcriptional regulator
A0A140N3Y5_ECOBD	Transcriptional regulator, LysR fam...	ECBD_0 633		Transcriptional regulator
A0A140N479_ECOBD	**DNA mismatch repair protein MutS**	**mutS** ECBD_0 991	mismatch repair	DNA repair
A0A140N6R1_ECOBD	**DNA-binding protein**	ECBD_1 050	regulation of transcription, DNA-templated	Transcriptional regulator
A0A140N7H8_ECOBD	Transcriptional regulator, DeoR fam...	ECBD_0 611		Transcriptional regulator
A0A140N7Q6_ECOBD	**DNA-binding protein Fis**	**fis** ECBD_0 484		Transcriptional regulator
A0A140N7T4_ECOBD	**DNA topoisomerase 4 subunit A**	**parC** ECBD_0 720	chromosome segregation; DNA topological change	DNA relaxation
A0A140N8H2_ECOBD	**Nucleoid-associated protein YejK**	**yejK** ECBD_1 471		DNA relaxation
A0A140N8I2_ECOBD	Transcriptional regulator, DeoR fam...	ECBD_2 335		Transcriptional regulator
A0A140N9L7_ECOBD	ROK family protein	ECBD_2 986		Transcriptional regulator
A0A140NAS5_ECOBD	Transcriptional regulator, AsnC fam...	ECBD_2 706		Transcriptional regulator
A0A140NB26_ECOBD	Transcriptional regulator, DeoR fam...	ECBD_1 874		Transcriptional regulator
A0A140NBN2_ECOBD	CI repressor	ECBD_2 773	negative regulation of transcription, DNA-templated; protein complex oligomerization	Transcriptional regulator
A0A140NBP1_ECOBD	**Sugar fermentation stimulation prot...**	**sfsA** ECBD_3 473		Carbohydrate metabolism
A0A140NBR4_ECOBD	**Chromosome partition protein MukB**	**mukB** ECBD_2 671	cell cycle; cell division; chromosome condensation; chromosome segregation; DNA replication	DNA replication
A0A140NC75_ECOBD	Phage shock protein B	ECBD_2 312	phage shock; regulation of transcription, DNA-templated	Transcriptional regulator
A0A140NC83_ECOBD	**Transcriptional regulatory protein ...**	ECBD_2 294	regulation of transcription, DNA-templated	Transcriptional regulator

A0A140NCD3_ECOBD	**Recombination-associated protein Rd...**	**rdgC** ECBD_3 268	DNA recombination		DNA recombination
A0A140NDA9_ECOBD	Transcriptional regulator, LysR fam...	ECBD_3 416			Transcriptional regulator
A0A140NDV2_ECOBD	**Integration host factor subunit bet...**	**ihfB** himD, ECBD_2 683	DNA recombination; regulation of transcription, DNA-templated; regulation of translation		DNA recombination
A0A140NEJ9_ECOBD	Transcriptional regulator, IclR fam...	ECBD_3 151	regulation of transcription, DNA-templated		Transcriptional regulator

S0/L1 proteins in GO terms with GO FDR <0.05

Bioreactor vs.
Commercial extract

Bioreactor proteins with GO FDR <0.05

Entry name	Protein name	Gene name	GO (biological process)	Keyword
A0A140N487_ECOBD	**2-amino-3-ketobutyrate coenzyme A l...**	**kbl** ECBD_0 108	biosynthetic process; L-threonine catabolic process to glycine	Amino acid related
A0A140N5K6_ECOBD	Endoribonuclease L-PSP	ECBD_0 627		Nuclease
A0A140N6I5_ECOBD	**Selenide, water dikinase**	**selD** ECBD_1 880	selenocysteine biosynthetic process	Other
A0A140N6K1_ECOBD	**Amino-acid acetyltransferase**	**argA** ECBD_0 907	arginine biosynthetic process	Amino acid related
A0A140N725_ECOBD	Phosphoadenosine phosphosulfate red...	**cysH** ECBD_0 967	hydrogen sulfide biosynthetic process; sulfate assimilation, phosphoadenylyl sulfate reduction by phosphoadenylyl-sulfate reductase (thioredoxin)	Other
A0A140N770_ECOBD	4-hydroxy-tetrahydrodipicolinate sy...	**dapA** ECBD_1 211	diaminopimelate biosynthetic process; lysine biosynthetic process via diaminopimelate	Amino acid related
A0A140N7T9_ECOBD	Cystathionine beta-lyase	ECBD_0 732	transsulfuration	Other
A0A140N8G9_ECOBD	**Sulfite reductase [NADPH] flavoprot...**	**cysJ** ECBD_0 965	cysteine biosynthetic process; hydrogen sulfide biosynthetic process; sulfate assimilation	Amino acid related
A0A140N8V9_ECOBD	**Adenylyl-sulfate kinase**	**cysC** ECBD_0 974	hydrogen sulfide biosynthetic process; sulfate assimilation	Other
A0A140NE69_ECOBD	3-isopropylmalate dehydratase large...	**leuC** ECBD_3 544	branched-chain amino acid biosynthetic process; leucine biosynthetic process	Amino acid related
A0A140NEC9_ECOBD	**Homoserine O-succinyltransferase**	**metAS** ECBD_4 024	L-methionine biosynthetic process from homoserine via O-succinyl-L-homoserine and cystathionine	Amino acid related

myTXTL proteins with GO FDR <0.05

Table A.9: Enriched proteins derived from the comparison of S0/L1 vs. S5/L1, S0/L1 vs. S15/L1 and of bioreactor S5/L1 against the commercial extract

Shaking flask vs. bioreactor

same lysis setting (S5/L1), different culture conditions

	SF S5/L0.5	BR S5/L0.5
total number of proteins	1480	1536
unique proteins n=1,2,3 n=0	65	0
thereof n=3 n=0	11	0
Proteins with FDR <0.05	0	8
Proteins in GO with FDR <0.05	3	0
thereof n=3 n=0	3	0

S0/L1 vs S5/L1

same lysozyme concentration, but 0 vs 5 sonication cycles

	S0/L1	S5/L1
total number of proteins	1426	1469
unique proteins n=1,2,3 n=0	42	85
thereof n=3 n=0	2	27
Proteins with FDR <0.05	0	0
Proteins in GO with FDR <0.05	0	7
thereof n=3 n=0	0	7

S0/L1vs S15/L1

same lysozyme concentration, but 0 vs. 15 sonication cycles

	S0/L1	S15/L1
total number of proteins	1426	1469
unique proteins n=1,2,3 n=0	57	120
thereof n=3 n=0	3	47
Proteins with FDR <0.05	1	6
Proteins in GO with FDR <0.05	0	21
thereof n=3 n=0	0	19

commercial extract vs. Bioreactor S5/L0.5

	myTXTL	Bioreactor
total number of proteins	1566	1536
unique proteins n=1,2,3 n=0	215	185
thereof n=3 n=0	121	30
Proteins with FDR <0.05	5	8
Proteins in GO with FDR <0.05	0	11
thereof n=3 n=0	0	5

commercial versus shaking flask S5/L0.5

	myTXTL	S5/L0.5
total number of proteins	1566	1480
unique proteins n=1,2,3 vs. n=0	214	128
thereof n=3 n=0	116	53
Proteins with FDR <0.05	309	356
Proteins in GO with FDR <0.05	172	31
thereof n=3 n=0	15	0

Table A.10: Summary of protein numbers for each comparison: total number of proteins found in the extracts, number of unique proteins, number of proteins with FDR<0.05 and numbers of proteins in GO terms with GO FDR<0.05

A.4 Plasmid sequences of reporter plasmids
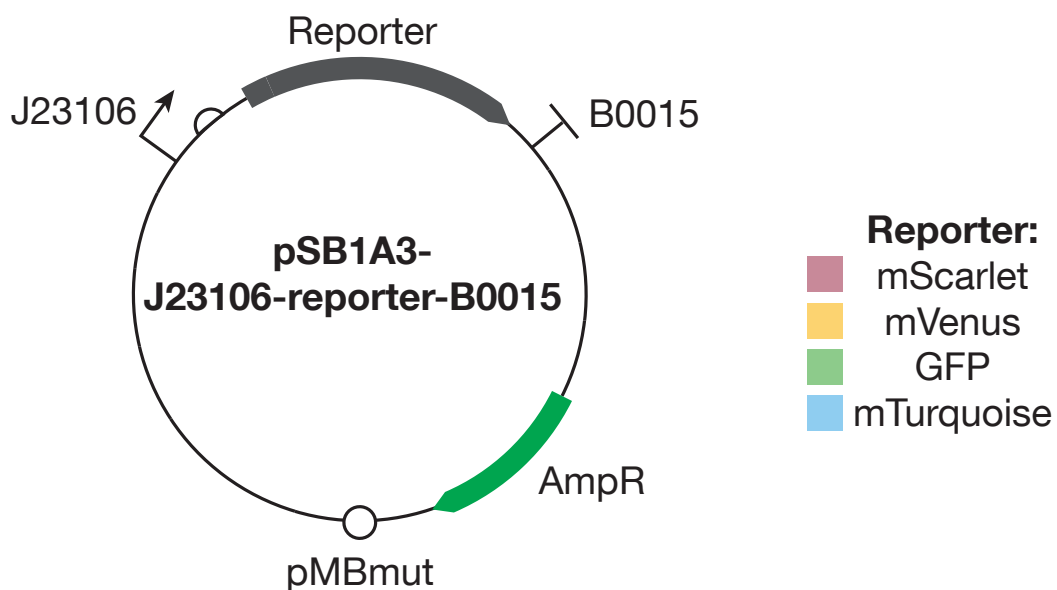
Table A.11: Plasmid maps for TXTL tests. A mScarlet, mVenus, E0040 GFP or mTurquoise reporter was cloned in a pSB1A3 backbone containing a J23106 promoter, a B0034 ribosome binding site (RBS) and a B0015 terminator.

cccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcata
gctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttat
ccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctt
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatctttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaa
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaa
tcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttacca
tctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc
gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaa
cgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatc
ccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagca
ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccg
agttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa
aactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgg
gtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggggataaagggcgacacggaaatgttgaatactcatactcttccttttttcaatatta
ttgaagcatttatcaggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaa
aagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcagaatttcagataaaaaaaaatc
cttagctttcgctaaggatgatttctggaattcgcggccgcttctagagtttacggctagctcagtcctaggtatagtgctagctactagagaaagag
gagaaactcgagatggtgtcaaagggagaggcggttatcaaggaatttatgcgctttaaagtccacatggagggcagcatgaacgggcacg
agtttgaaattgagggggagggggagggccgtccttatgaaggtactcagactgctaaactgaaggtgacaaaaggtggcccccttgcctttctc
gtgggacatcctgtcgccacaattcatgtacgggagccgcgcctttatcaaacatccccgcagatattcctgattactataaacaatctttcccgga
aggtttcaaatgggaacgcgtcatgaatttgaggacggggggcgctgtcacagttactcaggacacctccttggaagacggcacattgatttaca
aggttaagttgcgcggcacaaacttcccccctgacgggccagtaatgcaaaagaaaactatgggttgggaggcgtctacagaacgtttatacc
ccgaagacggggtgctgaaaggtgacattaagatggccctgccctgaaggacgggggtcgctatcttgccgactttaaaactacttataagg
ctaaaaaaccagtccagatgccaggcgcctataatgttgaccgcaagttagacatcacctcacataatgaagactataccgttgtagaacaat
acgagcgcagcgagggtcgtcacagtaccggggggatggatgaattatacaaataataacggatccgctgtccgccaggcatcaaataaaa
cgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggcct
ttctgcgtttatatactagtagcggccgctgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcact
caaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaacc
gtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaac
ccgacaggactataaagataccaggcgtttc

Table A.12: pSB1A3-J23106-B0034-mScarlet-B0015

cccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcata
gctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttat
ccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctt
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaa
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaa
tcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttacca
tctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc
gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaa
cgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatc
ccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagca
ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccg
agttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa
aactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgg
gtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatatta
ttgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaa
aagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcagaatttcagataaaaaaaatc
cttagctttcgctaaggatgatttctggaattcgcggccgcttctagagtttacggctagctcagtcctaggtatagtgctagctactagagaaagag
gagaaactcgagatgcaaaagagcaaaggcgaagaactgttcacgggtgtggttccgatcctggttgaactggatggcgatgtgaacggtca
taaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactgacgctgaaactgatttgcaccacgggtaaactgccggttccg
tggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacccggatcacatgaaacgccatgatttctttaaatctgcgatgcc
ggaaggctatgtgcaggaacgtaccatcttttttcaaagatgatggtaactacaaaacccgcgcggaagttaaatttgaaggcgatacgctggtg
aaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctgggtcacaaactggaatacaactacaacagtcataacgtgtaca
ttaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtcacaacatcgaagatggcggtgttcagctggccgatcattac
cagcagaacacccgattggcgatggtccggtgctgctgccggataatcattatctgagttaccagagcaaactgtctaaagatccgaatgaa
aaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacgcatggtatggatgaactgtataaataatgaggatccgctgtccg
ccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactgg
ctcaccttcgggtgggcctttctgcgtttatatactagtagcggccgctgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcgg
cgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcca
gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctca
agtcagaggtggcgaaacccgacaggactataaagataccaggcgtttc

Table A.13: pSB1A3-J23106-B0034-mVenus-B0015

cccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcata
gctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttat
ccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctt
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaa
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaa
tcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttacca
tctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc
gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaa
cgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatc
ccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagca
ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccg
agttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa
aactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgg
gtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatatta
ttgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggggttccgcgcacatttccccgaa
aagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcagaatttcagataaaaaaaaatc
cttagctttcgctaaggatgatttctggaattcgcggccgcttctagagtttacggctagctcagtcctaggtatagtgctagctactagagaaagag
gagaaactcgagatgcgtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaattttctgt
cagtggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacctgttccatggccaacacttgt
cactactttcggttatggtgttcaatgctttgcgagatacccagatcatatgaaacagcatgactttttcaagagtgccatgcccgaaggttatgtac
aggaaagaactatattttttcaaagatgacgggaactacaagacacgtgctgaagtcaagtttgaaggtgatacccttgttaatagaatcgagtta
aaaggtattgattttaaagaagatggaaacattcttggacacaaattggaatacaactataactcacacaatgtatacatcatggcagacaaac
aaaagaatggaatcaaagttaacttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaacaaaatactcca
attggcgatggccctgtcctttttaccagacaaccattacctgtccacacaatctgccctttcgaaagatcccaacgaaaagagagaccacatgg
tccttcttgagtttgtaacagctgctgggattacacatggcatggatgaactatacaaataataacggatccgctgtccgccaggcatcaaataaa
acgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggc
ctttctgcgtttatatactagtagcggccgctgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctca
ctcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaac
cgtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa
cccgacaggactataaagataccaggcgtttc

Table A.14: pSB1A3-J23106-B0034-GFP-B0015

cccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcata
gctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttat
ccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttacctt
cggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatctttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaa
aaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaa
tcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttacca
tctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc
gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaa
cgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatc
ccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagca
ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccg
agttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa
aactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgg
gtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatatta
ttgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaa
aagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcagaatttcagataaaaaaaatc
cttagctttcgctaaggatgatttctggaattcgcggccgcttctagagtttacggctagctcagtcctaggtatagtgctagctactagagaaagag
gagaaactcgagatggttagcaagggcgaagaactgttcaccggcgtcgtgccgattctggttgagctggatggtgatgtcaacggtcacaagt
ttagcgttagcggtgagggcgagggcgacgccacctacggtaaattgaccctgaagttatctgcacgaccggtaagctgccggttccgtggcc
gaccctggtgacgactctgtcgtggggcgtgcaatgtttcgcgcgctatccggatcacatgaaacagcatgacttctttaagagcgcgatgccgg
aaggctacgttcaggaacgtacgatctttttcaaagacgacggtaactataagacccgcgcagaagtcaagttcgagggtgacacgctggtga
atcgtattgagctgaaaggtattgactttaaagaggacggtaacatcctgggtcacaaactggagtataattacttcagcgacaatgtgtacatca
ccgctgataaacagaaaaacggcattaaagcaaacttcaagatccgtcacaatattgaagatggcggcgtgtgcaattggccgatcactatcaa
cagaacaccccgattggcgatggtccggtcctgctgccagataatcactacttgagcacgcaatccaaactgtccaaagatccgaacgaaaa
acgtgaccacatggtcctgctggaatttgttaccgcggcgggtatcacgtcgtgggtatggacgaactgtacaagtaattaacggatccgctgtccg
ccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactgg
ctcaccttcgggtgggcctttctgcgtttatatactagtagcggccgctgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcgg
cgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcca
gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctca
agtcagaggtggcgaaacccgacaggactataaagataccaggcgtttc

Table A.15: pSB1A3-J23106-B0034-mTurquoise-B0015

Appendix B

Supplementary figures and information for chapter 4

B.1 Supplementary figures
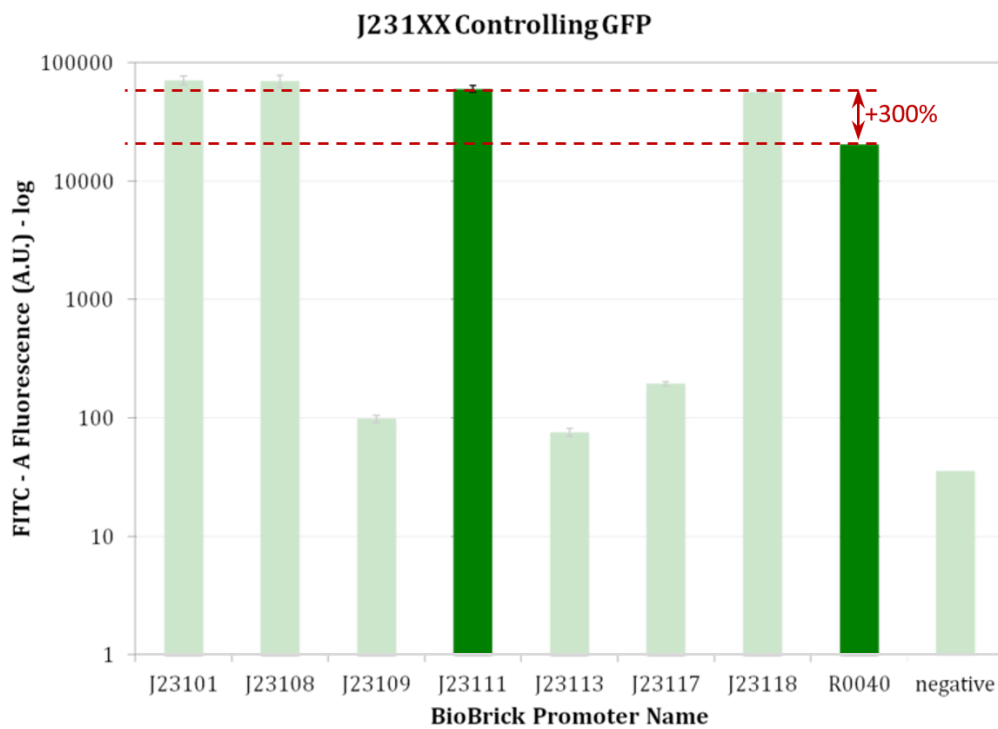
Figure B.1: The promoter strength of J23111 is about three times higher than for pLTet-O1 [151] which is registered as R0040 in the iGEM repository. Figure was adapted from the iGEM team Boston 2012 [169].
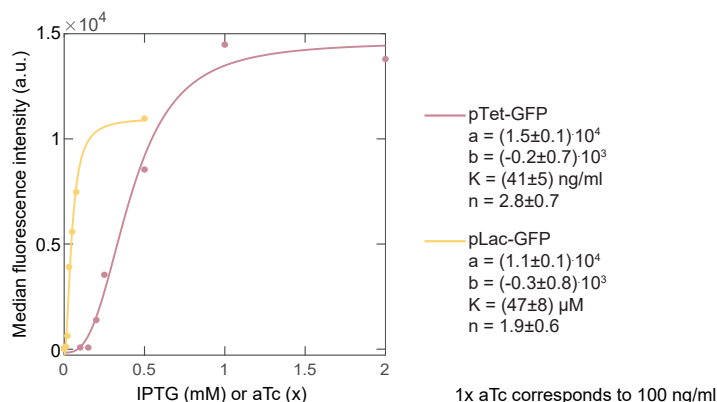
Figure B.2: Hill curves for pTet-GFP and pLac-GFP. A plasmid containing pTet-GFP or pLac-GFP was transformed to DH5alphaZ1. Overnight cultures were diluted 1:100 in M9 medium and induced with up to 200 ng/ml aTc or up to 0.5 mM IPTG at OD 0.3. After overnight induction samples were measured in a flow cytometer (CyFlow Cube 8, Sysmex) and the median fluorescence signal was plotted against the inducer concentration. Hill curves were fitted to the data and the fitting parameters are shown in the plots. The Tet promoter strength is about 36% higher, which confirms the results of Bujard et al. [151]
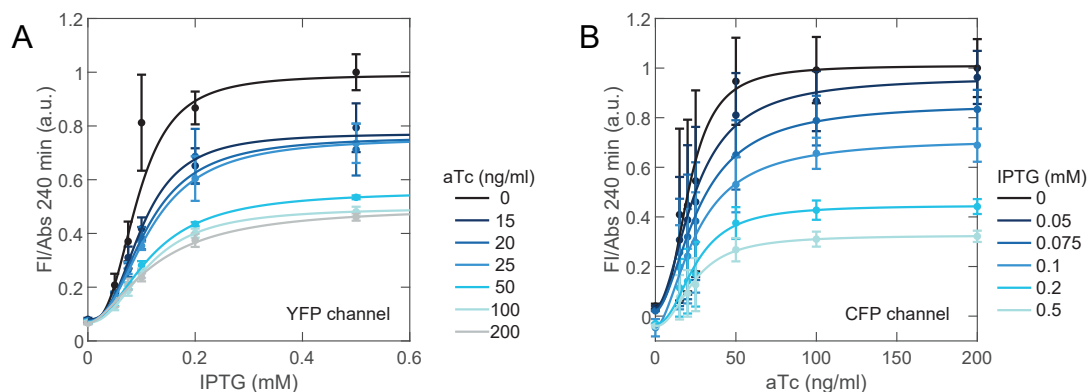
Figure B.3: A two-input two-output (2I/2O) logic gate was assembled by combining a T2AT pLac design including a YFP reporter (2I/2O YFP) and a convergent pTet design including a CFP reporter (2I/2O CFP). The FI/Abs signals after 240 min of induction were normalized to the signal at I0.5a0 or I0a200 respectively and Hill curves were fitted. 2I/2O YFP (A) shows a higher offset compared to 2I/2O CFP (B). aTc induction causes a signal decrease down to 60% of the maximum (compare I0.5a0 and I0.5a200). 2I/2O CFP can be repressed to less than 40% of the maximum under addition of IPTG (compare I0a200 and I0.5a200).

B.2 Supplementary tables

	ON-OFF	Δ(ON-OFF)	ON/OFF	Δ(ON/OFF)	ON/OFF induction	Δ(ON/OFF induction)	Distance pAT to J23111	Distance pTr to J23111
Tandem pTet long DT	4875	1483	3.0	0.6	22	7	405	858
Tandem pTet long ST	6374	604	3.8	0.5	20	5	404	857
Tandem pTet short DT	5152	505	3.0	0.3	20.7	0.6	243	696
Tandem pTet short ST	9835	504	3.5	0.2	21	9	208	661
convergent pTet	11217	141	4.1	0.1	19	3	223	360
T2AT pTet	10410	1008	4.4	0.3	7.0	0.7	223	360
Tandem pLac long DT	4870	162	1.42	0.01	31	2	446	751
Tandem pLac long ST	5163	96	1.46	0.01	30	2	407	712
Tandem pLac short DT	4213	278	1.37	0.03	31	2	229	534
Tandem pLac short ST	3755	1563	1.21	0.09	17.6	0.7	169	474
convergent pLac	5334	1354	1.35	0.09	39	2	180	309
T2AT pLac	13681	1516	2.48	0.29	12.6	0.4	229	362
2I2O CFP (normalized)	0.71	0.06	3.3	0.2	27	5	223	360
2I2O YFP (normalized)	0.55	0.07	2.2	0.2	13	5	229	362

Table B.1: Parameters calculated with equations B.1 to B.5

ON/OFF ratios (induction curves) are defined as

$$\frac{\text{ON}}{\text{OFF}} = \frac{\alpha + \beta}{\beta} \tag{B.1}$$

Errors were propagated using a quadratic addition:

$$\Delta\frac{\text{ON}}{\text{OFF}} = \Delta\frac{\alpha + \beta}{\beta} = \sqrt{\left(\frac{\Delta\alpha}{\beta}\right)^2 + \left(\frac{\alpha \cdot \Delta\beta}{\beta^2}\right)^2} \tag{B.2}$$

ON/OFF ratios (repression curves) are defined as

$$\frac{\text{ON}}{\text{OFF}} = \frac{\alpha_2 + \beta_2}{\beta_2} \tag{B.3}$$

Errors were propagated using a quadratic addition:

$$\Delta\frac{\text{ON}}{\text{OFF}} = \Delta\frac{\alpha_2 + \beta_2}{\beta_2} = \sqrt{\left(\frac{\Delta\alpha_2}{\beta_2}\right)^2 + \left(\frac{\alpha_2 \cdot \Delta\beta_2}{\beta_2^2}\right)^2} \tag{B.4}$$

The repression capability of the system is represented by the fitting parameter α_2:

$$\text{ON} - \text{OFF} = \alpha_2 \tag{B.5}$$

Promoter distances were determined from the center of the promoter sites to be independent on promoter orientation.

B.3 Plasmid sequences

aataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtcATAAATGTGAGCGGATAACATTGACAT
TGTGAGCGGATAACAAGATACTGAGCACAtcaACTGACTATTCTGTGCAATAGACAGTAAAGCAGG
GATAAACGAGATAGATAAGATAAGAATCCTACCATGATCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACA
CTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAtgatgcgtagggagatccggaatcattggcctatgtcacctc
gagTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACATCATGGTA
GGATTCTTATCTTATCTATCTCGTTTATCCCTGCTTTACTGTCTATTGCACAGAATAGTCAGTggaC
GCAAAAAACCCCGCTTCGGCGGGGTTTTTTCGCgtgcgctgatctgagAAAAAAAAAACCCCGCCCCTGA
CAGGGCGGGGTTTTTTTTggcaagggtctgatgttcactcagccgaatcaggctgctggatacatggctaagctgatttgggaatct
gtgagcgtgacggtggtagctgcggttgaagcaatgaactggcttaagtctgctgctaagctgctggctgctgaggtcaaagataagaagactg
gagagattcttcgcaagcgttgcgctgcgcattgggtaactcctgatggtttccctgtgtggcaggaatacaaggTTGACGGCTAGCTC
AGTCCTAGGTATAGTGCTAGCTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGG
AGATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGATGAGCAAAGGCGAAGAACTGT
TCACGGGTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGTGAACGGTCATAAATTTAGCGTG
TCTGGTGAAGGCGAAGGTGATGCGACCTACGGCAAACTGACGCTGAAACTGATTTGCACCACGG
GTAAACTGCCGGTTCCGTGGCCGACCCTGGTGACCACGCTGGGTTATGGTCTGATGTGTTTCGC
ACGTTACCCGGATCACATGAAACGCCATGATTTCTTTAAATCTGCGATGCCGGAAGGCTATGTGC
AGGAACGTACCATCTTTTTTCAAAGATGATGGTAACTACAAAACCCGCGCGGAAGTTAAATTTGAAG
GCGATACGCTGGTGAACCGTATTGAACTGAAAGGTATCGATTTCAAAGAAGATGGCAATATTCTG
GGTCACAAACTGGAATACAACTACAACAGTCATAACGTGTACATTACCGCCGATAAACAGAAAAA
CGGTATCAAAGCAAACTTCAAAATCCGTCACAACATCGAAGATGGCGGTGTTCAGCTGGCCGATC
ATTACCAGCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCTGCCGGATAATCATTATCTGAGT
TACCAGAGCAAACTGTCTAAAGATCCGAATGAAAAACGCGATCACATGGTTCTGCTGGAATTTGT
GACCGCGGCCGGCATTACGCATGGTATGGATGAACTGTATAAAtaaCCAGGCATCAAATAAAACGA
AAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTA
GAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagg
gcaaggtgtcaccaccctgccctttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttc
ggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa
aaggccagcaaaaggccaggaaccgtaaaaaggccgCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGA
CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC
CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGAT
ACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC
AGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA
AGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTG
GTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTA
TCAAAAAGGATCTTCACCTAGATCCTTTTAaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctg
acagttaCCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGG
AAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC
ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTA
TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccac
tcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaata
agggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatg
tatttagaaa

Table B.2: pSB1A3-tandem pLac long DT

tagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtcATAAATGTGAGCGGATAACATTG
ACATTGTGAGCGGATAACAAGATACTGAGCACAtcaACTGACTATTCTGTGCAATAGACAGTAAAG
CAGGGATAAACGAGATAGATAAGATAAGAatcctaccatgatCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACA
CTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATgatgtcgtagggagatccggaatctattggcctatgtcacctc
gagTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACAtcatggtaggat
TCTTATCTTATCTATCTCGTTTATCCCTGCTTTACTGTCTATTGCACAGAATAGTCAGTggaCGCAAA
AAACCCCGCTTCGGCGGGGTTTTTTCGCgtgcgctgatctgagggcaagggtctgatgttcactcagccgaatcaggctgct
ggatacatggctaagctgatttgggaatctgtgagcgtgacggtggtagctgcggttgaagcaatgaactggcttaagtctgctgctaagctgctg
gctgctgaggtcaaagataagaagactggagagattcttcgcaagcgttcgcgctgcgcattgggtaactcctgatggtttccctgtgtggcagga
atacaaggTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCtcttatcttatctatctcgtttatccctgcatacagaaa
cagaggagatacgcaatgataaacgagaacctggcggcagcgcaaaagATGAGCAAAGGCGAAGAACTGTTCACGG
GTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGTGAACGGTCATAAATTTAGCGTGTCTGGT
GAAGGCGAAGGTGATGCGACCTACGGCAAACTGACGCTGAAACTGATTTGCACCACGGGTAAAC
TGCCGGTTCCGTGGCCGACCCTGGTGACCACGCTGGGTTATGGTCTGATGTGTTTCGCACGTTA
CCCGGATCACATGAAACGCCATGATTTCTTTAAATCTGCGATGCCGGAAGGCTATGTGCAGGAAC
GTACCATCTTTTTCAAAGATGATGGTAACTACAAAACCCGCGCGGAAGTTAAATTTGAAGGCGATA
CGCTGGTGAACCGTATTGAACTGAAAGGTATCGATTTCAAAGAAGATGGCAATATTCTGGGTCAC
AAACTGGAATACAACTACAACAGTCATAACGTGTACATTACCGCCGATAAACAGAAAAACGGTATC
AAAGCAAACTTCAAAATCCGTCACAACATCGAAGATGGCGGTGTTCAGCTGGCCGATCATTACCA
GCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCTGCCGGATAATCATTATCTGAGTTACCAG
AGCAAACTGTCTAAAGATCCGAATGAAAAACGCGATCACATGGTTCTGCTGGAATTTGTGACCGC
GGCCGGCATTACGCATGGTATGGATGAACTGTATAAAtaaCCAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCAC
ACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtc
accaccctgccctttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggc
gagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccag
caaaaggccaggaaccgtaaaaaggccgCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT
GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC
GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
GGATCTTCACCTAGATCCTTTTAaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaCC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA
TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATC
CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacc
caactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgac
acggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatt

Table B.3: pSB1A3-tandem pLac long ST

tttagaaaaataaacaaatagggggttccgcgcacatttccccgaaaagtgccacctgacgtcATAAATGTGAGCGGATAACATT
GACATTGTGAGCGGATAACAAGATACTGAGCACAtcaACTGACTATTCTGTGCAATAGACAGTAAA
GCAGGGATAAACGAGATAGATAAGATAAGAatcctaccatgatCCAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCAC
ACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATgatgcgtagggagatccggaatcattggcctatgtcacct
cgagTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACAtcatggtagga
tTCTTATCTTATCTATCTCGTTTATCCCTGCTTTACTGTCTATTGCACAGAATAGTCAGTggaCGCAA
AAAACCCCGCTTCGGCGGGGTTTTTTTCGCgtgcgctgatctgagAAAAAAAAAACCCCGCCCCTGACAGG
GCGGGGTTTTTTTTggcaagggtctgatgttcacgTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCtctt
atcttatctatctcgtttatccctgcatacagaaacagaggagatacgcaatgataaacgagaacctggcggcagcgcaaaagATGAGCA
AAGGCGAAGAACTGTTCACGGGTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGTGAACGG
TCATAAATTTAGCGTGTCTGGTGAAGGCGAAGGTGATGCGACCTACGGCAAACTGACGCTGAAA
CTGATTTGCACCACGGGTAAACTGCCGGTTCCGTGGCCGACCCTGGTGACCACGCTGGGTTATG
GTCTGATGTGTTTCGCACGTTACCCGGATCACATGAAACGCCATGATTTCTTTAAATCTGCGATGC
CGGAAGGCTATGTGCAGGAACGTACCATCTTTTTCAAAGATGATGGTAACTACAAAACCCGCGCG
GAAGTTAAATTTGAAGGCGATACGCTGGTGAACCGTATTGAACTGAAAGGTATCGATTTCAAAGA
AGATGGCAATATTCTGGGTCACAAACTGGAATACAACTACAACAGTCATAACGTGTACATTACCGC
CGATAAACAGAAAAACGGTATCAAAGCAAACTTCAAAATCCGTCACAACATCGAAGATGGCGGTG
TTCAGCTGGCCGATCATTACCAGCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCTGCCGGA
TAATCATTATCTGAGTTACCAGAGCAAACTGTCTAAAGATCCGAATGAAAAACGCGATCACATGGT
TCTGCTGGAATTTGTGACCGCGGCCGGCATTACGCATGGTATGGATGAACTGTATAAAtaaCCAGG
CATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTG
AACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgct
gcagtccggcaaaaaagggcaaggtgtcaccaccctgcccttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactg
actcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcag
gaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgCGTTGCTGGCGTTTTTCCACAGGC
TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC
TGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA
CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT
GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAaattaaaaatgaagttttaaatcaatctaaagtat
atatgagtaaacttggtctgacagttaCCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAA
CCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
CATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCT
CCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCG
CCACATAGCAGAACTTTAAAAGTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgag
atccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaa
tgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcat
gagcggatacatatttgaatgta

Table B.4: pSB1A3-tandem pLac short DT

gacgtcATAAATGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGATACTGAGCACAtcaACT
GACTATTCTGTGCAATAGACAGTAAAGCAGGGATAAACGAGATAGATAAGATAAGAATCCTACCAT
GATCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTT
TGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTAT
AtgatgcgtagggagatccggaatctattggcctatgtcacctcgagTCCCTATCAGTGATAGAGATTGACATCCCTATC
AGTGATAGAGATACTGAGCACATCATGGTAGGATTCTTATCTTATCTATCTCGTTTATCCCTGCTTT
ACTGTCTATTGCACAGAATAGTCAGTggaCGCAAAAAACCCCGCTTCGGCGGGGTTTTTTCGCgtgc
gctgatctgagTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCTCTTATCTTATCTATCTCGTTTAT
CCCTGCATACAGAAACAGAGGAGATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGA
TGAGCAAAGGCGAAGAACTGTTCACGGGTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGT
GAACGGTCATAAATTTAGCGTGTCTGGTGAAGGCGAAGGTGATGCGACCTACGGCAAACTGACG
CTGAAACTGATTTGCACCACGGGTAAACTGCCGGTTCCGTGGCCGACCCTGGTGACCACGCTGG
GTTATGGTCTGATGTGTTTCGCACGTTACCCGGATCACATGAAACGCCATGATTTCTTTAAATCTG
CGATGCCGGAAGGCTATGTGCAGGAACGTACCATCTTTTTCAAAGATGATGGTAACTACAAAACC
CGCGCGGAAGTTAAATTTGAAGGCGATACGCTGGTGAACCGTATTGAACTGAAAGGTATCGATTT
CAAAGAAGATGGCAATATTCTGGGTCACAAACTGGAATACAACTACAACAGTCATAACGTGTACAT
TACCGCCGATAAACAGAAAAACGGTATCAAAGCAAACTTCAAAATCCGTCACAACATCGAAGATG
GCGGTGTTCAGCTGGCCGATCATTACCAGCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCT
GCCGGATAATCATTATCTGAGTTACCAGAGCAAACTGTCTAAAGATCCGAATGAAAAACGCGATC
ACATGGTTCTGCTGGAATTTGTGACCGCGGCCGGCATTACGCATGGTATGGATGAACTGTATAAA
taaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTT
GTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA
actagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgcccttttctttaaaaccgaaaagattacttcgcgttatgcaggcttc
ctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcaGCTCACTCAAAGGCGGTAATacggttatccacag
aatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttc
cacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggc
gtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct
catagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgc
cttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgag
gtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagt
taccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcag
aaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagatt
atcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaCCAA
TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTC
CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC
CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGA
GCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA
GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATG
ATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCC
GTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAG
TGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcaccca
actgatcttcagcatctttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataaaggcgacac
ggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcaggGTTATTGTCTCATGAGCGGAtacatatttgaatgt
atttagaaaaataaacaaatagggggttccgcgcacatttccccgaaaagtgccacct

Table B.5: pSB1A3-tandem pLac short ST

cGCGAAAAAACCCCGCCGAAGCGGGGTTTTTTGCGATAAATGTGAGCGGATAACATTGACATTGT
GAGCGGATAACAAGATACTGAGCACAtccACTGACTATTCTGTGCAATAGACAGTAAAGCAGGGAT
AAACGAGATAGATAAGATAAGAATGGAACCATGATGTGCTCAGTATCTCTATCACTGATAGGGAT
GTCAATCTCTATCACTGATAGGGACCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGG
GCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCG
GGTGGGCCTTTCTGCGTTTATAgaattcTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCgggTC
TTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGATACGCAATGATAAACGAGAA
CCTGGCGGCAGCGCAAAAGATGAGCAAAGGCGAAGAACTGTTCACGGGTGTGGTTCCGATCCTG
GTTGAACTGGATGGCGATGTGAACGGTCATAAATTTAGCGTGTCTGGTGAAGGCGAAGGTGATG
CGACCTACGGCAAACTGACGCTGAAACTGATTTGCACCACGGGTAAACTGCCGGTTCCGTGGCC
GACCCTGGTGACCACGCTGGGTTATGGTCTGATGTGTTTCGCACGTTACCCGGATCACATGAAAC
GCCATGATTTCTTTAAATCTGCGATGCCGGAAGGCTATGTGCAGGAACGTACCATCTTTTTCAAAG
ATGATGGTAACTACAAAACCCGCGCGGAAGTTAAATTTGAAGGCGATACGCTGGTGAACCGTATT
GAACTGAAAGGTATCGATTTCAAAGAAGATGGCAATATTCTGGGTCACAAACTGGAATACAACTAC
AACAGTCATAACGTGTACATTACCGCCGATAAACAGAAAAACGGTATCAAAGCAAACTTCAAAATC
CGTCACAACATCGAAGATGGCGGTGTTCAGCTGGCCGATCATTACCAGCAGAACACCCCGATTG
GCGATGGTCCGGTGCTGCTGCCGGATAATCATTATCTGAGTTACCAGAGCAAACTGTCTAAAGAT
CCGAATGAAAAACGCGATCACATGGTTCTGCTGGAATTTGTGACCGCGGCCGGCATTACGCATG
GTATGGATGAACTGTATAAAtaaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC
CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGG
TGGGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgccctttttctttaaaacc
gaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaagg
cggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacag
gactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtccgcctttctcccttt
cgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccg
ttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaa
caggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatct
gcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaa
gcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaa
gggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaactt
ggtctgacagttaCCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGT
GCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA
GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatata
acccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaatgccgcaaaaaag
ggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatat
ttgaatgtatttagaaaaataaacaaatagggggttccgcgcacatttccccgaaaagtgccacctgacgt

Table B.6: pSB1A3-convergent pLac

actcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttc
cgcgcacatttccccgaaaagtgccacctgacgtctaagaagacctcGCGAAAAAACCCCGCCGAAGCGGGGTTTTTT
GCGcgaccacgacggctAAAAAAAACCCCGCCCTGTCAGGGGCGGGGTTTTTTTTTTgagcGGGACTGAC
TATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAGATAAGATAAGATAGATCCTACCAT
GATTTAAACAAAATTATTTGTAGAGGACTGTTTCGGCCCTTTTGGGCCATCGTCAGGTCGGATACA
CATCCGGCGACAGTCTctcgagATAAATGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGA
TACTGAGCACAgggactgactattctgtgcaatagtcagtaaagcagggataaacgagatagataagataagatagatcctaccatga
tGTGCTCAGTATCTCTATCACTGATAGGGATGTCAATCTCTATCACTGATAGGGAgagaCCAGGCAT
CAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAAC
GCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGtgcgctgacaagtc
cattaccaaggagtgcgatgcagttactcagaattcTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCgggTCT
TATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGATACGCAATGATAAACGAGAAC
CTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacgggtgtggttccgatcctggttgaactggatggcgatgtga
acggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactgacgctgaaactgatttgcaccacgggtaaactgc
cggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacccggatcacatgaaacgccatgatttctttaaatct
gcgatgccggaaggctatgtgcaggaacgtaccatcttttttcaaagatgatggtaactacaaaacccgcgcggaagttaaatttgaaggcgat
acgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctgggtcacaaactggaatacaactacaacagtcata
acgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtcacaacatcgaagatggcggtgttcagctggcc
gatcattaccagcagaacacccgattggcgatggtccggtgctgctgccggataatcattatctgagttaccagagcaaactgtctaaagatcc
gaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacgcatggtatggatgaactgtataaataaCCAGGC
ATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGA
ACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgctg
cagtccggcaaaaaagggcaaggtgtcaccaccctgcccttttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactga
ctcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcagg
aaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctga
cgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctc
gtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctccttcgggaagcgtggcgctttctcatagctcacgctgtaggtat
ctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg
agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacaga
gttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg
tagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaa
gatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcaccta
gatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctat
ctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctg
caatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctg
caactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgcta
caggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccatgttgtgcaa
aaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctta
ctgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccg
gcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatct
taccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaaca
ggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcat

Table B.7: pSB1A3-T2AT pLac

tgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaa
ataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaagacttcTCCCTATCAGTGATAGAGATTGACAT
CCCTATCAGTGATAGAGATACTGAGCACatcaGGGACTGACTATTCTGTGCAATAGTCAGTAAAGC
AGGGATAAACGAGATAGATAAGATAAGATAGATCCTACCATGATCCAGGCATCAAATAAAACGAA
AGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAG
AGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATgatgcgtagggagatcctttaggcggtgctta
ctcttacataaaggggctgttagtattaccccgcgaggattacatgcctatatttgtctctttgccggcttatatggacaagcatagcattcgaaaag
gtgagccggaaacatatgctgtctctggcacgttcattggcctatgtcacctcgagATAAATGTGAGCGGATAACATTGACAT
TGTGAGCGGATAACAAGATACTGAGCACAATCATGGTAGGATCTATCTTATCTTATCTATCTCGTT
TATCCCTGCTTTACTGACTATTGCACAGAATAGTCAGTCCCggaCGCAAAAAACCCCGCTTCGGCG
GGGTTTTTTCGCtcgtactgctagtagaggcAAAAAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTTTag
ccgtcgtggtgaagtccgtacacggtatgatcggacgcctcgtgagatcaatacgtataccaggtgtcctgtgagcagcgaaagactcacagg
ctccgtaatactagcatgcgataagtccctaactgactatggccttctgattgatgccacggcgatctccgtgcgctgatctgaggaattcTTGA
CGGCTAGCTCAGTCCTAGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATA
CAGAAACAGAGGAGATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcg
aagaactgttcacgggtgtggttccgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgat
gcgacctacggcaaactgacgctgaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggt
ctgatgtgtttcgcacgttacccggatcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatcttttt
caaagatgatggtaactacaaaacccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaa
agaagatggcaatattctgggtcacaaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatca
aagcaaacttcaaaatccgtcacaacatcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattggcgatggtccg
gtgctgctgccggataatcattatctgagttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtg
accgcggccggcattacgcatggtatggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGA
AAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCT
CACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgc
ccttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatca
gctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca
ggaaccgtaaaaaggccgcgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtgg
cgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacct
gtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt
gcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggca
gcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaag
aacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcg
gtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaa
cgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaa
gtatatatgagtaaacttggtctgacagttaCCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATT
TCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCAT
CTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAAT
AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAG
TCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTT
GCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC
CTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCA
TTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCG
CGCCACATAGCAGAACTTTAAAAGTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgtt
gagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggca
aaatgccgcaaaaaagggaataaaggggcgacacggaaatgt

Table B.8: pSB1A3-tandem pTet long DT

tgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaa
atagggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaagacttcTCCCTATCAGTGATAGAGATTGACAT
CCCTATCAGTGATAGAGATACTGAGCACatcaGGGACTGACTATTCTGTGCAATAGTCAGTAAAGC
AGGGATAAACGAGATAGATAAGATAAGATAGATCCTACCATGATCCAGGCATCAAATAAAACGAA
AGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAG
AGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAtgatgcgtagggagatcctttaggcggtgctta
ctcttacataaaggggctgttagtattaccccgcgaggattacatgcctatatttgtctctttgccggcttatatggacaagcatagcattcgaaaag
gtgagccggaaacatatgctgtctctggcacgttctattggcctatgtcacctcgagATAAATGTGAGCGGATAACATTGACAT
TGTGAGCGGATAACAAGATACTGAGCACAATCATGGTAGGATCTATCTTATCTTATCTATCTCGTT
TATCCCTGCTTTACTGACTATTGCACAGAATAGTCAGTCCCggaCGCAAAAAACCCCGCTTCGGCG
GGGTTTTTTCGCtcgtactgctagtagaggcgaagtgttcaacgcgttcgccgtatggaatgtgcgcgttgccgtcgtggtgaagtccgta
cacggtatgatcggacgcctcgtgagatcaatacgtataccaggtgtcctgtgagcagcgaaagactcacaggctccgtaatactagcatgcg
ataagtccctaactgactatggccttctgattgatgccacggcgatctccgtgcgctgatctgaggaattcTTGACGGCTAGCTCAGT
CCTAGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGA
GATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacgggtgtg
gttccgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactga
cgctgaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacc
cggatcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatcttttttcaaagatgatggtaactaca
aaacccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctggg
tcacaaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtc
acaacatcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattggcgatggtccggtgctgctgccggataatcatta
tctgagttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacgcat
ggtatggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC
GTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGG
CCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgccctttttctttaaaaccgaaaa
gattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaa
tacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccg
cgttgctggcgtttttccacaggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactat
aaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcggga
agcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccccgttcagc
ccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggat
tagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctc
tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagc
agattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggat
tttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtct
gacagttaCCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGG
AAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC
ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTA
TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccac
tcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaata
agggcgacacggaaatgt

Table B.9: pSB1A3-tandem pTet long ST

tgaatactcatactcttccttttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaa
ataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaagacttcTCCCTATCAGTGATAGAGATTGACAT
CCCTATCAGTGATAGAGATACTGAGCACAtcaGGGACTGACTATTCTGTGCAATAGTCAGTAAAGC
AGGGATAAACGAGATAGATAAGATAAGATAGatcctaccatgatCCAGGCATCAAATAAAACGAAAGGC
TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCA
CACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAtgatgcgtagggagatcctttaggcggtgcttactcttacat
aaaggggctgttagtattacccgcgaggattacatgcctatatttgtctctttgccggcttatatggacaagcatagcattcgaaaaggtgagccg
gaaacatatgctgtctctggcacgttctattggcctatgtcacctcgagATAAATGTGAGCGGATAACATTGACATTGTGAG
CGGATAACAAGATACTGAGCACAatcatggtaggatCTATCTTATCTTATCTATCTCGTTTATCCCTGCTT
TACTGACTATTGCACAGAATAGTCAGTCCCggaCGCAAAAAACCCCGCTTCGGCGGGGTTTTTTCG
CtcgtactgctagtagaggcAAAAAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTTTagccgtcgtggtgaagt
ccgtaccTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTAT
CCCTGCATACAGAAACAGAGGAGATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGat
gagcaaaggcgaagaactgttcacgggtgtggttccgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaag
gcgaaggtgatgcgacctacggcaaactgacgctgaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacg
ctgggttatggtctgatgtgtttcgcacgttacccggatcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaac
gtaccatcttttcaaagatgatggtaactacaaaacccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggt
atcgatttcaaagaagatggcaatattctgggtcacaaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaa
aacggtatcaaagcaaacttcaaaatccgtcacaacatcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattgg
cgatggtccggtgctgctgccggataatcattatctgagttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgc
tggaatttgtgaccgcggccggcattacgcatggtatggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACA
CTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtca
ccaccctgccctttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcg
agcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggggataacgcaggaaagaacatgtgagcaaaaggccagc
aaaaggccaggaaccgtaaaaaggccgCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCAT
CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT
GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC
GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
GGATCTTCACCTAGATCCTTTTAaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaCC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA
TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATC
CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacc
caactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgac
acggaaatgt

Table B.10: pSB1A3-tandem pTet short DT

tttagaaaaataaacaaatagggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaagacttcTCCCTATCAGTGAT
AGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACAtcaGGGACTGACTATTCTGTGCAAT
AGTCAGTAAAGCAGGGATAAACGAGATAGATAAGATAAGATAGatcctaccatgatCCAGGCATCAAATA
AAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATgatgcgtagggagatcctttag
gcggtgcttactcttacataaaggggctgttagtattaccccgcgaggattacatgcctatatttgtctctttgccggcttatatggacaagcatagca
ttcgaaaaggtgagccggaaacatatgctgtctctggcacgttcattggcctatgtcacctcgagATAAATGTGAGCGGATAACAT
TGACATTGTGAGCGGATAACAAGATACTGAGCACAatcatggtaggatCTATCTTATCTTATCTATCTCG
TTTATCCCTGCTTTACTGACTATTGCACAGAATAGTCAGTCCCggaCGCAAAAAACCCCGCTTCGG
CGGGGTTTTTTCGCtcgtactgctagtagaggcgatctccgtgcgctgatctgaggaattcTTGACGGCTAGCTCAGTCC
TAGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGA
TACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacgggtgtggtt
ccgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactgacg
ctgaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacccg
gatcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatctttttcaaagatgatggtaactacaaa
acccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctgggtc
acaaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtca
caacatcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattggcgatggtccggtgctgctgccggataatcattat
ctgagttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacgcatg
gtatggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG
TTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGC
CTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgcccttttttctttaaaaccgaaaaga
ttacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatac
ggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgCG
TTGCTGGCGTTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCA
GAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG
TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG
AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGC
GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAC
GAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAa
attaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaCCAATGCTTAATCAGTGAGGCACCT
ATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACG
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC
TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
TAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAA
GCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCAT
GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATcattggaaaacgttcttcggggcg
aaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttct
gggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaata
ttattgaagcatttatcaggggttattgtctcatgagcggatacatatttgaatgta

Table B.11: pSB1A3-tandem pTet short ST

aggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttttcaatattattgaagcatttatcagggt
tattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggggttccgcgcacatttccccgaaaagtgccacctgacgtct
aagaaGCGAAAAAACCCCGCCGAAGCGGGGTTTTTTTGCGcgacccggctAAAAAAAACCCCGCCCTGT
CAGGGGCGGGGTTTTTTTTTTgatagacTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATA
GAGATACTGAGCACtattGGGACTGACTATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGAT
AGATAAGATAAGATAGATCCTACCATGATTGTGCTCAGTATCTTGTTATCCGCTCACAATGTCAAT
GTTATCCGCTCACATTTATgctagCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC
CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGG
TGGGCCTTTCTGCGTTTATAttacctgaagaataagtcagccgtgtaacccgatgaggaattcTTGACGGCTAGCTCA
GTCCTAGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAG
GAGATACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacggg
tgtggttccgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaac
tgacgctgaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgtta
cccggatcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatctttttcaaagatgatggtaacta
caaaacccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctg
ggtcacaaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatcc
gtcacaacatcgaagatggcggtgttcagctggccgatcattaccagcagaacacccgattggcgatggtccggtgctgctgccggataatc
attatctgagttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacg
catggtatggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTT
TCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTG
GGCCTTTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgcccttttttctttaaaaccgaa
aagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggt
aatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggc
cgCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGA
AGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAA
GCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA
GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT
TTTAaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaaacttggtctgacagttaCCAATGCTTAATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAA
CTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTC
ACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCT
GCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTG
GTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC
AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATC
ACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT
GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATcattggaaaacgttc
ttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatcttttactttcac
cagcgtttctgggtgagcaaaaacagga

Table B.12: pSB1A3-convergent pTet

ctcttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggggttcc
gcgcacatttccccgaaaagtgccacctgacgtctaagaagacctcGCGAAAAAACCCCGCCGAAGCGGGGTTTTTTG
CGcgaccacgacggctAAAAAAAACCCCGCCCTGTCAGGGGCGGGGTTTTTTTTgagcGGGACTGACT
ATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAGATAAGATAAGATAGATCCTACCATG
ATTTAAACAAAATTATTTGTAGAGGACTGTTTCGGCCCTTTTGGGCCATCGTCAGGTCGGATACAC
ATCCGGCGACAGTCTctcgaggacTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGA
TACTGAGCACTattGGGACTGACTATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAGAT
AAGATAAGATAGATCCTACCATGATTGTGCTCAGTATCTTGTTATCCGCTCACAATGTCAATGTTAT
CCGCTCACATTTATgctagCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC
GTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGG
CCTTTCTGCGTTTATAttacctgaagaataagtcagccgtgtaacccgatgaggaattcTTGACGGCTAGCTCAGTCCT
AGGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGAT
ACGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacgggtgtggttc
cgatcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactgacgct
gaaactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacccgga
tcacatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatcttttcaaagatgatggtaactacaaaac
ccgcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctgggtcac
aaactggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtcaca
acatcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattggcgatggtccggtgctgctgccggataatcattatctg
agttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgcggccggcattacgcatggta
tggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTT
TTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCT
TTCTGCGTTTATAactagtgcgctgcagtccggcaaaaaagggcaaggtgtcaccaccctgccctttttctttaaaaccgaaaagatta
cttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacgg
ttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgCGTT
GCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA
GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCT
TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCG
CAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAaa
ttaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaCCAATGCTTAATCAGTGAGGCACCTA
TCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGA
TACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC
TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
TAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAA
GCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCAT
GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATcattggaaaacgttcttcggggcg
aaaactctcaaggatcttaccgctgttgagatccagttcgatataacccactcgtgcacccaactgatcttcagcatctttactttcaccagcgtttct
gggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcata

Table B.13: pSB1A3-T2AT pTet

tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgt
agttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaact
gtcagaccaagtttactcatatatactttagattgatttaaaacttcattttaatttaaaaggatctaggtgaagatcctttttgataatctcatgaccaa
aatcccttaacgtgagttttcgttccactgagcgtcagaccccttaataagatgatcttcttgagatcgttttggtctgcgcgtaatctcttgctctgaaa
acgaaaaaaccgccttgcagggcggtttttcgaaggttctctgagctaccaactctttgaaccgaggtaactggcttggaggagcgcagtcacc
aaaacttgtcctttcagtttagccttaaccggcgcatgacttcaagactaactcctctaaatcaattaccagtggctgctgccagtggtgctttttgcat
gtctttccgggttggactcaagacgatagttaccggataaggcgcagcggtcggactgaacggggggttcgtgcatacagtccagcttggagc
gaactgcctacccggaactgagtgtcaggcgtggaatgagacaaacgcggccataacagcggaatgacaccggtaaaccgaaaggcag
gaacaggagagcgcacgagggagccgcccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccaccactgatttgagcgtcagat
ttcgtgatgcttgtcaggggggcggagcctatggaaaaacggctttgccgcggccctctcacttccctgttaagtatcttcctggcatcttccagga
aatctccgccccgttcgtaagccatttccgctcgccgcagtcgaacgaccgagcgtagcgagtcagtgagcgaggaagcggaatatatcctgt
atcacatattctgctgacgcaccggtgcagccttttttctcctgccacatgaagcacttcactgacaccctcatcagtgccaacatagtaagccagt
atacactccgctagcgctgaggtctgcctcgtgaagaaggtgttgctgactcataccaggcctgaatcgccccatcatccagccagaaagtga
gggagccacggttgatgagagctttgttgtaggtggaccagttggtgattttgaactttgctttgccacggaacggtctgcgttgtcgggaagatgc
gtgatctgatccttcaactcagcaaaagttcgatttattcaacaaagccacgttgtgtctcaaaatctctgatgttacattgcacaagataaaaatat
atcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgtttactagaggttgatcgggcacgtaagaggttccaactttca
ccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaaaATGGAGAAAAAAATCA
CGGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAG
TTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGA
AAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAACGCTCACCCGG
AGTTTCGTATGGCCATGAAAGACGGTGAGCTGGTGATCTGGGATAGTGTTCACCCTTGTTACACC
GTTTTCCATGAGCAAACTGAAACGTTTTCGTCCCTCTGGAGTGAATACCACGACGATTTCCGGCA
GTTTCTCCACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAG
GGTTTATTGAGAATATGTTTTTTGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAA
CGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACGATGGGCAAATATTATACGCAAGGCG
ACAAGGTGCTGATGCCGCTGGCGATCCAGGTTCATCATGCCGTTTGTGATGGCTTCCATGTCGG
CCGCATGCTTAATGAATTACAACAGTACTGTGATGAGTGGCAGGGCGGGGCGTAAtaatactagctccg
gcaaaaaaacgggcaaggtgtcaccaccctgccctttttctttaaaaccgaaaagattacttcgcgtttgccacctgacgtctaagaaGCGA
AAAAACCCCGCCGAAGCGGGGTTTTTTGCGcgacccggctAAAAAAAACCCCGCCCCTGTCAGGGGCG
GGGTTTTTTTTTTgatagacTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTG
AGCACTattGGGACAGATCCACTGAGGCGTGGATCTGTGAACACTAAACTAAACGACAATGTAATCA
AACTAACggatggctacttgTGTGCTCAGTATCTTGTTATCCGCTCACAATGTCAATGTTATCCGCTCAC
ATTTATgctagCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATC
TGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG
CGTTTATAttacctgaagaataagtcagccgtgtaacccgatgaggaattcTTGACGGCTAGCTCAGTCCTAGGTATAG
TGCTAGCGGGAGTTTGATTACATTGTCGTTTAGTTTAGTGATACATAAACAGAGGAGATATCACAT
GACTAAACGAAACCTGGCGGCAGCGCAAAAGatggtgagtaaaggcgaagagctgttcacaggggttgttccgattctgg
tcgaactggacggggacgttaatggtcacaaattcagcgttagcggtgagggcgagggtgatgccacttatggtgaaactgaccctgaaattcat
ctgtaccaccggcaaactgcctgttccttggcctacactggttacaacactgacttggggtgttcaatgtttgctcgctatccggatcacatgaaac
agcacgatttcttcaaaagcgccatgcctgaaggttatgtccaagagctgtacgatcttcttttaaagacgacggcaactataaaaccgtgccga
ggtgaaattcgaaggtgataccctggtaaaccgtatcgaactgaaagggatcgacttcaaagaggacgggaacattctgggccataaactgg
agtataacgccatcagcgataatgtgtatattaccgccgacaaacagaaaaacgggatcaaagccaacttcaaaatccgccacaacatcga
ggatggtagcgttcaactggccgatcactatcaacagaataccccgattggtgatggtcctgttctgtgctgcctgataaccactatctgagcaccca
gtctaaactgtccaaagacccgaacgagaaacgtgatcacatggttctgctggagtttgttaccgctgccggcattactctgggtatggatgaact
gtataaatagcATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGctgcagtccggcaaaaaag
ggcaaggtgtcaccaccctgccctttttctttaaaaccgaaaagattacttcgcgttatgcaggcttcctcgctcactgactcgctgcgctcggtcgtt
cggctgcggcgagcggtatcagctcactaagcttgacctcGCGAAAAAAACCCCGCCGAAGCGGGGTTTTTTGCGcga
ccacgacggctAAAAAAAAACCCCGCCCTGTCAGGGGCGGGGTTTTTTTTTgagcGGGACTGACTATTCT
GTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAGATAAGATAAGATAGATCCTACCATGATTTA
AACAAAATTATTTGTAGAGGACTGTTTCGGCCCTTTTGGGCCATCGTCAGGTCGGATACACATCC
GGCGACAGTCTctcgagATAAATGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGATACTG
AGCACAGGGACTGACTATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAGATAAGATA
AGATAGATCCTACCATGATGTGCTCAGTATCTCTATCACTGATAGGGATGTCAATCTCTATCACTG
ATAGGGAgagaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTAT
CTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCT
GCGTTTATAgtgcgctgacaagtccattaccaaggagtgcgatgcagttactcagaattcTTGACGGCTAGCTCAGTCCTA
GGTATAGTGCTAGCgggTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGATA

CGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAGatgagcaaaggcgaagaactgttcacgggtgtggttccg
atcctggttgaactggatggcgatgtgaacggtcataaatttagcgtgtctggtgaaggcgaaggtgatgcgacctacggcaaactgacgctga
aactgatttgcaccacgggtaaactgccggttccgtggccgaccctggtgaccacgctgggttatggtctgatgtgtttcgcacgttacccggatc
acatgaaacgccatgatttctttaaatctgcgatgccggaaggctatgtgcaggaacgtaccatctttttcaaagatgatggtaactacaaaaccc
gcgcggaagttaaatttgaaggcgatacgctggtgaaccgtattgaactgaaaggtatcgatttcaaagaagatggcaatattctgggtcacaa
actggaatacaactacaacagtcataacgtgtacattaccgccgataaacagaaaaacggtatcaaagcaaacttcaaaatccgtcacaac
atcgaagatggcggtgttcagctggccgatcattaccagcagaacaccccgattggcgatggtccggtgctgctgccggataatcattatctga
gttaccagagcaaactgtctaaagatccgaatgaaaaacgcgatcacatggttctgctggaatttgtgaccgccggccggcattacgcatggtat
ggatgaactgtataaataaCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTT
TATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTT
TCTGCGTTTATAactagtagcggccgctgcaggagtcactaagggttagttagttagattagcagaaagtcaaaagcctccgaccgga
ggcttttgactaaaacttcccttggggttatcattggggctcactcaaaggcggtaatcagataaaaaaaatccttagctttcgctaaggatgatttc
tgctagagatggaatagactggatgggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc

Table B.14: pSB3C5-2I/2O (convergent pTet-Tr2 J23111-TS2-mCerulean + T2AT pLac-Tr J23111-TS1-mVenus)

Abbreviations

3-PGA	3-Phosphoglyceric acid
AA	amino acids
ABS	absorbance
AGC	automatic gain control
AT	anti-trigger; antagonist to Tr
aTc	anhydrotetracycline
ATP	adenosine triphosphate
BCA	bicinchoninic acid
CFP	cyan fluorescent protein
CoA	coenzyme A
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleic acid
DT	double terminator
DTT	dithiothreitol
ETC	electron transport chain
FDR	false discovery rate
FI	fluorescence intensity
FMN	flavin mononucleotide
G1P	glucose-1-phosphate
G6P	glucose-6-phosphate
GFP	green fluorescent protein
GO	gene ontology
HDC	higher energy collision induced dissociation
HPLC	high performance liquid chromatography
iGEM	international genetically engineered machine
IPTG	Isopropyl β- d-1-thiogalactopyranoside
Ixay	inducer concentration of x mM IPTG in combination with y ng/ml aTc
mCer	mCerulean, a CFP
mS	mScarlet, a RFP
mV	mVenus, a YFP

MWCO	molecular weight cutoff
NAD	Nicotinamide adenine dinucleotide
nfH$_2$O	nuclease free water
NTP	nucleosid triphosphate
OD$_{600}$	optical densitiy = absorbance measurement at a wave length of 600 nm
ori	origin of replication
PANOx	PEP, AA, NAD, oxalic acid
PBS	phospate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEP	phosphoenolpyruvate
PFU/ml	phage forming units per milliliter
PSM	peptide spectrum match
RBS	ribosome binding site
RFP	red fluorescent protein
RNA	ribonucleic acid (mRNA: messenger RNA, tRNA: transfer-RNA, ncRNA: non-coding RNA, sRNA: small RNA)
RNAP	RNA polymerase
Sx/Ly	Cell lysis protocol with x sonication cycles at a lysozyme concentration of y mg/mL
TCA	tricarboxylic acid cycle
TI	transcriptional interference
Tr	*trans*-acting trigger RNA, activates toehold switch
TS	toehold switch
TXTL	transcription and translation
TXTL test	cell-free transcription-translation (gene expression) test
YFP	yellow fluorescent protein

Bibliography

[1] E. Falgenhauer, S. von Schönberg, C. Meng, A. Mückl, K. Vogele, Q. Emslander, C. Ludwig, and F. C. Simmel, "Evaluation of an e. coli cell extract prepared by lysozyme-assisted sonication via gene expression, phage assembly and proteomics.," *Chembiochem: a European journal of chemical biology*, vol. 22, no. 18, pp. 2805–2813, 2021.

[2] E. Falgenhauer, A. Mückl, M. Schwarz-Schilling, and F. C. Simmel, "Transcriptional interference in toehold switch-based rna circuits," *ACS Synthetic Biology*, vol. 11, no. 5, pp. 1735–1745, 2022.

[3] E. C. D. G. for Research, *Synthetic Biology: Applying Engineering to Biology : Report of a NEST High-Level Expert Group.* Community research: Project Report, Office for Official Publications of the European Communities, 2005.

[4] F. Jacob and J. Monod, "Genetic regulatory mechanisms in the synthesis of proteins," *Journal of molecular biology*, vol. 3, no. 3, pp. 318–356, 1961.

[5] W. Arber and S. Linn, "Dna modification and restriction," *Annual review of biochemistry*, vol. 38, no. 1, pp. 467–500, 1969.

[6] V. C. Bode and A. Kaiser, "Changes in the structure and activity of λ dna in a superinfected immune bacterium," *Journal of molecular biology*, vol. 14, no. 2, pp. 399–417, 1965.

[7] N. Cozzarelli, N. E. Melechen, T. M. Jovin, and A. Kornberg, "Polynucleotide cellulose as a substrate for a polynucleotide ligase induced by phage t4," *Biochemical and biophysical research communications*, vol. 28, no. 4, pp. 578–586, 1967.

[8] M. L. Gefter, A. Becker, and J. Hurwitz, "The enzymatic repair of dna, formation of circular lambda-dna.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 58, no. 1, p. 240, 1967.

[9] M. Gellert, "Formation of covalent circles of lambda dna by e. coli extracts.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 57, no. 1, p. 148, 1967.

[10] B. M. Olivera and I. Lehman, "Linkage of polynucleotides through phosphodiester bonds by an enzyme from escherichia coli.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 57, no. 5, p. 1426, 1967.

[11] K. B. Mullis and H. A. Erlich, "Primer-directed enzymatic amplification of dna with a thermostable dna polymerase," *Science*, vol. 239, no. 4839, pp. 487–491, 1988.

[12] F. Sanger, S. Nicklen, and A. R. Coulson, "Dna sequencing with chain-terminating inhibitors," *Proceedings of the national academy of sciences*, vol. 74, no. 12, pp. 5463–5467, 1977.

[13] iGEM Foundation; https://parts.igem.org/, "Registry of standard biological parts." Accessed: 2021-09-28.

[14] P. E. Purnick and R. Weiss, "The second wave of synthetic biology: from modules to systems," *Nature reviews Molecular cell biology*, vol. 10, no. 6, pp. 410–422, 2009.

[15] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–342, 2000.

[16] B. P. Kramer, A. U. Viretta, M. Daoud-El Baba, D. Aubel, W. Weber, and M. Fussenegger, "An engineered epigenetic transgene switch in mammalian cells," *Nature biotechnology*, vol. 22, no. 7, pp. 867–870, 2004.

[17] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[18] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty, "A fast, robust and tunable synthetic gene oscillator," *Nature*, vol. 456, no. 7221, pp. 516–519, 2008.

[19] S. Hooshangi, S. Thiberge, and R. Weiss, "Ultrasensitivity and noise propagation in a synthetic transcriptional cascade," *Proceedings of the National Academy of Sciences*, vol. 102, no. 10, pp. 3581–3586, 2005.

[20] A. A. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, and C. A. Voigt, "Genetic circuit design automation," *Science*, vol. 352, no. 6281, 2016.

[21] D. Del Vecchio, A. J. Dy, and Y. Qian, "Control theory meets synthetic biology," *Journal of The Royal Society Interface*, vol. 13, no. 120, p. 20160380, 2016.

[22] M. Weiss, J. P. Frohnmayer, L. T. Benk, B. Haller, J.-W. Janiesch, T. Heitkamp, M. Börsch, R. B. Lira, R. Dimova, R. Lipowsky, *et al.*, "Sequential bottom-up assembly of mechanically stabilized synthetic cells by microfluidics," *Nature materials*, vol. 17, no. 1, pp. 89–96, 2018.

[23] O. wet ware; https://openwetware.org/wiki/E. coli genotypes, "*E. coli* genotypes." Accessed: 2021-11-19.

[24] M. Roberts, R. Cranenburgh, M. Stevens, and P. Oyston, "Synthetic biology: biology by design," *Microbiology*, vol. 159, no. Pt 7, p. 1219, 2013.

[25] E. Karzbrun, J. Shin, R. H. Bar-Ziv, and V. Noireaux, "Coarse-grained dynamics of protein synthesis in a cell-free system," *Physical review letters*, vol. 106, no. 4, p. 048104, 2011.

[26] A. D. Silverman, A. S. Karim, and M. C. Jewett, "Cell-free gene expression: an expanded repertoire of applications," *Nature Reviews Genetics*, vol. 21, no. 3, pp. 151–170, 2020.

[27] M. W. Nirenberg and J. H. Matthaei, "The dependence of cell-free protein synthesis in e. coli upon naturally occurring or synthetic polyribonucleotides," *Proceedings of the National Academy of Sciences*, vol. 47, no. 10, pp. 1588–1602, 1961.

[28] R. S. Fuller, J. M. Kaguni, and A. Kornberg, "Enzymatic replication of the origin of the escherichia coli chromosome," *Proceedings of the National Academy of Sciences*, vol. 78, no. 12, pp. 7370–7374, 1981.

[29] T. Preiss and M. W. Hentze, "Dual function of the messenger rna cap structure in poly (a)-tail-promoted translation in yeast," *Nature*, vol. 392, no. 6675, pp. 516–520, 1998.

[30] F. Caschera and V. Noireaux, "Synthesis of 2.3 mg/ml of protein with an all escherichia coli cell-free transcription–translation system," *Biochimie*, vol. 99, pp. 162–168, 2014.

[31] J. Chappell, K. Jensen, and P. S. Freemont, "Validation of an entirely in vitro approach for rapid prototyping of dna regulatory elements for synthetic biology," *Nucleic acids research*, vol. 41, no. 5, pp. 3471–3481, 2013.

[32] R. Kelwick, L. Ricci, S. M. Chee, D. Bell, A. J. Webb, and P. S. Freemont, "Cell-free prototyping strategies for enhancing the sustainable production of polyhydroxyalkanoates bioplastics," *Synthetic Biology*, vol. 3, no. 1, p. ysy016, 2018.

[33] D. Garenne, S. Thompson, A. Brisson, A. Khakimzhan, and V. Noireaux, "The all-e. coli txtl toolbox 3.0: new capabilities of a cell-free synthetic biology platform," *Synthetic Biology*, vol. 6, no. 1, p. ysab017, 2021.

[34] P. Zhang, J. Wang, X. Ding, J. Lin, H. Jiang, H. Zhou, and Y. Lu, "Exploration of the tolerance ability of a cell-free biosynthesis system to toxic substances," *Applied biochemistry and biotechnology*, vol. 189, no. 4, pp. 1096–1107, 2019.

[35] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, and T. Ueda, "Cell-free translation reconstituted with purified components," *Nature biotechnology*, vol. 19, no. 8, pp. 751–755, 2001.

[36] L. Grasemann, B. Lavickova, M. C. Elizondo-Cantú, and S. J. Maerkl, "Onepot pure cell-free system.," *Journal of Visualized Experiments: Jove*, no. 172, 2021.

[37] G. Zubay, "In vitro synthesis of protein in microbial systems," *Annual review of genetics*, vol. 7, no. 1, pp. 267–287, 1973.

[38] F. Caschera and V. Noireaux, "A cost-effective polyphosphate-based metabolism fuels an all e. coli cell-free expression system," *Metabolic engineering*, vol. 27, pp. 29–37, 2015.

[39] H. Niederholtmeyer, V. Stepanova, and S. J. Maerkl, "Implementation of cell-free biological networks at steady state," *Proceedings of the National Academy of Sciences*, vol. 110, no. 40, pp. 15985–15990, 2013.

[40] J. Shin, P. Jardine, and V. Noireaux, "Genome replication, synthesis, and assembly of the bacteriophage t7 in a single cell-free reaction," *ACS synthetic biology*, vol. 1, no. 9, pp. 408–413, 2012.

[41] S. D. Cole, A. E. Miklos, A. C. Chiao, Z. Z. Sun, and M. W. Lux, "Methodologies for preparation of prokaryotic extracts for cell-free expression systems," *Synthetic and Systems Biotechnology*, vol. 5, no. 4, pp. 252–267, 2020.

[42] K. Sitaraman, D. Esposito, G. Klarmann, S. F. Le Grice, J. L. Hartley, and D. K. Chatterjee, "A novel cell-free protein synthesis system," *Journal of biotechnology*, vol. 110, no. 3, pp. 257–263, 2004.

[43] J. L. Dopp, Y. R. Jo, and N. F. Reuel, "Methods to reduce variability in e. coli-based cell-free protein expression experiments," *Synthetic and systems biotechnology*, vol. 4, no. 4, pp. 204–211, 2019.

[44] F. Caschera, A. S. Karim, G. Gazzola, A. E. d'Aquino, N. H. Packard, and M. C. Jewett, "High-throughput optimization cycle of a cell-free ribosome assembly and protein synthesis system," *ACS synthetic biology*, vol. 7, no. 12, pp. 2841–2853, 2018.

[45] O. Borkowski, M. Koch, A. Zettor, A. Pandi, A. C. Batista, P. Soudier, and J.-L. Faulon, "Large scale active-learning-guided exploration for in vitro protein production optimization," *Nature communications*, vol. 11, no. 1, pp. 1–8, 2020.

[46] A. M. Miguez, M. P. McNerney, and M. P. Styczynski, "Metabolic profiling of escherichia coli-based cell-free expression systems for process optimization," *Industrial & engineering chemistry research*, vol. 58, no. 50, pp. 22472–22482, 2019.

[47] D. Garenne, M. C. Haines, E. F. Romantseva, P. Freemont, E. A. Strychalski, and V. Noireaux, "Cell-free gene expression," *Nature Reviews Methods Primers*, vol. 1, no. 1, pp. 1–18, 2021.

[48] T. Yamane, Y. Ikeda, T. Nagasaka, and H. Nakano, "Enhanced cell-free protein synthesis using a s30 extract from escherichia coli grown rapidly at 42° c in an amino acid enriched medium," *Biotechnology progress*, vol. 21, no. 2, pp. 608–613, 2005.

[49] L. R. Burrington, K. R. Watts, and J. P. Oza, "Characterizing and improving reaction times for e. coli-based cell-free protein synthesis," *ACS Synthetic Biology*, vol. 10, no. 8, pp. 1821–1829, 2021.

[50] Y.-C. Kwon and M. C. Jewett, "High-throughput preparation methods of crude extract for robust cell-free protein synthesis," *Scientific reports*, vol. 5, no. 1, pp. 1–8, 2015.

[51] A. L. Koch, "Shrinkage of growing escherichia coli cells by osmotic challenge," *Journal of bacteriology*, vol. 159, no. 3, pp. 919–924, 1984.

[52] Y. Chisti and M. Moo-Young, "Disruption of microbial cells for intracellular products," *Enzyme and Microbial Technology*, vol. 8, no. 4, pp. 194–204, 1986.

[53] H. Lange, P. Taillandier, and J.-P. Riba, "Effect of high shear stress on microbial viability," *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology*, vol. 76, no. 5, pp. 501–505, 2001.

[54] P. Gray, M. LILLY, and P. DUNNILL, "Kinetics of beta-galactosidase production by a constitutive mutant of escherichia-coli," *JOURNAL OF FERMENTATION TECHNOLOGY*, vol. 50, no. 6, p. 381, 1972.

[55] W. Broothaerts, J. McPherson, B. Li, E. Randall, W. D. Lane, and P. A. Wiersma, "Fast apple (malus× domestica) and tobacco (nicotiana tobacum) leaf polyphenol oxidase activity assay for screening transgenic plants," *Journal of agricultural and food chemistry*, vol. 48, no. 12, pp. 5924–5928, 2000.

[56] A. Bolano, S. Stinchi, R. Preziosi, F. Bistoni, M. Allegrucci, F. Baldelli, A. Martini, and G. Cardinali, "Rapid methods to extract dna and rna from cryptococcus neoformans," *FEMS yeast research*, vol. 1, no. 3, pp. 221–224, 2001.

[57] K. Bielawski, A. Bernat, M. Własiuk, and B. Falkiewicz, "Hcv-rna detection in liver bioptates–a comparison of automatic and home-made protocols combined with a new procedure of hcv-rna extraction.," *Medical science monitor: international medical journal of experimental and clinical research*, vol. 7, pp. 197–201, 2001.

[58] S. Fujimoto, Y. Nakagami, and F. Kojima, "Optimal bacterial dna isolation method using bead-beating technique," *Memoirs Kyushu Univ Dep Of Health Scis Of Medical Sch*, vol. 3, pp. 33–38, 2004.

[59] Z. Z. Sun, C. A. Hayes, J. Shin, F. Caschera, R. M. Murray, and V. Noireaux, "Protocols for implementing an escherichia coli based tx-tl cell-free expression system for synthetic biology," *Journal of Visualized Experiments*, 2013.

[60] P. Shrestha, T. M. Holland, and B. C. Bundy, "Streamlined extract preparation for escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing," *Biotechniques*, vol. 53, no. 3, pp. 163–174, 2012.

[61] E. Uhlmann, D. Oberschmidt, A. Spielvogel, K. Herms, M. Polte, J. Polte, and A. Dumke, "Development of a versatile and continuously operating cell disruption device," *Procedia CIRP*, vol. 5, pp. 119–123, 2013.

[62] S. Ferdous, J. L. Dopp, and N. F. Reuel, "Optimization of e. coli tip-sonication for high-yield c ell-free extract using finite element modeling," *AIChE Journal*, vol. 67, no. 10, p. e17389, 2021.

[63] L. Stryer, *Biochemie, 4. Auflage*. Spektrum Akademischer Verlag, 1996.

[64] W. Vollmer, D. Blanot, and M. A. De Pedro, "Peptidoglycan structure and architecture," *FEMS microbiology reviews*, vol. 32, no. 2, pp. 149–167, 2008.

[65] R. Irvin, T. MacAlister, and J. Costerton, "Tris (hydroxymethyl) aminomethane buffer modification of escherichia coli outer membrane permeability," *Journal of bacteriology*, vol. 145, no. 3, pp. 1397–1403, 1981.

[66] K. Fujiwara and N. Doi, "Biochemical preparation of cell extract for cell-free protein synthesis without physical disruption," *PloS one*, vol. 11, no. 4, p. e0154614, 2016.

[67] D. Garenne, C. L. Beisel, and V. Noireaux, "Characterization of the all-e. coli transcription-translation system mytxtl by mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 33, no. 11, pp. 1036–1048, 2019.

[68] D.-M. Kim and J. R. Swartz, "Oxalate improves protein synthesis by enhancing atp supply in a cell-free system derived from escherichia coli," *Biotechnology Letters*, vol. 22, no. 19, pp. 1537–1542, 2000.

[69] D. Siegal-Gaskins, Z. A. Tuza, J. Kim, V. Noireaux, and R. M. Murray, "Gene circuit performance characterization and resource usage in a cell-free "breadboard"," *ACS synthetic biology*, vol. 3, no. 6, pp. 416–425, 2014.

[70] M. Schwarz-Schilling, A. Dupin, F. Chizzolini, S. Krishnan, S. S. Mansy, and F. C. Simmel, "Optimized assembly of a multifunctional rna-protein nanostructure in a cell-free gene expression system," *Nano letters*, vol. 18, no. 4, pp. 2650–2657, 2018.

[71] F. Chizzolini, M. Forlin, N. Yeh Martín, G. Berloffa, D. Cecchi, and S. S. Mansy, "Cell-free translation is more variable than transcription," *ACS synthetic biology*, vol. 6, no. 4, pp. 638–647, 2017.

[72] J. H. Davis, A. J. Rubin, and R. T. Sauer, "Design, construction and characterization of a set of insulated bacterial promoters," *Nucleic acids research*, vol. 39, no. 3, pp. 1131–1141, 2011.

[73] V. K. Mutalik, J. C. Guimaraes, G. Cambray, C. Lam, M. J. Christoffersen, Q.-A. Mai, A. B. Tran, M. Paull, J. D. Keasling, A. P. Arkin, *et al.*, "Precise and reliable gene expression via standard transcription and translation initiation elements," *Nature methods*, vol. 10, no. 4, pp. 354–360, 2013.

[74] C. Lou, B. Stanton, Y.-J. Chen, B. Munsky, and C. A. Voigt, "Ribozyme-based insulator parts buffer synthetic circuits from genetic context," *Nature biotechnology*, vol. 30, no. 11, pp. 1137–1142, 2012.

[75] L. Qi, R. E. Haurwitz, W. Shao, J. A. Doudna, and A. P. Arkin, "Rna processing enables predictable programming of gene expression," *Nature biotechnology*, vol. 30, no. 10, pp. 1002–1006, 2012.

[76] D. Del Vecchio, "Modularity, context-dependence, and insulation in engineered biological circuits," *Trends in biotechnology*, vol. 33, no. 2, pp. 111–119, 2015.

[77] D.-M. Kim and J. R. Swartz, "Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis," *Biotechnology and bioengineering*, vol. 74, no. 4, pp. 309–316, 2001.

[78] Y. Wang and Y. P. Zhang, "Cell-free protein synthesis energized by slowly-metabolized maltodextrin," *BMC biotechnology*, vol. 9, no. 1, pp. 1–8, 2009.

[79] R. Kim and C. Choi, "Expression-independent consumption of substrates in cell-free expression system from escherichia coli," *Journal of biotechnology*, vol. 84, no. 1, pp. 27–32, 2000.

[80] D.-M. Kim and J. R. Swartz, "Prolonging cell-free protein synthesis with a novel atp regeneration system," *Biotechnology and Bioengineering*, vol. 66, no. 3, pp. 180–188, 1999.

[81] T.-W. Kim, H.-C. Kim, I.-S. Oh, and D.-M. Kim, "A highly efficient and economical cell-free protein synthesis system using the s12 extract of escherichia coli," *Biotechnology and Bioprocess Engineering*, vol. 13, no. 4, pp. 464–469, 2008.

[82] H.-C. Kim, T.-W. Kim, and D.-M. Kim, "Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source," *Process Biochemistry*, vol. 46, no. 6, pp. 1366–1369, 2011.

[83] A. C. Chiao, R. M. Murray, and Z. Z. Sun, "Development of prokaryotic cell-free systems for synthetic biology," *bioRxiv*, p. 048710, 2016.

[84] M. K. Takahashi, J. Chappell, C. A. Hayes, Z. Z. Sun, J. Kim, V. Singhal, K. J. Spring, S. Al-Khabouri, C. P. Fall, V. Noireaux, *et al.*, "Rapidly characterizing the fast dynamics of rna genetic circuitry with cell-free transcription-translation (tx-tl) systems," *ACS synthetic biology*, vol. 4, no. 5, pp. 503–515, 2015.

[85] M. V. Olson, "The human genome project.," *Proceedings of the National Academy of Sciences*, vol. 90, no. 10, pp. 4338–4344, 1993.

[86] P. P. Amaral and J. S. Mattick, "Noncoding rna in development," *Mammalian genome*, vol. 19, no. 7, pp. 454–492, 2008.

[87] R. C. Wilson and J. A. Doudna, "Molecular mechanisms of rna interference," *Annual review of biophysics*, vol. 42, pp. 217–239, 2013.

[88] F. Hille, H. Richter, S. P. Wong, M. Bratovič, S. Ressel, and E. Charpentier, "The biology of crispr-cas: backward and forward," *Cell*, vol. 172, no. 6, pp. 1239–1259, 2018.

[89] C. Zhang, "Novel functions for small rna molecules," *Current opinion in molecular therapeutics*, vol. 11, no. 6, p. 641, 2009.

[90] M. K. Thomason and G. Storz, "Bacterial antisense rnas: how many are there, and what are they doing?," *Annu Rev Genet*, vol. 44, pp. 167–88, 2010.

[91] D. H. Lenz, K. C. Mok, B. N. Lilley, R. V. Kulkarni, N. S. Wingreen, and B. L. Bassler, "The small rna chaperone hfq and multiple small rnas control quorum sensing in vibrio harveyi and vibrio cholerae," *Cell*, vol. 118, no. 1, pp. 69–82, 2004.

[92] E. Massé and S. Gottesman, "A small rna regulates the expression of genes involved in iron metabolism in escherichia coli," *Proceedings of the National Academy of Sciences*, vol. 99, no. 7, pp. 4620–4625, 2002.

[93] H. Wang, A. Iacoangeli, S. Popp, I. A. Muslimov, H. Imataka, N. Sonenberg, I. B. Lomakin, and H. Tiedge, "Dendritic bc1 rna: functional role in regulation of translation initiation," *Journal of Neuroscience*, vol. 22, no. 23, pp. 10232–10241, 2002.

[94] M. Beltran, I. Puig, C. Peña, J. M. García, A. B. Álvarez, R. Peña, F. Bonilla, and A. G. de Herreros, "A natural antisense transcript regulates zeb2/sip1 gene expression during snail1-induced epithelial–mesenchymal transition," *Genes & development*, vol. 22, no. 6, pp. 756–769, 2008.

[95] A. Castello, B. Fischer, K. Eichelbaum, R. Horos, B. M. Beckmann, C. Strein, N. E. Davey, D. T. Humphreys, T. Preiss, L. M. Steinmetz, *et al.*, "Insights into rna biology from an atlas of mammalian mrna-binding proteins," *Cell*, vol. 149, no. 6, pp. 1393–1406, 2012.

[96] A. Castello, M. W. Hentze, and T. Preiss, "Metabolic enzymes enjoying new partnerships as rna-binding proteins," *Trends in Endocrinology & Metabolism*, vol. 26, no. 12, pp. 746–757, 2015.

[97] E. Chu, D. M. Koeller, J. L. Casey, J. C. Drake, B. A. Chabner, P. C. Elwood, S. Zinn, and C. J. Allegra, "Autoregulation of human thymidylate synthase messenger rna translation by thymidylate synthase.," *Proceedings of the National Academy of Sciences*, vol. 88, no. 20, pp. 8977–8981, 1991.

[98] S. Dabo and E. F. Meurs, "dsrna-dependent protein kinase pkr and its role in stress, signaling and hcv infection," *Viruses*, vol. 4, no. 11, pp. 2598–2635, 2012.

[99] S. Mazurek, F. Hugo, K. Failing, and E. Eigenbrodt, "Studies on associations of glycolytic and glutaminolytic enzymes in mcf-7 cells: Role of p36," *Journal of cellular physiology*, vol. 167, no. 2, pp. 238–250, 1996.

[100] M. Ramanathan, D. F. Porter, and P. A. Khavari, "Methods to study rna-protein interactions," *Nature methods*, vol. 16, no. 3, pp. 225–234, 2019.

[101] W. Winkler, A. Nahvi, and R. R. Breaker, "Thiamine derivatives bind messenger rnas directly to regulate bacterial gene expression," *Nature*, vol. 419, no. 6910, pp. 952–956, 2002.

[102] W. C. Winkler, S. Cohen-Chalamish, and R. R. Breaker, "An mrna structure that controls gene expression by binding fmn," *Proceedings of the National Academy of Sciences*, vol. 99, no. 25, pp. 15908–15913, 2002.

[103] W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins, and R. R. Breaker, "Control of gene expression by a natural metabolite-responsive ribozyme," *Nature*, vol. 428, no. 6980, pp. 281–286, 2004.

[104] F. J. Isaacs, D. J. Dwyer, and J. J. Collins, "Rna synthetic biology," *Nature biotechnology*, vol. 24, no. 5, pp. 545–554, 2006.

[105] N. Sudarsan, J. K. Wickiser, S. Nakamura, M. S. Ebert, and R. R. Breaker, "An mrna structure in bacteria that controls gene expression by binding lysine," *Genes and development*, vol. 17, no. 21, pp. 2688–2697, 2003.

[106] M. Mandal, B. Boese, J. E. Barrick, W. C. Winkler, and R. R. Breaker, "Riboswitches control fundamental biochemical pathways in bacillus subtilis and other bacteria," *Cell*, vol. 113, no. 5, pp. 577–586, 2003.

[107] B. Suess, B. Fink, C. Berens, R. Stentz, and W. Hillen, "A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo," *Nucleic acids research*, vol. 32, no. 4, pp. 1610–1614, 2004.

[108] J. Li, A. A. Green, H. Yan, and C. Fan, "Engineering nucleic acid structures for programmable molecular circuitry and intracellular biocomputation," *Nature chemistry*, vol. 9, no. 11, pp. 1056–1067, 2017.

[109] F. J. Isaacs, D. J. Dwyer, C. Ding, D. D. Pervouchine, C. R. Cantor, and J. J. Collins, "Engineered riboregulators enable post-transcriptional control of gene expression," *Nature biotechnology*, vol. 22, no. 7, pp. 841–847, 2004.

[110] M. K. Takahashi and J. B. Lucks, "A modular strategy for engineering orthogonal chimeric rna transcription regulators," *Nucleic acids research*, vol. 41, no. 15, pp. 7577–7588, 2013.

[111] A. A. Green, P. A. Silver, J. J. Collins, and P. Yin, "Toehold switches: De-novo-designed regulators of gene expression," *Cell*, 2014.

[112] K. Pardee, A. A. Green, M. K. Takahashi, D. Braff, G. Lambert, J. W. Lee, T. Ferrante, D. Ma, N. Donghia, M. Fan, *et al.*, "Rapid, low-cost detection of zika virus using programmable biomolecular components," *Cell*, vol. 165, no. 5, pp. 1255–1266, 2016.

[113] J. Kim, Y. Zhou, P. D. Carlson, M. Teichmann, S. Chaudhary, F. C. Simmel, P. A. Silver, J. J. Collins, J. B. Lucks, P. Yin, *et al.*, "De novo-designed translation-repressing riboregulators for multi-input cellular logic," *Nature chemical biology*, vol. 15, no. 12, pp. 1173–1182, 2019.

[114] J. Chappell, M. K. Takahashi, and J. B. Lucks, "Creating small transcription activating rnas," *Nature chemical biology*, vol. 11, no. 3, pp. 214–220, 2015.

[115] J. Chappell, A. Westbrook, M. Verosloff, and J. B. Lucks, "Computational design of small transcription activating rnas for versatile and dynamic gene regulation," *Nature communications*, vol. 8, no. 1, pp. 1–12, 2017.

[116] K. Shearwin, B. Callen, and J. Egan, "Transcriptional interference – a crash course," *Trends in Genetics*, vol. 21, pp. 339–345, Jun 2005.

[117] A. J. Bird, M. Gordon, D. J. Eide, and D. R. Winge, "Repression of adh1 and adh3 during zinc deficiency by zap1-induced intergenic rna transcripts," *The EMBO journal*, vol. 25, no. 24, pp. 5726–5734, 2006.

[118] C. F. Hongay, P. L. Grisafi, T. Galitski, and G. R. Fink, "Antisense transcription controls cell fate in saccharomyces cerevisiae," *Cell*, vol. 127, no. 4, pp. 735–745, 2006.

[119] S. Petruk, Y. Sedkov, K. M. Riley, J. Hodgson, F. Schweisguth, S. Hirose, J. B. Jaynes, H. W. Brock, and A. Mazo, "Transcription of bxd noncoding rnas promoted by trithorax represses ubx in cis by transcriptional interference," *Cell*, vol. 127, no. 6, pp. 1209–1221, 2006.

[120] X. Hu, S. Eszterhas, N. Pallazzi, E. E. Bouhassira, J. Fields, O. Tanabe, S. A. Gerber, M. Bulger, J. D. Engel, M. Groudine, *et al.*, "Transcriptional interference among the murine β-like globin genes," *Blood*, vol. 109, no. 5, pp. 2210–2216, 2007.

[121] I. Abarrategui and M. S. Krangel, "Noncoding transcription controls downstream promoters to regulate t-cell receptor α recombination," *The EMBO journal*, vol. 26, no. 20, pp. 4380–4390, 2007.

[122] Y. Han, Y. B. Lin, W. An, J. Xu, H.-C. Yang, K. O'Connell, D. Dordai, J. D. Boeke, J. D. Siliciano, and R. F. Siliciano, "Orientation-dependent regulation of integrated hiv-1 expression by host gene transcriptional readthrough," *Cell host & microbe*, vol. 4, no. 2, pp. 134–146, 2008.

[123] C. N. Merrikh and H. Merrikh, "Gene inversion potentiates bacterial evolvability and virulence," *Nature communications*, vol. 9, no. 1, pp. 1–10, 2018.

[124] S. A. Hoffmann, S. M. Kruse, and K. M. Arndt, "Long-range transcriptional interference in e. coli used to construct a dual positive selection system for genetic switches," *Nucleic acids research*, vol. 44, no. 10, pp. e95–e95, 2016.

[125] A. A. Krylov, V. V. Shapovalova, E. A. Miticheva, M. S. Shupletsov, and S. V. Mashko, "Universal actuator for efficient silencing of escherichia coli genes based on convergent transcription resistant to rho-dependent termination," *ACS Synthetic Biology*, vol. 9, no. 7, pp. 1650–1664, 2020.

[126] X. Xu, X. Li, Y. Liu, Y. Zhu, J. Li, G. Du, J. Chen, R. Ledesma-Amaro, and L. Liu, "Pyruvate-responsive genetic circuits for dynamic control of central metabolism," *Nature Chemical Biology*, vol. 16, no. 11, pp. 1261–1268, 2020.

[127] B. P. Callen, K. E. Shearwin, and J. B. Egan, "Transcriptional interference between convergent promoters caused by elongation over the promoter," *Mol Cell*, vol. 14, pp. 647–56, Jun 2004.

[128] A. E. Bordoy, U. S. Varanasi, C. M. Courtney, and A. Chatterjee, "Transcriptional interference in convergent promoters as a means for tunable gene expression," *ACS Synthetic Biology*, vol. 5, pp. 1331–1341, Dec 2016.

[129] A. C. Palmer, A. Ahlgren-Berg, J. B. Egan, I. B. Dodd, and K. E. Shearwin, "Potent transcriptional interference by pausing of rna polymerases over a downstream promoter," *Molecular Cell*, vol. 34, pp. 545–555, Jun 2009.

[130] A. C. Palmer, J. B. Egan, and K. E. Shearwin, "Transcriptional interference by rna polymerase pausing and dislodgement of transcription factors," *Transcription*, vol. 2, no. 1, pp. 9–14, 2011.

[131] N. J. O'Connor, A. E. Bordoy, and A. Chatterjee, "Engineering transcriptional interference through rna polymerase processivity control," *ACS Synthetic Biology*, vol. 10, no. 4, pp. 737–748, 2021.

[132] K. Sneppen, I. B. Dodd, K. E. Shearwin, A. C. Palmer, R. A. Schubert, B. P. Callen, and J. B. Egan, "A mathematical model for transcriptional interference by rna polymerase traffic in escherichia coli," *Journal of molecular biology*, vol. 346, no. 2, pp. 399–409, 2005.

[133] N. Hao, S. Krishna, A. Ahlgren-Berg, E. E. Cutts, K. E. Shearwin, and I. B. Dodd, "Road rules for traffic on dna—systematic analysis of transcriptional roadblocking in vivo," *Nucleic acids research*, vol. 42, no. 14, pp. 8861–8872, 2014.

[134] E. Vitaly, T. Francine, A. R. Rahmouni, B. Sergei, and N. Evgeny, "Transcription through the roadblocks: the role of rna polymerase cooperation," *The EMBO Journal*, vol. 22, no. 18, pp. 4719–4727, 2003.

[135] J. A. Brophy and C. A. Voigt, "Antisense transcription as a tool to tune gene expression," *Molecular systems biology*, vol. 12, no. 1, p. 854, 2016.

[136] A. E. Bordoy, N. J. O'Connor, and A. Chatterjee, "Construction of two-input logic gates using transcriptional interference," *ACS synthetic biology*, vol. 8, no. 10, pp. 2428–2441, 2019.

[137] A. Wong, H. Wang, C. L. Poh, and R. I. Kitney, "Layering genetic circuits to build a single cell, bacterial half adder," *BMC biology*, vol. 13, no. 1, pp. 1–16, 2015.

[138] W. J. Kelly and K. R. Muske, "Optimal operation of high-pressure homogenization for intracellular product recovery," *Bioprocess and biosystems engineering*, vol. 27, no. 1, pp. 25–37, 2004.

[139] A. Espah Borujeni, A. S. Channarasappa, and H. M. Salis, "Translation rate is controlled by coupled trade-offs between site accessibility, selective rna unfolding and sliding at upstream standby sites," *Nucleic acids research*, vol. 42, no. 4, pp. 2646–2659, 2014.

[140] I. Hindennach and H. Jockusch, "Peptide mapping of phage qβ proteins using cell-free synthesis," *Virology*, vol. 60, no. 2, pp. 327–341, 1974.

[141] N. Edwards, S. Beeton, A. Bull, and J. Merchuk, "A novel device for the assessment of shear effects on suspended microbial cultures," *Applied microbiology and biotechnology*, vol. 30, no. 2, pp. 190–195, 1989.

[142] A. Nieß, J. Failmezger, M. Kuschel, M. Siemann-Herzberg, and R. Takors, "Experimentally validated model enables debottlenecking of in vitro protein synthesis and identifies a control shift under in vivo conditions," *ACS synthetic biology*, vol. 6, no. 10, pp. 1913–1921, 2017.

[143] M. C. Jewett, K. A. Calhoun, A. Voloshin, J. J. Wuu, and J. R. Swartz, "An integrated cell-free metabolic platform for protein production and synthetic biology," *Molecular systems biology*, vol. 4, no. 1, p. 220, 2008.

[144] E. Falgenhauer, "Characterization of toehold riboregulator switches in different gene expression systems," master's thesis, Technische Universität München, 2017.

[145] M. Rustad, A. Eastlund, R. Marshall, P. Jardine, and V. Noireaux, "Synthesis of infectious bacteriophages in an e. coli-based cell-free expression system," *JoVE (Journal of Visualized Experiments)*, no. 126, p. e56144, 2017.

[146] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, "Andromeda: a peptide search engine integrated into the maxquant environment," *Journal of proteome research*, vol. 10, no. 4, pp. 1794–1805, 2011.

[147] S. Tyanova, T. Temu, and J. Cox, "The maxquant computational platform for mass spectrometry-based shotgun proteomics," *Nature protocols*, vol. 11, no. 12, pp. 2301–2319, 2016.

[148] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[149] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using david bioinformatics resources," *Nature protocols*, vol. 4, no. 1, pp. 44–57, 2009.

[150] "Uniprot: the universal protein knowledgebase in 2021," *Nucleic acids research*, vol. 49, no. D1, pp. D480–D489, 2021.

[151] R. Lutz and H. Bujard, "Independent and tight regulation of transcriptional units in escherichia coli via the lacr/o, the tetr/o and arac/i1-i2 regulatory elements," *Nucleic acids research*, vol. 25, no. 6, pp. 1203–1210, 1997.

[152] B. K. Chan, S. T. Abedon, and C. Loc-Carrillo, "Phage cocktails and the future of phage therapy," *Future microbiology*, vol. 8, no. 6, pp. 769–783, 2013.

[153] F. W. Twort, "An investigation on the nature of ultra-microscopic viruses," *Acta Kravsi*, 1961.

[154] F. d'Herelle, "On an invisible microbe antagonistic to dysentery bacilli," *CR Acad Sci*, vol. 165, pp. 373–375, 1917.

[155] J. E. Somerville, L. Cassiano, B. Bainbridge, M. D. Cunningham, R. P. Darveau, *et al.*, "A novel escherichia coli lipid a mutant that produces an antiinflammatory lipopolysaccharide.," *The Journal of clinical investigation*, vol. 97, no. 2, pp. 359–365, 1996.

[156] K. C. Murphy, "The λ gam protein inhibits recbcd binding to dsdna ends," *Journal of molecular biology*, vol. 371, no. 1, pp. 19–24, 2007.

[157] iGEM team Munich 2018; https://2018.igem.org/Team:Munich, "Phactory." Accessed: 2022-05-27.

[158] K. Vogele and P. Grossman, "Invitris; http://invitris.org." Accessed: 2022-05-27.

[159] J. Zawada, B. Richter, E. Huang, E. Lodes, A. Shah, and J. Swartz, "High-density, defined media culture for the production of escherichia coli cell extracts," ACS Publications, 2003.

[160] K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey, and A. R. Ferre-D'Amare, "Structural basis for activity of highly efficient rna mimics of green fluorescent protein," *Nature structural & molecular biology*, vol. 21, no. 8, pp. 658–663, 2014.

[161] J. S. Paige, K. Y. Wu, and S. R. Jaffrey, "Rna mimics of green fluorescent protein," *Science*, vol. 333, no. 6042, pp. 642–646, 2011.

[162] J. R. Babendure, S. R. Adams, and R. Y. Tsien, "Aptamers switch on fluorescence of triphenylmethane dyes," *Journal of the American Chemical Society*, vol. 125, no. 48, pp. 14716–14717, 2003.

[163] D. S. Bindels, L. Haarbosch, L. Van Weeren, M. Postma, K. E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M. A. Hink, *et al.*, "mscarlet: a bright monomeric red fluorescent protein for cellular imaging," *Nature methods*, vol. 14, no. 1, pp. 53–56, 2017.

[164] G.-J. Kremers, J. Goedhart, E. B. van Munster, and T. W. Gadella, "Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and fret förster radius," *Biochemistry*, vol. 45, no. 21, pp. 6570–6580, 2006.

[165] B. P. Cormack, R. H. Valdivia, and S. Falkow, "Facs-optimized mutants of the green fluorescent protein (gfp)," *Gene*, vol. 173, no. 1, pp. 33–38, 1996.

[166] J. Goedhart, D. Von Stetten, M. Noirclerc-Savoye, M. Lelimousin, L. Joosen, M. A. Hink, L. Van Weeren, T. W. Gadella, and A. Royant, "Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%," *Nature communications*, vol. 3, no. 1, pp. 1–9, 2012.

[167] G. T. Hanson, T. B. McAnaney, E. S. Park, M. E. Rendell, D. K. Yarbrough, S. Chu, L. Xi, S. G. Boxer, M. H. Montrose, and S. J. Remington, "Green fluorescent protein variants as ratiometric dual emission ph sensors.," *Biochemistry*, vol. 41, no. 52, pp. 15477–15488, 2002.

[168] P. Puigbò, I. G. Bravo, and S. Garcia-Vallve, "Caical: a combined set of tools to assess codon usage adaptation," *Biology direct*, vol. 3, no. 1, pp. 1–8, 2008.

[169] iGEM12 BostonU; http://parts.igem.org/Part:BBa K783076, "Promoter strength j23111 and ptet." Accessed: 2012-10-03.

Acknowledgments

Wer schnell gehen will, geht allein.
Wer aber weit gehen will, geht mit anderen.
(afrikanisches Sprichwort)

Herzlich bedanken möchte ich mich bei meinem Doktorvater Prof. Dr. Friedrich C. Simmel, an dessen Lehrstuhl ich seit meiner Bachelorarbeit tätig war. Sein breit angelegtes Forschungfeld ermöglichte es mir im Laufe der Zeit Projekte aus den sehr unterschiedlichen Bereichen der DNA-Nanotechnologie und der synthetischen Biologie zu bearbeiten. Er räumte mir dabei stets einen großen Handlungsspielraum in der Projektplanung und -umsetzung ein, lieferte aber auch neue Ansätze und Ideen zu deren Weiterentwicklung.

Des Weiteren gilt mein besonderer Dank Dr. Andrea Mückl, mit der ich meine Projekte im Hinblick auf deren Publikation diskutieren konnte und die mich beim Schreiben der Manuskripte und beim Durchführen einiger Experimente auf dem Toeholdswitch-Projekt unterstützte.

Ganz besonders bedanken möchte ich mich auch bei meiner ehemaligen Werksstudentin Sophie von Schönberg für ihre wertvolle Hilfe bei der Herstellung von einigen Chargen Zellextrakt und bei der Durchführung von Experimenten zur Qualitätssicherung.

Großer Dank gilt auch Dr. Cheng Meng und Dr. Christina Ludwig für die erfolgreiche Kollaboration und die fruchtbaren Diskussionen. Chen Meng führte zusammen mit Quirin Emslander die Massenspektrometrie-Messungen für das Zellextrakt-Projekt am BayBioMS-Zentrum durch. Im Anschluss unterstützte er mich bei der Auswertung und Interpretation der gewonnenen Daten.

Ein herzliches Dankeschön gilt auch Dr. Kilian Vogele, der für das Zellextrakt-projekt Phagen assemblierte, und Dr. Matthaeus Schwarz-Schilling, der mit mir während meiner Masterarbeit den Grundstein für das Toeholdswitch-Projekt legte.

155

Danke sagen möchte ich auch Dr. Jonathan List, der mir seit meiner Bachelorarbeit mit Rat und Tat zur Seite steht, obwohl sich unsere Forschungsschwerpunkte weit voneinander entfernt haben.

Ebenfalls dankbar bin ich Florian Katzmeier für die Unterstützung bei der Anwendung von t-Tests und allen, die sich für das Korrekturlesen dieser Arbeit zur Verfügung gestellt haben.

Ich möchte mich auch bei allen anderen Mitgliedern des E14-Lehrstuhls für die gute Arbeitsatmosphäre und das ausgesprochen gute, freundschaftliche Verhältnis bedanken.

Auch meiner Familie und meinem Freundeskreis gilt Anerkennung dafür, dass sie mir eine gute Stütze waren, wenn es nach einem langen Arbeitstag neue Kräfte zu sammeln galt.