Engineering standardised and modular biological controllers for efficient design and easy implementation in synthetic genetic circuits

Ari Dwijayanti

Department of Life Sciences

Imperial College London

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

2019

Declaration

I hereby certify that all of material in this thesis is my own work, unless otherwise stated. All external contributions and any information derived from other sources has been appropriately acknowledged. The work or any part of the work has not been previously submitted anywhere for any degree or diploma.

Ari Dwijayanti

Abstract

Synthetic biology has the vision to re-design biology in an easier, faster, more robust, efficient, and reliable fashion by applying engineering principles to living systems. To this end, a modular design approach enables rapid prototyping and manufacturing of various synthetic genetic circuit libraries. This transformative approach needs expansion of interchangeable, standardised, and well-characterised genetic components required for composing higher-level functional circuits. Additionally, compatibility of biological parts into modular design assemblies is highly desirable to streamline fabrication of synthetic genetic circuits. This project is focused on the *in vivo* characterisation, standardisation, modularisation, and implementation of a set of biomolecular regulators in *Escherichia coli*, specifically at the transcriptional level through the standardised promoter architecture and the post-transcriptional level via modular Artificial RNA interference (mARi). These regulatory systems were rationally developed within a modular design and DNA assembly framework to facilitate their easy adoption and implementation. The regulatory properties of both controllers were further characterised towards a range of typical genetic and cellular contexts that are important for diverse applications. Additionally, extensibility and orthogonality of these controllers allow for multiplexed and simultaneous regulation of multi-gene systems with alternative configurations. As a demonstration, a standardised inducible promoter was employed to express stress-inducing recombinant proteins. Furthermore, the production of these proteins was improved up to 5-fold by the use of an adaptive and dynamic negative feedback system, which is governed by mARi and driven by the host-stress response. Ultimately, this improvement was robustly maintained in different tested perturbations. Owing to its modularity, this feedback system could potentially improve the production of any recombinant protein of interest without specifically tuning the system or requiring strain modification. Collectively, the genetic regulatory platforms presented in this project greatly provide valuable resources for developing the next-generation of engineered biological circuits.

Acknowledgements

First of all, I would like to thank Indonesia Endowment Funds for Education (LPDP) for the generous PhD scholarship. I thank my family for understanding, support, and prayer. I would thank Dr Geoff Baldwin and Prof Guy-Bart Stan for hosting me with the opportunity to work in the labs for the last four years, especially with the freedom to shape the project. I would also thank Prof Mark Isalan and Prof Tom Ellis as my progress review panels, and Dr Tom Gorochowski and Dr Karen Polizzi as my thesis examiners. I also am grateful for being supervised by a mentor such as Dr Marko Storch. I have to thank you for 'The Daddy and daughter's chat' as well as chocolates and sweets when needed. Apologies for taking a lot of your time and bother you as always. A special thanks goes to Dr Myriam Haltalli and Dr Ian Yunus for constructive discussion, great help, and emotional support during thesis writing and beyond. I am sure you would be great and successful PIs for your own research group one day. I also want to thank all current and former members of Baldwin, Stan, Jones, Lo Celso, Brady, Borkoulas lab and the Imperial College Centre for Synthetic Biology. Thank you all for so many fun moments, positive attitude, advice, help, and nice memories that we created. On behalf of my parents, I would send a big thanks to Amreen, Babar, and family who look after and help me to survive well in London. Lastly, to all residents of Wisma Siswa Merdeka, thank you for sharing the everlasting story, inspiration, courage, and authentic Indonesian cuisine.

Table of contents

List of Figures

List of Tables

List of Abbreviations

3OC6-HSL	3-oxohexanoyl-homoserine lactone
3OHC14-AHL	3-hydroxytetradecanoyl-homoserine lactone
aa	amino acid
Ab^R	Antibiotic resistance
Abs600	Absorbance at 600 nm
Ara	L-arabinose
aTc	anhydrotetracycline
au	arbitrary unit
BASIC	Biopart Assembly Standard for Idempotent Cloning
BFP	Blue Fluorescent Protein
BioFoundry	Biological foundry
Biopart	Biological part
bp	base pair
cDNA	complementary DNA
CDS	Coding sequence
CIDAR	Cross-disciplinary Integration of Design Automation Research lab
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRi	CRISPR interference
Cuma	Cuminic acid
DAPG	2,4-diacetylphophloroglucinol
DBTL	Design-Build-Test-Learn
DNA	deoxyribonucleic acid
DOE	Design of Experiment
GOI	Gene of Interest
gRNA	guide RNA
Hfq	Host factor-1
iGEM	international Genetically Engineered Machine
IPTG	isopropyl β-d-1-thiogalactopyranoside
*i*P	BASIC prefix
*i*S	BASIC suffix
kb	kilo base pair
kDa	kilo Dalton
LB	Luria-Bertani
mARi	modular Artificial RNA interference
mCherry	Monomeric Cherry Fluorescent Protein
MoClo	Molecular Cloning
MoPET	Modular Protein Expression Toolbox
mRNA	Messenger RNA
NOMAD	Nucleic Acid Ordered Assembly with Directionality
OLMA	Oligo-linker Mediated Assembly
ORF	Open Reading Frame
ori	origin of replication
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
pH	Power of hydrogen
PODAC	Protected Oligonucleotide Duplex Assisted Cloning
PNPase	polynucleotide phosphorylase

QGA	Quick Gene Assembly
RACE	Rapid Amplification of Complementary DNA End
RBS	Ribosome Binding Site
RNA	ribonucleic acid
RNAi	RNA interference
RNAP	RNA polymerase
RNase E	Ribonuclease E
SDS-PAGE	Sodium Dodecyl Sulphate-Poly Acrylamide Gel Electrophoresis
RppH	RNA pyrophosphohydrolase
SD	Standard Deviation
SEVA	Standard European Vectors Architecture
sfGFP	Superfolder Green Fluorescent Protein
sRNA	Small non-coding RNA
STAR	Small Transcription Activating RNA
sTSRV	satellite RNA of tobacco ringspot virus
TALE	Transcription-Activator-Like Effector
TetR	Tet repressor
TF	Transcription factor
TIR	Translation Initiation Region
tRNA	transfer RNA
TSS	Transcription Start Site
TU	Transcriptional unit
UTR	Un-translated region
Van	vanillic acid

Chapter 1

Introduction

1.1 *Escherichia coli* as a programmable chassis for synthetic biology purposes

Living cells are able to automatically sense, compute, and respond to various changing conditions such as nutrient starvation, pH, and osmotic stress[1–5]. This adaptive strategy is enabled by coordinated stochastic interactions of cellular components and regulatory networks. Beyond their complexity, living cells can be viewed as a programmable chassis. This approach has inspired the creation of genetic circuits to perform novel functionalities that do not exist in nature. The ability to rationally re-design biological systems by implementing engineering principles can be defined as synthetic biology. A synthetic biology approach relies on the bottom-up construction of synthetic genetic circuits that function in host cells.

Here, *Escherichia coli* is chosen as a model organism of interest. *E. coli* is a Gram-negative and facultative anaerobic bacterium. As a model organism, its genomic deoxyribonucleic acid (DNA) has been completely sequenced[6,7] and its physiology has been extensively studied[5,8–10]. Due to its rapid growth rate and simple cultivation procedure, *E. coli* has been widely used for many biotechnological applications[11,12]. The genetic tools and molecular biology techniques are also well characterised, enabling us to manipulate its genetic contents. Indeed, *E. coli* has been widely used as a workhorse for industrially-relevant biochemicals including amino acid, organic acids, and biofuels[13]. In addition, nearly 30% of approved recombinant therapeutic proteins are produced in *E. coli*[11,14]. Beyond that, *E. coli* has also been engineered and used for advanced applications, such as cellular diagnostics[15,16], smart therapeutics[17], living materials[18], and complex logic computations[19,20]. Altogether, this makes *E. coli* an attractive programmable chassis for genetic engineering and synthetic biology purposes.

1.2 The central dogma in a biological system

As a self-replicating unit, cells pass on the genetic materials to their progeny. All information that dictates biological functions in living cells is stored in the DNA. Inside an organism (*in vivo*), the DNA sequence can be self-replicated during a replication process catalysed by DNA Polymerases. To express a protein, the DNA sequence that encodes a Gene of Interest (GOI)

has to undergo several sequential processes (e.g. transcription and translation) (Figure 1.1). These sequential processes are generally known as the central dogma in the biological system.

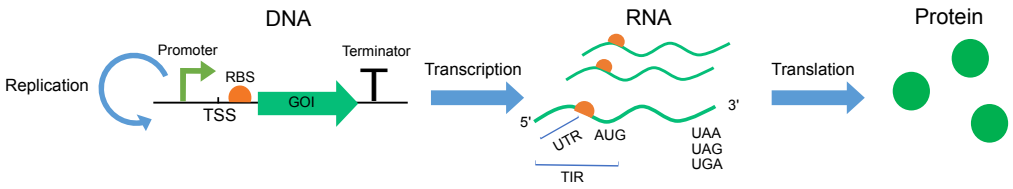


Figure 1.1 The central dogma in a biological system that transforms the information from DNA through RNA to protein. These processes include DNA replication, transcription, and translation. The structure of a transcriptional unit is shown with a promoter, RBS, GOI, and terminator. The structure of an mRNA is shown with UTR, TIR, start codon (AUG), and stop codons (UAA, UGA, and UAG).

The first step, transcription, is a process where the stored information from a DNA sequence is transformed into a ribonucleic acid (RNA) sequence. In this step, a specific region of the DNA sequence called the promoter is recognised by RNA Polymerase (RNAP). The binding of RNAP onto the promoter region initiates the transcription of functional RNA molecules, i.e. messenger RNA (mRNA) and small non-coding RNA (sRNA). The synthesis of RNA by RNAP typically starts from the Transcription Start Site (TSS). The transcription process is attenuated by a terminator. Based on its molecular mechanism, the termination of transcription in bacteria can be classified into intrinsic and Rho-dependent termination[21]. An intrinsic terminator relies on a GC-rich stem-loop structure in the DNA strand to release the transcription complexes[22]. By contrast, the Rho-dependent termination is facilitated by an RNA helicase, namely Rho to dissociate the transcription complexes[23].

The second step is translation, a process which converts mRNA into protein. This process is facilitated by the presence of a short Shine-Dalgarno of Ribosomal Binding Site (RBS) which is part of the 5’ Untranslated Region (UTR) within mRNA sequences. The RBS sequence is recognised by a ribosome, a highly complex molecular machine for protein synthesis. Once the ribosome binds to the RBS, it moves along the Translational Initiation Region (TIR) to scan the start codon located downstream of the RBS. To synthesise a protein, the ribosome reads the triplets of encoded mRNA sequence (codon) and pairs it with the associated cognate amino acid from the transfer RNA. In prokaryotes, protein synthesis predominantly begins at the start codon (AUG) and halts at the stop codon (UAA, UGA, and UAG triplets). The resulting

protein undergoes a complex folding and maturation process that is crucial for its biological functions.

The same central dogma processes and host machinery (i.e. DNA Polymerase, RNA Polymerase, ribosome, and transfer RNA (tRNA)) are typically shared between gene expression from the host cell and any synthetic constructs. As such, the development of genetic components to represent biological functions is based on those in natural systems such as promoter, RBS, and terminator. This smallest basic genetic unit is known as a biological part (biopart)[24]. A set of bioparts can be assembled to create a higher level function such as modules and complex circuits[24–26]. This has paved the way to engineer desirable biological functions by introducing synthetic circuits with a specific design whilst utilising the existing regulatory machinery from the host cell.

1.3 Biological design exploration through an iterative and integrated Design-Build-Test-Learn (DBTL) cycle

Like in many engineering frameworks, the creation of diverse synthetic biological circuits is done using a Design-Build-Test-Learn (DBTL) cycle[27–29] (Figure 1.2). Through this iterative process, one can create novel genetic circuits using a modular circuit construction from the standardised and modular biopart repositories.

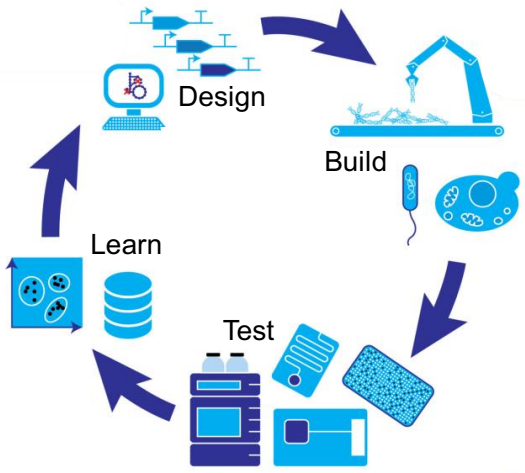


Figure 1.2 Illustration of typical iterative Design-Build-Test-Learn (DBTL) cycle. The figure was adapted from Jessop-Fabre & Sonnenschein, 2019.

This engineering cycle typically starts at the Design step, in which conceptualisation of genetic circuits is converted into biopart list and DNA sequences. The bioparts can be sourced from pre-characterised libraries or generated through DNA *de novo* synthesis with additional consideration of chassis selection and DNA assembly methods[28]. Typically, these assembly strategies utilise a set of standardised bioparts that can be stored, shared, and reused for the design and construction of many synthetic circuits[26,30]. The selection of regulatory elements plays a crucial role to balance gene expression and introduce genetic modulation for design space exploratory purposes. To guide the selection of bioparts and predict their performance, numerous computational tools are available such as RBS calculator v2.0[31,32], EMOPEC[33], R2oDNA Designer[34], and NUPACK[35]. Further, the automated design of genetic circuits for cellular computation has been exemplified using Cello, a design environment to create the circuit diagrams, select logic gates, and simulate their performance[36].

The physical construction of the designed construct is performed in the Build step. In earlier efforts, the *ad hoc* construction of synthetic genetic circuits was achieved using a conventional approach based on a combination of multiple cloning sites and bespoke ligation steps. Although this method has been routinely used in basic molecular cloning, it is a laborious process and lacks standardisation and modularity for further iterations. Therefore, various innovations in DNA assembly methods have been developed (further discussed in Section 1.4) and implemented in the Build step. The development of robust and scalable workflows for DNA assembly has significantly increased the capability and capacity in prototyping of synthetic circuits. As a result, various designs can be generated in parallel and much quicker than conventional approaches. At the end of the Build step, the assembled circuits are typically inserted and expressed into the organism of interest.

The performance of the circuits is examined in the Test step. This step involves verification at different levels including genomic, transcriptomic, proteomic, and metabolomics[27]. Although the Test step remains a rate-limiting step in the DBTL cycle, the current development of advanced analytical assays has tried to potentially increase the throughput and shorten turn-around times.

The data produced from the Test step are then processed and analysed in the Learn step. The output of the Learn step is crucial to realise data-driven design in the next iterative process. Despite this, these exhaustive processes are mainly performed using manual analysis. To efficiently handle a large set of data, an automated workflow assisted by a computer algorithm is urgently required. For instance, a multivariate approach using the Design of Experiment (DOE) has been implemented and integrated into a DBTL cycle, helping with the discovery and optimisation of biosynthetic pathways[37,38]. Furthermore, a robot scientist defined as artificial intelligent coupled to a robotic system has successfully guided biological experimentation in biosynthesis of aromatic amino acid[39].

To enable high experimental throughput, extract meaningful outcomes, and refine a design in a short period of time and cost-effective manner, the DBTL framework also needs to be supported by a standardised workflow and scalable technology platform. The integration of the DBTL pipeline in a biological foundry (BioFoundry) environment has become important to automate these processes along with a reduction in time, labour, and overall cost[27,37,40–42]. Moreover, the integration of DBTL with automation is also envisioned to improve reproducibility[43,44].

1.4 Standardised and modular DNA assembly accelerates prototyping of synthetic circuits

Since the cost of DNA synthesis has continued to fall, it is likely we will be able to directly synthesis long and complex genetic circuits in the near future. However, the genetic components within this construction are typically non-interchangeable thus cannot be reused in other constructions. This may impede the realisation of rapid and efficient prototyping of synthetic circuits, where a set of exchangeable bioparts can be reused in various design combinations and complexities[45]. To overcome this issue, methods in physical DNA assembly have been a fundamental enabling technology to practically implement modular circuit construction from a collection of well-characterised, standardised, and modular bioparts. A number of DNA assembly methods have been recently developed. A summary of innovative DNA assembly methods is listed in Table 1.1.

Table 1.1 Selected standardised and modular DNA assembly methods.

Physical standard	Forbidden Restriction Sites	Repository of standardised bioparts	Multi-part assembly	Hierarchical assembly	Assembly scar	Efficiency[a]	Accuracy[b]	BioFoundry adoption
NOMAD[46]	*Bsa*I, *Bsm*BI, *Bsm*AI	No	Yes	Yes	4 bp between parts	n.d.	n.d.	n.d.
BioBrick[47]	*Xba*I, *Spe*I, *Pst*I, *EcoR*I	Yes (iGEM distribution)	No	Yes	6-8 bp	n.d.	n.d.	n.d.
QGA[48]	*Xba*I, *Spe*I, *Pst*I, *EcoR*I	Yes (iGEM distribution)	No	Yes	6 bp between parts	n.d.	90-92%	n.d.
ePathBrick[49]	*AvrII, Xba*I, *Spe*I, *Nhe*I	Yes (iGEM distribution)	Yes	Yes	6-8 bp	n.d.	>70%	n.d.
Golden Gate[50,51]	*Bsa*I	No	Yes	Yes	4 bp between parts	7,320-19,683 colonies	>90%	Yes[52–54]
MoClo[55]	*Bsa*I, *BpiI, Esp3*I	No	Yes	Yes	4 bp between parts	150-33,000 colonies	90-99.8%	n.d.
MoPET[56]	*Bsa*I, *Bpi*I, *Sap*I	No	Yes	Yes	4 bp between parts	~400 colonies	95-100%	n.d.
CIDAR MoClo[57]	*Bsa*I, *Bbs*I	Yes (Addgene collection)	Yes	Yes	4 bp between parts	n.d.	>95%	n.d.
EcoFlex[58,59]	*Bsa*I, *Bsm*BI, *Bpi*I	Yes (Addgene collection)	Yes	Yes	4 bp between parts	n.d.	>99%	Yes[60]
MetClo[61]	*Bsa*I, *Bpi*I, *Lgu*I	Yes (Addgene collection)	Yes	Yes	n.d.	n.d.	n.d.	n.d.
BASIC[62,63]	*Bsa*I	No	Yes	Yes	4 bp (upstream) and 6 bp (downstream)	40-10,000 colonies	50-100%	Yes[38]
OLMA[64]	*Bsa*I	No	Yes	n.d.	Scar-less	4 parts assembly, 500 cfu/ug DNA	10-99.9%	n.d.
PODAC[65]	*Bsa*I	No	No	Yes	8 bp between parts	>24,500 colonies	>98%	n.d.
Start-stop assembly[66]	*Sap*I, *Bsa*I, *Bbs*I	Yes (Addgene collection)	Yes	Yes	Scar-less at CDS boundaries	n.d.	n.d.	n.d.

NOMAD, Nucleic Acid Ordered Assembly with Directionality; QGA, Quick Gene Assembly; MoClo, Molecular Cloning; MoPET, Modular Protein Expression Toolbox; CIDAR, Cross-disciplinary Integration of Design Automation Research lab; BASIC, Biopart Assembly Standard for Idempotent Cloning; OLMA, Oligo-linker Mediated Assembly; PODAC, Protected Oligonucleotide Duplex Assisted Cloning; [a]The number of colonies; [b]Percentage of the correct colonies; n.d., not defined

Historically, the concept of standardised and modular DNA assembly was first introduced over 20 years ago and termed as Nucleic Acid Ordered Assembly with Directionality (NOMAD)[46]. NOMAD facilitates the fully modular arrangement of DNA fragments by using type IIs restriction endonuclease. The type IIs restriction endonuclease cuts the DNA a few bases away from its recognition site and produces user-defined overhangs. This modular DNA design technique removes the limitation in the traditional cloning that mainly uses digestion and ligation strategy with fixed recognition sequences. Additionally, this initial effort was intended to be easily adopted through accessible modules, vectors, and databases.

To further progress the propagation of reusable and standardised physical components in the DNA assembly, a technical standard called BioBrick assembly was developed[47]. The BioBrick assembly employs standard flanking sequences with four restriction sites. The BioBrick assembly has been widely promoted and disseminated through the international Genetically Engineered Machine (iGEM) competition (www.igem.org). This deployment benefits from a global contribution from the synthetic biology community, especially with the expansion, characterisation, refinement, and submission of bioparts to the Registry of Standard Biological Parts (parts.igem.org). The collection of bioparts in the iGEM distribution has led to various derivatives of the BioBrick standard with improved accuracy such as Quick Gene Assembly (QGA)[48] and ePathBrick[49].

As the construction of synthetic circuits moves towards more complex and diverse designs, parallel assembly of multiple bioparts in a one-pot reaction is highly desirable. This has been demonstrated by Golden-Gate assembly[50,51] and its derivatives including Molecular Cloning (MoClo)[55], Modular Protein Expression Toolbox (MoPET)[56], Cross-disciplinary Integration of Design Automation Research lab (CIDAR) MoClo[57], EcoFlex[58,59], MetClo[61], Biopart Assembly Standard for Idempotent Cloning (BASIC)[62,63], Oligo-linker Mediated Assembly (OLMA)[64] and Start-stop assembly[66]. These assembly standards use type IIs restriction enzymes that generate short and predefined sticky ends to determine the assembly order of final constructs[45].

While these DNA assembly methods provide modular, combinatorial, and scalable DNA assembly, ‘scar’ sequences between bioparts remain in the final construct. The BioBrick-

derived assembly methods result in 6-8 bp scar sequences in the final construct as a result of two fused BioBricks[47,49]. The Golden Gate-derived assembly typically leaves 4 bp scars when two bioparts are religated[50,51,55–57,59]. In some cases, these undesirable scar sequences affect the evolutionary stability of the engineered circuits[67]. To avoid these scars at the coding sequence boundaries, a modular and multipart DNA assembly method called Start-Stop assembly has been introduced[66]. This assembly standard employs type IIs endonuclease-based approach to generate 3 bp overhangs. These overhangs are designed as start and stop codons in the coding sequences. Therefore, it provides scar-less and modular assembly in the construction of a range of expression units, including monocistronic, operon system, and hybrid configurations.

Beyond these methods, the development of modular assembly heavily depends on the downstream applications. For instance, the ePathBrick assembly method as an extension of the BioBrick standard and is specifically developed for modular, combinatorial, and optimised multi-gene pathway engineering[49]. Whereas the EcoFlex[58,59], MetClo[61], and Protected Oligonucleotide Duplex Assisted Cloning (PODAC)[65] are derived from Golden Gate assembly[50,51] and are predominantly used in the construction of various synthetic circuits for metabolic engineering purposes. The recursive assembly of PODAC has also been demonstrated to create artificial guide RNA (gRNA) array.

These modern DNA assembly frameworks often rely on a library of interchangeable and standardised bioparts. These bioparts within standardised format are stored in an entry vector and can be reused for the other designs and constructions at will. Ideally, they are deposited into a public inventory platform such as iGEM repository (http://parts.igem.org) and Addgene collection (https://www.addgene.org). Therefore, the repository of standardised bioparts can be accessed and expanded alongside the adoption of standard assembly by the synthetic biology community.

Additionally, these assembly strategies have been proven to be simple, efficient, and reliable as reflected from the high number of colonies obtained (efficiency) and correct assembly event (accuracy). For those reasons, several BioFoundries have adopted these standardised

and modular DNA assembly methods to enable a high-throughput automation pipelines[38,52–54,60].

1.5 BASIC: a DNA linker-guided assembly for standardised and modular construction of synthetic genetic circuits

The modular assembly of synthetic circuits throughout this project is based on a high-throughput DNA assembly method, called BASIC DNA assembly[62]. The assembly process is accomplished by type IIs restriction endonucleases to enables parallel and multi-part assembly without a Polymerase Chain Reaction (PCR) amplification step[45,62].

The sequences of bioparts within the BASIC assembly are designed and synthesised by avoiding a forbidden restriction site for *Bsa*I[62]. To ensure their reusability, these bioparts are stored in an entry vector with a standardised format (Figure 1.3). In particular, the core structure of bioparts contain unique and short flanking prefix (*i*P) and suffix (*i*S) sequences that encoding a *Bsa*I recognition site. These flanking sequences enable excision of bioparts of interest from storage plasmids during a restriction-ligation step. This produces a DNA fragment with a predefined 4 bp overhang sequences that can be used as adhesive ends for further assembly steps.

Another key feature of BASIC DNA assembly is that the process is guided by a set of orthogonal DNA linkers[62]. These DNA linkers allow efficient assembly due to their long overlapping base-pairs (21 bp). Each linker is divided into two parts (prefix and suffix-linkers) that are separately attached to corresponding complementary overhangs. The prefix-linker attached to the *i*P sequence and the suffix linker added to the *i*S sequences (Figure 1.3). In this case, the use of DNA linkers of choice is independent of the type of bioparts. Additionally, all BASIC DNA linkers have been computationally designed and validated using R2oDNA Designer software to avoid the complex secondary structures during the assembly process[34]. They are also designed to be orthogonal with a reduced cross-reactivity towards other linkers, bioparts, and the host genome[34].

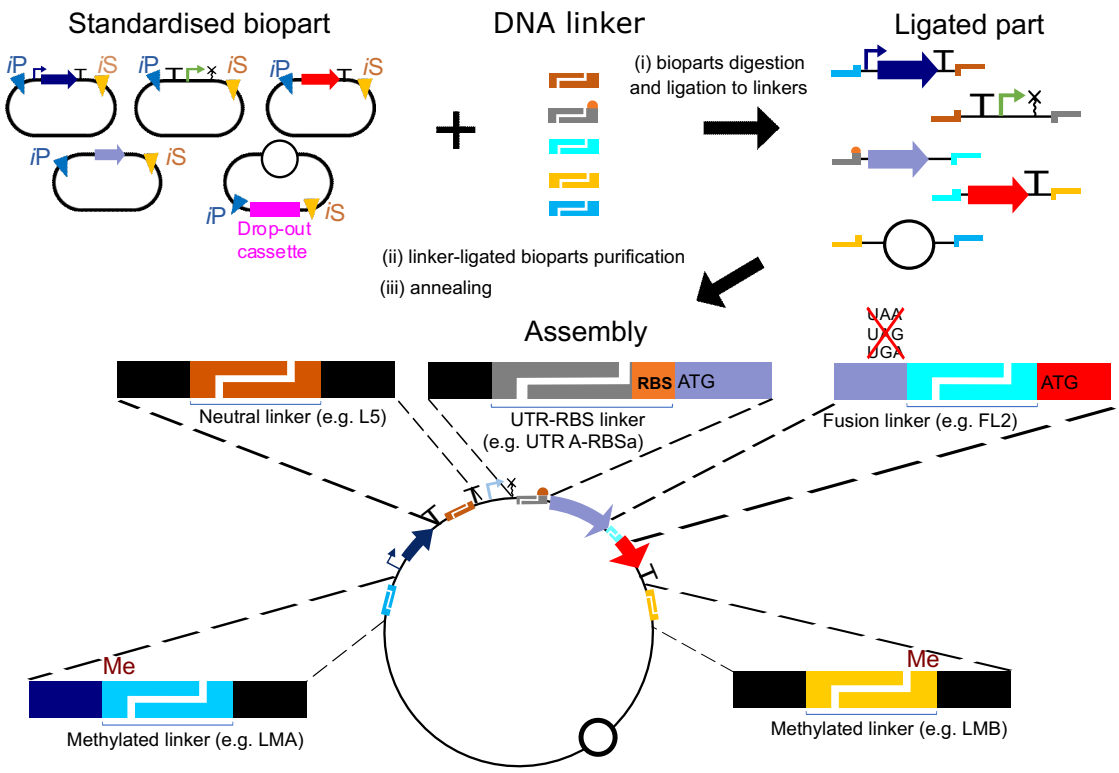


Figure 1.3 A schematic overview of the BASIC DNA assembly method. The standardised BASIC bioparts contains BASIC prefix (*i*P) and BASIC suffix (*i*S). This method involves three simple steps: (i) the digestion-ligation step of a set of standardised bioparts to particular BASIC linkers, (ii) the purification of linker-ligated bioparts, and (iii) the annealing of purified BASIC biopart-linkers to generate the final assemblies.; UTR, Untranslated Region; RBS, Ribosome Binding Site; ATG indicates start codon; UAA, UAG, and UGA indicate stop codons; Me, methylated.

To date, four types of BASIC orthogonal linkers are available to facilitate the construction of synthetic circuits. The first type is BASIC neutral linkers that are generally used to join two bioparts or expression cassettes. The second type is BASIC methylated linkers which contain 5' methylation in their *Bsa*I site to protect the linkers from *in vitro Bsa*I digestion during the assembly step. Therefore, the region flanked by these linkers can be reused for the further hierarchical assembly due to its idempotency. The third type is UTR-RBS BASIC linkers that contain untranslated region (UTR) and Ribosome Binding Site (RBS) commonly found in the 5' region of mRNA. The prefix and suffix linker pairs of the UTR-RBS linker can be substituted in flexible fashion to exchange UTR and RBS type independently, thus supporting easy tuning of gene expression. This unique feature also facilitates flexible combination of different promoters, UTRs, RBSs, and GOIs for modular gene expression. This type of linker is used to connect a promoter and a gene of interest (GOI) such as in a functional transcriptional unit. In addition to a single expression cassette, the UTR-RBS linker could also be employed to combine different cistrons in a multigene system such as an operon system. The last type of BASIC linker is BASIC fusion linkers used to connect two functional domains of a protein for

biotechnological applications. To create an *in-frame* domain arrangement, the stop codon (e.g. UAA, UGA, and UAG) at the end of the first GOI has to be excluded. The GOI without stop codon can be directly connected to the start codon (AUG) of the first GOI using a BASIC fusion linker. All types of BASIC DNA linkers are used throughout the construction of various synthetic circuits in this project.

To perform modular circuit construction, the BASIC DNA assembly process involves three main steps[62] (Figure 1.3). Firstly, bioparts digestion and ligation to orthogonal linkers. The *Bsa*I digestion and T4 ligation are simultaneously performed in a single-pot reaction. Secondly, purification of linker-ligated bioparts is performed to discard small un-ligated fragments. Finally, assembly is performed by an annealing process.

This assembly method yields over 90% accuracy for seven parts assemblies[62]. Such a strategy in one-pot reaction with high efficiency and accuracy is important for scalability of the DNA assembly protocol. BASIC DNA assembly is also compatible with automated workflows in a BioFoundry environment[38] and miniaturised laboratory automation using OT-2 liquid handler[68]. Therefore, BASIC assembly can provide a rapid prototyping pipeline for combinatorial, standardised, and modular construction of synthetic genetic circuits.

1.6 Library of molecular tools for engineering synthetic circuits

To accelerate the construction of synthetic genetic circuits, the DNA assembly methods have to be supported by the availability of diverse biomolecular toolboxes. The standardised and interchangeable bioparts such as plasmid backbones and genetic switches, at the transcriptional and post-transcriptional levels, used in the expression cassettes were the primary interest and will be further described.

1.6.1 Standardised plasmid backbone

Genetic circuits can be encoded in the chromosome of a chosen host or as extrachromosomal vectors. In general, genetic modification and integration into host-genome offers the benefit

of evolutionary stability, whereas genetic manipulation via an extrachromosomal vector (i.e. heterologous plasmid) provides flexibility and easy manipulation of the circuit design. Aiming to accelerate the DBTL cycle, the majority of the genetic modification throughout this project is based on plasmid-based systems.

A plasmid has the ability to self-replicate such that a host cell harbours multiple copies depending on the origin of replication (ori)[69,70]. An antibiotic resistance marker (Ab^R) acts as a selection for positive clones. In addition to these main elements, DNA segments containing specific functional modules can be inserted between the ori and Ab^R and are often termed as cargo modules. A number of accessories can also be added for distinct functionalities. In commercial plasmids, the existence of multiple cloning sites allows for the insertion of various GOIs through traditional cloning methods. Furthermore, protein tags are also often included to simplify downstream protein purification.

Despite their widespread use, concise and standardised plasmid backbones are not common, but needed to assist robust and predictable engineering of synthetic genetic circuits[71,72]. Standardised vector platforms have been developed by the synthetic biology community and disseminated through the iGEM registry (http://parts.igem.org). Typically, these plasmid backbones adopt the BioBrick standard[47]. Therefore, they are compatible with the BioBrick assembly[47,73].

Further effort in standardising plasmid backbones has also been initiated and documented through the Standard European Vector Architecture (SEVA) catalogue (http://seva.cnb.csic.es)[71,72]. SEVA plasmids (pSEVA) consists of a simple compositional structure with three main components: Ori, Ab^R, and cargo[71,72] (Figure 1.4A). To date, there are nine Oris, six Ab^Rs, and 13 cargos that can be combined and are available in the database[72]. Seven Oris from these standardised backbones have been quantified using cell sorting and droplet digital PCR, demonstrating a range of 2-40 plasmid copy numbers in *E. coli*[70]. The additional accessories in this standardised vector platform are two transcriptional terminators (T0 and T1) adjacent to the cargo, a minimum sequence for replication initiation (*oriV*), and the origin of conjugative transfer (*oriT*) to maintain the portability of pSEVA across different bacterial species[71].

Along with this standardised format, the sequences flanking cargo within pSEVA have been specifically modified to leverage its adoption. Previously, 20 combinations of rare restriction sites were embedded in these flanking sequences to improve their compatibility with conventional enzymatic restriction/ligation, and various modern DNA assembly methods including Gibson assembly, uracil excision cloning, and a nicking enzyme-based DNA assembly[74].

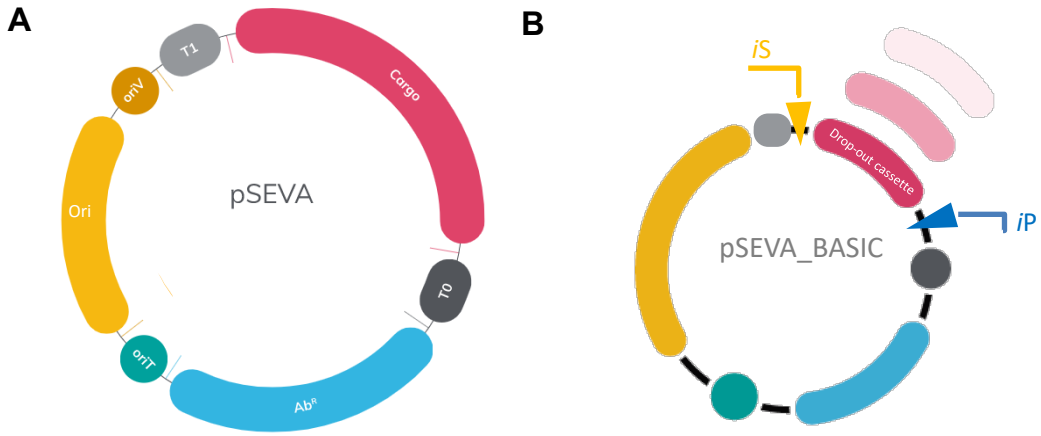


Figure 1.4 Illustration of pSEVA and pSEVA_BASIC. (A) The structure of pSEVA carrying an antibiotic resistance (Ab^R), the origin of replication (Ori), and cargo. The additional accessories: T0, T1, *oriT*, and *oriV* were identified. (B) The simplified structure of pSEVA_BASIC harbouring a drop-off cassette (e.g. mScarlett expression cassette) is shown. This drop-off cassette is flanked by BASIC prefix (*i*P) and suffix (*i*S) which can be released during the assembly process. Figures were modified from SEVA repositories (http://seva.cnb.csic.es/).

These standardised plasmid backbones have also been adopted for use with the BASIC DNA assembly method[62] by modifying the cargo and its flanking sequences (Storch *et al.*, manuscript in preparation). Specifically, a drop-out cassette (i.e. mScarlett expression cassette) was created as a selection marker by Dr Marko Storch and members of the Baldwin lab. This expression cassette is flanked by BASIC prefix and suffix sequences that can be excised during the BASIC digestion-ligation step (Section 1.5). Using this strategy, the negative clones can be easily distinguished by their visible red colour. The map of modified pSEVA backbone in BASIC format (pSEVA_BASIC) is illustrated in Figure 1.4B.

1.6.2 Transcriptional control using promoter

A promoter is a DNA sequence at the 5' region of an expression cassette that functions as a recognition site for RNA Polymerase (RNAP). In prokaryotes, the functionality of promoters is

governed by sequences at the -35 and -10 positions[75]. Promoters play a crucial role in balancing gene expression, especially in determining the timing and level of gene expression. Based on a promoter activity, it can be classified as either constitutive or inducible.

A constitutive promoter exhibits continuous transcription of a downstream GOI. The transcription strength of a constitutive promoter is determined by the sequences at -35 and -10 position, which dictate the binding affinity of RNAP[76]. Interestingly, many genes in their native systems employ constitutive promoter with intermediate expression[77]. In engineered systems, constitutive promoters are typically used to provide a static level of gene expression[78]. To enable the constitutive gene expression in the engineered circuits, a set of constitutive promoters with different strengths (i.e. Anderson promoter libraries) have been deposited in the iGEM registry (http://parts.igem.org) and extensively characterised[79].

In contrast to constitutive promoters, the inducible promoters are conditionally expressed under specific circumstances. The core of an inducible promoter has DNA operator sites facilitating the binding of factors, i.e. Transcription Factor (TF) or sigma (σ) factor. To dynamically control gene expression, a TF has to be present[80,81]. For many TFs, the binding and unbinding to the DNA operator sequences can be stimulated by addition of external inducers (i.e. chemical inducers) or internal changing conditions that lead to initiating the transcription process.

To date, a number of native inducible promoters have been identified and characterised in *E. coli*[82–85]. One of notable inducible systems in *E. coli* can be exemplified by the arabinose (Ara)-based inducible system. The Ara-responsive promoter (P_{BAD}) was extracted from araBAD operon[86]. This inducible promoter is controlled by regulatory protein AraC and the presence of arabinose as a chemical inducer. AraC is naturally present in *E. coli* MG1655 and BL21(DE3) strains to catabolise arabinose[80]. Additionally, the activation of P_{BAD} is reduced when the engineered strains are cultured in a glucose-containing medium[81,86]. To avoid this effect, the AraC/P_{BAD} system may be cultured in growth medium with other carbon sources, such as glycerol[81].

To expand the variant of inducible promoters, these promoter sets can be engineered and tuned by modifying the sequences of native inducible promoters. The initial work to engineer inducible promoters started over 30 years ago by de Boer and colleagues[87]. A hybrid promoter, namely Tac promoter (P_{Tac}), was derived from the native inducible promoters, i.e. *trp* and *lac* UV5 promoters. This promoter can be controlled by the presence of Lac repressor and turned on by adding isopropyl β-d-1-thiogalactopyranoside (IPTG) to the growth medium due to the presence of a lac operator site.

As an alternative to IPTG-inducible promoter, a tightly regulated promoter by Tet repressor (TetR) was developed[88]. Since the TetR is not naturally encoded in the genotype of *E. coli* cells, the functionality of a Tet-based promoter (P_{Tet}) has been shown to be independent of host-strain background[80,88]. The P_{Tet} promoter produces a high level of gene expression when induced by anhydrotetracycline (aTc) that cannot be catabolised by *E. coli* wild type[80]. Owing to these regulatory properties, further exploration of Tet-based regulation has been attempted to expand a library of TetR-derived repression system with a reduced cross-interaction[89]. To date, 16 variants of TetR-derived regulation systems have been developed and characterised, including TetR and PhlF-based repression systems.

In addition to the previous examples of chemically inducible promoters, a stress-responsive promoter (P_{htpG1}) is also used in this work. P_{htpG1} is an endogenous promoter that is activated by the interaction of HtpG chaperone and heat-shock response σ-factor (σ^{32})[82,83]. In *E. coli*, this promoter is a part of native stress-responsive regulation towards the accumulation of misfolded proteins[83]. Recently, the up-regulation of P_{htpG1} has been identified upon over-expression of heterologous proteins[82].

1.6.3 Post-transcriptional control via RNA interference (RNAi) system

RNA interference (RNAi) system by small non-coding RNAs (sRNAs) has been identified as major post-transcriptional regulator in bacteria[90–92]. Typically, bacterial sRNAs are 50-250 bp in length and do not contain an open reading frame[93]. The RNAi system has a critical role in diverse native regulatory mechanisms, such as stress response, homeostasis, quorum-sensing, virulence system, and sugar metabolism[90,94–98]. This is supported by unique

regulatory properties of sRNAs including their short length, fast response, rapid signal propagation, and low-cost production[99–101].

Gene regulation via RNAi system is accomplished by a short base-pairing interaction of the complementary RNA sequences within the target site in the mRNA and the seed sequence in the sRNA. Based on the arrangement of sRNAs towards their targets, they can be grouped into *cis* and *trans*-acting sRNAs. The *cis*-acting sRNAs are encoded on the same strand with opposite direction to their target, and thus make complementary base-pairing to the mRNA they regulate. By contrast, *trans*-acting sRNAs are located separately in a distinct location and have limited complementarity to their target genes. Varying locations of the target binding sites within the mRNA target are hypothesised to result in different binding affinities that may lead to different silencing mechanisms and range of activities. The majority of target sites for *trans*-encoded sRNAs are located in the 5′ UTR and/or TIR[102–104]. Interestingly, a repression effect has also been observed when the target site is located at the junction between the coding region and the 3′ UTR[103]. One of possible reason is that the 5′ and 3′ UTR have a selective bias for the AU-rich region which may be favourable for the binding of an essential RNA chaperone called host factor-I (Hfq) protein [105]. However, some studies also found that the Open Reading Frame (ORF) of the mRNA target is an effective recognition site for RNAi system[106–108].

Despite this, *trans*-encoded sRNAs create an efficient gene regulation to the mRNA of their targets[102,107,109,110]. Most of the well-identified sRNAs in Gram-negative bacteria are *trans*-acting[93]. Interestingly, it has also been reported that *trans*-acting sRNAs have higher efficacy than *cis*-acting sRNAs in repressing gene expression[93]. One of the possible reasons is that the *trans*-acting sRNA has a longer lifetime than *cis*-acting sRNA in cytoplasm[93]. The *trans*-encoded sRNAs, therefore, have been widely applied as versatile genetically encoded controllers for metabolic engineering and synthetic biology purposes[27,40–42]. Based on these facts, this project focuses on the exploration and implementation of *trans*-acting sRNAs, especially for downregulating gene expression.

The majority of RNAi systems exhibit effective gene repression mediated by the presence of RNA chaperones such as an Hfq (host factor-I protein) and degradosome complex as highlighted in many studies[105,108,111–115]. The Hfq has many essential roles in RNA-RNA

interactions, especially for stabilising and protecting sRNAs from ribonuclease attack[93,116–118]; increasing the local concentration of sRNAs[111]; facilitating base pairing of the sRNAs to target mRNAs[91]; and destabilising the RNAi-mRNA complex by structural rearrangement[114,119]. The Hfq is highly conserved and abundant in a range of organisms, such as *E. coli*[120,121], *Salmonella typhimurium*[122], *Neisseria meningitidis*[123], *Anabaena*[124], *Synechocystis*[125], and *Yersinia enterocolitica*[126] due to its pivotal roles in stress-response regulation and virulence[127]. To maintain its cellular level during stress conditions, the Hfq is expressed under the control of heat-shock promoters, such as σ^{32} and σ^{70}-dependent promoters[128] and modulated by host growth phases[129–131]. Additionally, the accumulation of Hfq is known to be autoregulated by its synthesis[129].

In a native system, Hfq is known to recruit other helper proteins (degradosome complex), e.g. Ribonuclease E (RNase E)[105,108,111,114,117,118,132], polynucleotide phosphorylase (PNPase)[105,111], and RNA pyrophosphohydrolase (RppH)[105]. The degradosome complex is responsible for modulating sRNA-mRNA interactions. Specifically, RNase E has an essential role as an endonuclease that facilitates single-stranded RNAs cleavage and sRNA-mRNA degradation[111]. PNPase protects RNAs from RNAse E degradation by stabilising RNAs and regulating their accessibility from RNase E[111]. Finally, RppH provides monophosphate in the 5' UTR of RNAs that allows for RNase E-dependent degradation[105]. Moreover, these helper proteins are also highly conserved and abundantly available as a part of natural cellular stress response in a range of organisms[120,123–125,133,134].

In general, the *trans*-encoded sRNAs have a modular structure that consists of two main modules: a seed sequence and sRNA scaffold. The seed sequence has a function to guide the sRNAs towards the specific target, whereas the sRNA scaffold contains a recognition site for Hfq binding. This core structure is preceded by a promoter to drive the transcription of sRNA and followed by a transcriptional terminator (Figure 1.5A).

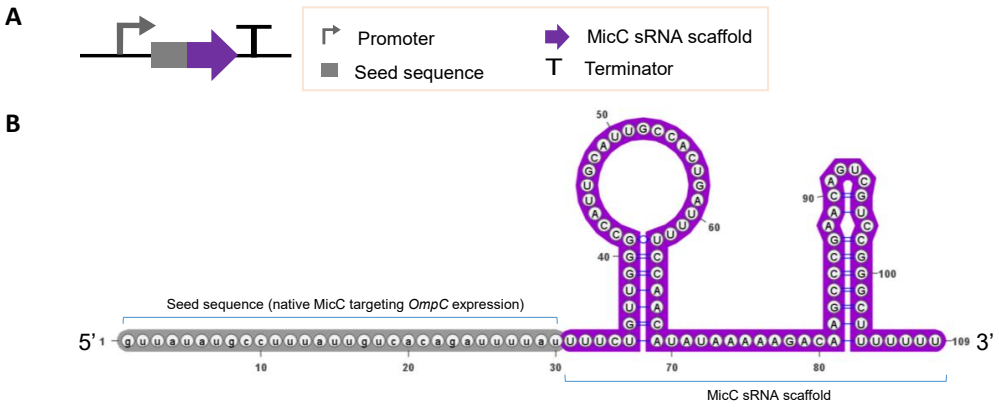


Figure 1.5 Illustration of MicC sRNA structure. (A) Schematic of a transcription unit of *Salmonella* MicC sRNA consisting of a promoter, seed sequence, MicC sRNA scaffold, and terminator. (B) Secondary structure of native MicC sRNA in *S. enterica* targeting OmpC. The secondary structure of MicC sRNA was adapted from Pfeiffer *et al*., 2009 with permission (see Appendix A).

Several studies have investigated the modular structure of the *Salmonella* MicC sRNAs in *E. coli* and have shown that these natural *trans*-encoded sRNAs differ in their seed sequences while using the same sRNA scaffold[108,135]. This sRNA regulation has been identified as a key regulator in the synthesis of major outer membrane proteins (e.g. OmpC[121] and OmpD[108]). MicC targeting OmpC has a 30 bp seed sequence (Figure 1.5B) whereas the seed sequence for OmpD regulation is only 12 bp[108]. For the OmpC, the target site is located in the TIR, whereas OmpD has a target site in the middle of coding sequence (CDS). These MicC sRNAs with different target sites employ different silencing mechanisms[108,135] (Figure 1.6). The first silencing mechanism is known as inhibition of translation initiation when target site is placed in the TIR (Figure 1.6A). The base-pairing interaction of MicC sRNA to the TIR of OmpC prevents the initiation of translation. This may further lead to rapid mRNA decay by RNase E degradation. The second silencing mechanism is by an endonucleolytic mRNA destabilisation leading to degradation of downstream mRNA (Figure 1.6B). In this mechanism, the MicC sRNA creates base-pairing in the middle of CDS of OmpD. This mechanism is perceived as RNase E-dependent mRNA decay without inhibition of translation initiation[108]. Interestingly, these two distinct mechanisms resulted in higher fold repression of MicC sRNA targeting OmpC than OmpD[108].

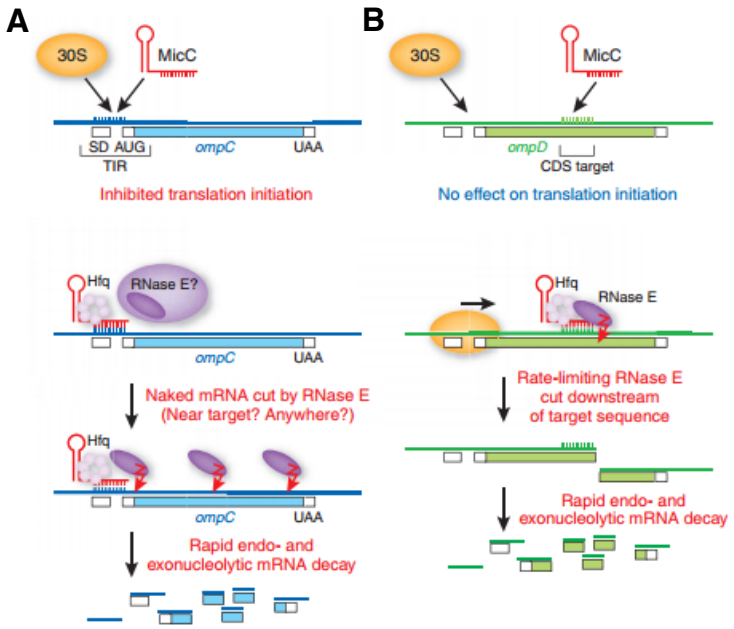


Figure 1.6 Two different silencing mechanisms of native MicC sRNA targeting endogenous outer membrane proteins (OmpC and OmpD), via (A) inhibition of translation initiation and (B) endonucleolytic mRNA destabilisation[135,136]. Figures were taken from Wagner, 2009 with permission (see Appendix A).

This modular architecture of bacterial sRNAs allows for easy exchange and modification of each constituent in an engineered RNAis as demonstrated in the previous studies[102,104,106,107,137]. Underpinned by the same principle to the MicC sRNAs, a library of artificial RNAis can be generated by substituting the seed sequence while retaining the native sRNA scaffold. With this strategy, the new variant of sRNAs can be redirected to specifically target different mRNAs of interest. This RNA-RNA interaction can be predicted *in silico* to minimise cross-interaction to non-cognate targets and its repression strength can be tuned based on its seed sequence[102,138,139]. Additionally, an advantage of these new variants of synthetic sRNA is that they can be recognised and modulated by existing RNA chaperones.

1.7 Interconnection of biomolecular controllers is essential to rewire complex regulatory circuits

The previously outlined genetic switches can be used to control gene expressions at a transcriptional level (Section 1.6.2) and a post-transcriptional level (Section 1.6.3). Owing to their unique natural modes of action and properties of regulation, these controllers are known to interact with each other in biomolecular level and create synergistic interactions[140–

[142]. Such interconnections between different types of biomolecular controllers are important, and exist in both natural systems and engineered genetic circuits (Figure 1.7).

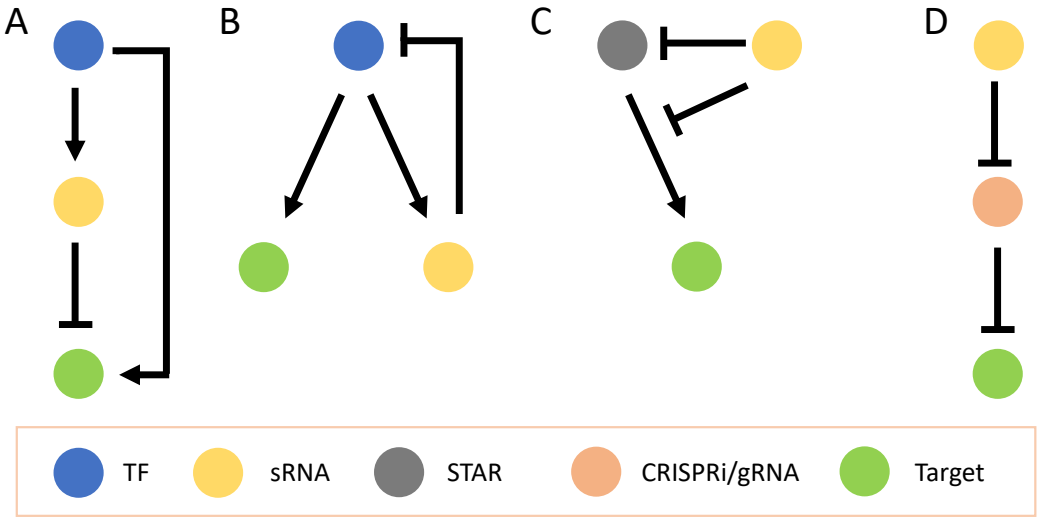


Figure 1.7 Illustration of interconnected biomolecular controllers. (A) The combination of the TF-RNAi systems in a natural system. (B) The combination of the TF-RNAi systems in an engineered circuit. (C) The combination of the RNAi-STAR systems in an engineered circuit. (D) The combination of the RNAi-CRISPRi/gRNA systems in an engineered circuit.

In nature, a commonly found biomolecular controller has the structure of a mixed feed-forward loop[143]. This mixed regulatory system is comprised of a three-node circuit and built using the transcriptional control via a TF and the post-transcriptional regulation via an RNAi system (Figure 1.7A). A common configuration of this feed-forward loop in *E. coli* comprises of a TF located at the top layer and RNAi system used in the lower layer. The TF can regulate the target through two different routes: direct and indirect via RNAi system. This double-layered regulation results in improved response and provides a higher degree of control by preventing transcriptional leakage and synchronising transcription and translation rates[143].

Likewise, a combination of TF and RNAi systems has also been demonstrated in a synthetic negative feedback loop[144,145] (Figure 1.7B). In this engineered circuit, the expression of a well-characterised TF actuates transcription of an engineered sRNA. This modified sRNA with a bespoke seed sequence specifically targets the TIR of the mRNA (i.e. TF), thus reducing the intracellular concentration of the TF. This synthetic feedback system is employed to fine-tune the circuit output and for adaptation of genetic modules towards changes in ribosome availability[144,145].

Another interconnection of biomolecular regulators in engineered genetic circuits is exemplified by the combination of an RNAi and Small Transcription Activating RNA (STAR) systems[146] (Figure 1.7C). In this programmable system, gene activation is achieved using the STAR, an RNA regulator that acts to trigger transcription of GOI[147–149]. Whereas, RNAi system is used to deactivate the gene expression at both transcriptional levels by sequestering STAR and post-transcriptional level through binding to an mRNA of interest. Importantly, this multilevel gene regulation shows reversible activation and de-activation which is an important feature for dynamic control of synthetic circuits.

In addition to this, multilevel gene regulation has also been demonstrated by integrating RNAi and Clustered Regularly Interspaced Short Palindromic Repeat interference (CRISPRi)/ gRNA systems[150] (Figure 1.7D). Strong repression of GOI is facilitated by the use of CRISPRi/gRNA system[151,152]. The binding of the dCas9-gRNA complex to its target promoter is modulated by sRNA that complements and sequesters gRNA. This derepression mechanism relies on the cleavage rate of the gRNA-RNAi complexes.

1.8 Coordinated feedback control interfacing synthetic circuits and host-responses

The introduction of a synthetic circuit into living cells often results in interactions between the synthetic construct and endogenous systems[145,153]. In some cases, these interactions can be harnessed to create coordinated feedback controls between the synthetic circuits and host-responses[154–156] (Figure 1.8).

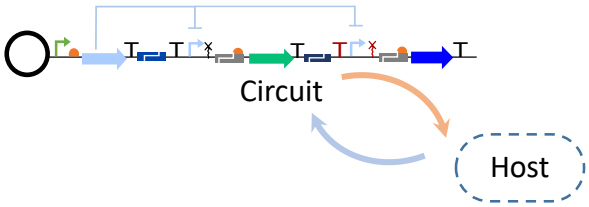


Figure 1.8 Illustration of dynamic interaction between synthetic circuits and host response.

One of the mutual interactions can be exemplified in the existence of an unnatural load due to over-expression of heterologous proteins against host-cell responses. It has been

experimentally validated that the over-expression of stress-inducing recombinant proteins from engineered constructs with less efficient design trigger the activation of host-stress responses[154,155,157]. The identification of stress-responsive promoters has led to the development of system employing negative feedback loops to control a circuit in response to host-stress[154,158,159]. As an output, this feedback control is used to automatically downregulate the expression of recombinant protein. For an actuator of this feedback system, Dragosits *et al*. have used transcriptional factor-based regulation (i.e. TetR) fused with a degradation tag to create transient repression[155]. More recently, Ceroni *et al*. used stress-inducible gRNA expression coupled with constitutive expression of dCas9 to automatically regulate heterologous protein production and maintain the fitness of host cells[154].

Similarly, physiological and environmental changes can be sensed and used as inputs to synthetic circuits to dynamically control endogenous metabolism, e.g. for acetate bioproduction[156]. In that study, three orthogonal sensors including glucose, oxygen, and acetate-responsive promoters are developed to continuously monitor and respond to multiple signals during bioproduction. These sensors are integrated to create combinatorial logic operations. The feedback loop of these logics is then implemented using a CRISPRi and targeted proteolysis-based gene down-regulation. The resulting synthetic regulatory controls automatically slow the accumulation of acetate at different growth phases depending on the activity of endogenous gene expression in the acetate pathway.

1.9 Challenges in up-scaling functional synthetic circuits

The ability to engineer biological systems has progressed beyond conventional genetic engineering by implementing key engineering principles in the design of biological systems. A combination of scalable and modular DNA assembly techniques, expansion of well-characterised bioparts, and forward design of synthetic circuits are envisioned to accelerate the prototyping process and potentially reduce the overall cost of producing synthetic circuits with desired functions. Despite this, several challenges remain that need further exploration to realise the creation of complex synthetic circuits in living cells.

Complex biological circuits require large number of bioparts. The diversity of bioparts, especially regulatory elements, available for tuning gene expression is one of particular importance. To date, a number of toolboxes to control gene expression have been developed including protein-based regulators of transcription[160–162], RNA-based regulators of transcription[147,148,151,152], and RNA-based regulators of post-transcription levels[104,106]. The selection of a particular regulatory mechanism in a circuit brings with its consequence in terms of evolutionary stability, e.g. long repetitive sequences are often mutated[163–165]. To this end, the development of smaller and more predictable regulators such as those RNA-based has gained interest[99,141,166]. This is not only because of their predictability but also due to the potential for implementing compact and efficient designs. Moreover, RNA-based regulators could be interconnected to other biomolecular controllers as discussed in Sections 1.7 and 1.8.

To further improve the versatility and usability of diverse regulatory elements in synthetic circuits, these regulators would ideally be supported by standardisation of their structure[28]. Compared to the expansion of the new types of biopart, progress in standardising bioparts is slower. Earlier efforts have focused on defining the physical composition of bioparts and their incorporation into the BioBrick assembly standard[24,47]. In addition to this, the reliability, tunability, orthogonality, and composability of bioparts have been added as important features for standardised bioparts[167]. However, much additional work is required to fill the demand for available standardised bioparts.

In addition to the development of standardised bioparts, their modularity is another key factor. Here, modularity is defined as the ability of discrete bioparts to be combined together (composability) and still function correctly. This can be leveraged by their integration into scalable DNA assembly methods. The development of interchangeable bioparts aims to provide reusable bioparts for different circuit design and consequently enable collaborative sharing bioparts between different labs (discussed in Sections 1.4, 1.5, and 1.6). Ideally, a collection of bioparts compatible with a specific DNA assembly method[28] are deposited into publicly available biological repositories for reuse, i.e. iGEM repository (http://parts.igem.org) or Addgene (https://www.addgene.org).

1.10 Objective and thesis outline

The main objective of this thesis is were to systematically develop and expand a set of standardised and modular genetic controllers for bacterial synthetic circuits. These biomolecular controllers were rationally designed and developed to be compatible with well-established modular design assembly frameworks, specifically the BASIC DNA assembly method[62]. This facilitates their easy adoption and enables the rapid construction of synthetic genetic circuits. These regulatory controllers are extensible so that variants of these regulators can be easily generated and employed in multi-gene systems. The regulatory properties and robustness of the regulators were characterised in a range of typical genetic and cellular contexts. As an exemplar biotechnology application, the controllers were used to create a modular expression system for recombinant protein production with improved and robust productivity. This was enabled by the use of dynamic negative feedback control that automatically optimises translational resource allocation for protein production and growth.

This thesis has been structured as follows. Chapter 1 introduces the background of the project including the driving motives of engineering the biological systems in the context of modular design and assembly frameworks and the challenges in developing next-generation synthetic biological circuits. Chapter 2 describes the construction of standardised transcriptional elements, specifically a series of constitutive and inducible bacterial promoters. In Chapter 3, a modular and orthogonal post-transcriptional regulator relying on a *trans*-encoding modular Artificial RNA interference (mARi) system is presented. In Chapter 4, the standardised transcriptional regulator and mARi-based controller are implemented to create an adaptive negative feedback system which can improve the productivity of stress-inducing protein expression. These result chapters synergistically support each other to address the main objectives of this project and are then comprehensively discussed in Chapter 5. Chapter 6 summarises the materials and methods that are used for computational designs, molecular cloning methods, *in vivo* assays, and data analysis in this project.

Chapter 2

Design and characterisation of a set of standardised bacterial promoters to enable modular gene expression

2.1 Introduction

As an initial step in standardising bioparts, the design of promoter transcriptional regulatory elements is investigated in this chapter. The promoter is one of the key regulatory elements in natural and engineered genetic circuits used to initiate transcription of GOIs. The activity of a promoter is governed by consensus sequences (-10 and -35 boxes) within the core promoter, which determine the binding affinity of RNAP. However, the performance of a promoter is often influenced by its interaction with local genetic contexts[36,168–172] (Figure 2.1A). Therefore, the predictability of gene expression driven by this regulatory element is an important and challenging topic for further exploration, especially in the rational design and engineering of synthetic genetic circuits.

Several strategies have been used to improve the predictability of gene expression by the use of promoter insulator sequences adjacent to core promoter sequence[168,171,172]. For instance, a flanking insulator sequence introduced upstream and downstream of the core promoter reduces unwanted contextual effects and improves predictability of gene expression[168]. Placing a 36-nucleotide insulator sequence in the upstream region of a promoter has also been shown to minimise the context-dependency of the promoter (for both constitutive and inducible promoters)[172].

To increase the predictability of gene expression, several strategies have been explored with a focus in standardising the sequence of mRNA transcripts. This is exemplified by the use of a bicistronic design, which includes a short standardised 5' UTR sequence in between double RBSs[171]. The first RBS drives the expression of a 16 aa leader peptide, whereas the second RBS is positioned within the coding sequence of this leader peptide. The stop codon of the short leader peptide is designed to overlap by 1 bp with the start codon of the second coding sequence. Within this architecture, ribosomes bind to the first RBS, synthesise the leader sequence and untangle mRNA structure in this short region. Subsequently, the binding of ribosomes to the second RBS reinitiates translation of the coding sequence. The presence of a bicistronic design in the downstream region of the core promoter has been demonstrated to improve the consistency of gene expression across different GOIs by up to 93%[171].

An alternative to this, a ribozyme-based insulator with auto-catalytic function has been used downstream of the core promoter sequence to remove variation in the mRNA sequence when expressed from different promoters with unknown TSS (denoted as +1)[36,169,173] (Figure 2.1B). Recently, the improved predictability of gene expression has also been achieved by using an RNA processing-platform (i.e. CRISPR-based RNA cleavage)[174].

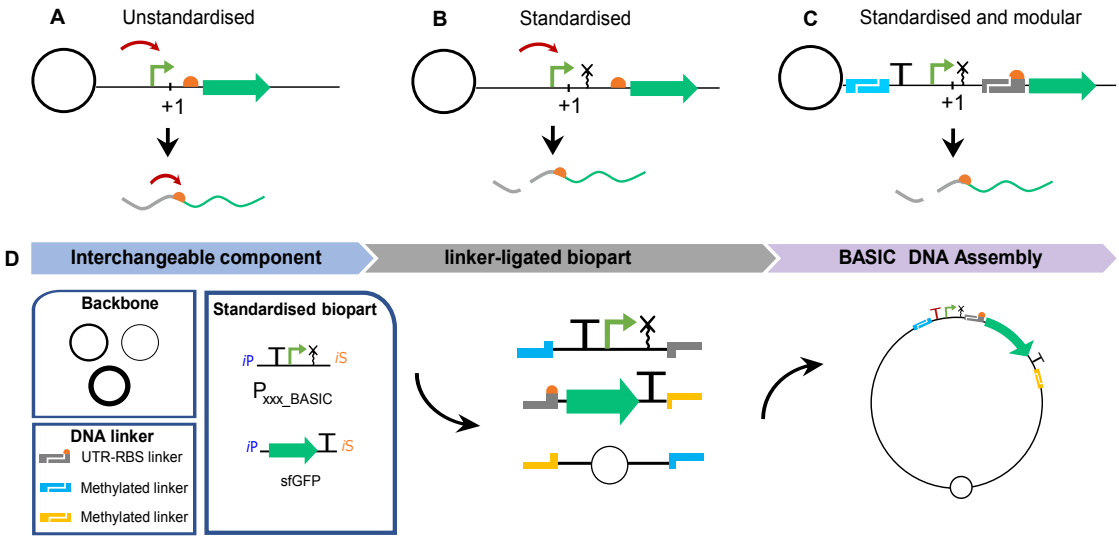


Figure 2.1 The general concept of standardised and modular bacterial promoters. (A) Unstandardised promoter with interaction of the upstream and downstream region of the core promoter that causes unpredictable gene expression. (B) The use of ribozyme insulator-based strategy to improve the predictability of gene expression. (C) The standardised and modular bacterial promoter allows predictable gene expression while providing an easy implementation in the modular construction of synthetic genetic circuits. (D) A schematic of standardised and modular promoter as an interchangeable biopart (P_{xxx_BASIC}) and its practical implementation in the BASIC DNA assembly. The Transcriptional start site is denoted as +1. The BASIC prefix (*i*P) and suffix (*i*S) are shown.

Despite the impressive progress that has been made, there is still a need to define the standardised structure of promoter element to use in modular assembly methods[175]. In this chapter, the design and characterisation of standardised bacterial promoters (Figures 2.1C) to enable modular construction of synthetic genetic circuits for gene expression is described.

To standardise the architecture of bacterial promoters, the key elements adjacent to a bacterial promoter were firstly defined, identified, and evaluated. Our standardised promoters were designed as BASIC bioparts (i.e. DNA parts with prefixes and suffixes at the 5' and 3' -boundaries) and integrated into the modular DNA BASIC assembly framework[62] to allow simple plasmid construction (Figures 2.1D). The superfolder green fluorescent protein (sfGFP) was used as a reporter to evaluate the performance of the standardised promoters.

Three selected promoter designs were then used to expand a library of bacterial constitutive and inducible promoters. For the standardised inducible promoters, their regulatory properties were tested with respect to different copy numbers of cognate transcription factor and target promoter. In addition to plasmid-borne transcription factors, the inducible promoter set was also characterised and implemented in host strains with genome-integrated transcription factors (marionette strains[80]). Finally, three standardised inducible promoters were employed for multiplexed multi-gene expression in the marionette strains.

2.2 Overview and selection of core elements in initial standardised promoter architecture

To investigate the requirements for a standardised promoter design, we defined a consistent promoter architecture composed of a set of defined modular components: an upstream terminator, an UP element between the terminator and the core promoter, and a ribozyme downstream of the core promoter (Figure 2.2).

The first main component is an upstream terminator that prevents transcriptional read-through from the previous gene by its stem-loop structure. For this initial assessment, a synthetic and well-characterised terminator L3S3P11[176] was used. In this case, L3S3P11 was modified to eliminate a forbidden *Bsa*I site by site-directed mutation of a nucleotide at position 45 from C to A[36]. This mutated terminator is referred to as L3S3P11*.

The terminator sequence was followed by a spacer sequence, non-coding DNA intended to act as a neutral sequence to reduce unwanted interaction between upstream sequences and the core promoter. Since this spacer is located upstream of the core promoter, it was termed an UP element. The UP element (Up A) with 45 bp length was generated using R2oDNA Designer[34].

A core promoter was next included as a recognition site for RNAP. Production of an mRNA from a promoter starts at the TSS. It has been shown that varying the promoter sequences changes the TSS and hence the 5’ UTR sequence produced[170]. Often this resulted in changes

in mRNA secondary structure that further affects complex interaction at the TIR and affects mRNA stability. Both of these factors have contributed to the unpredictability of gene expression to date[177].

To circumvent this issue, a self-cleaving ribozyme 75 bp long was introduced in the 3' region of the standardised promoter design. The ribozyme automatically cleaves the mRNA removing variable sequences at the 5'-end of mRNA arising from different TSSs[169]. We used natural ribozymes derived from a satellite RNA of tobacco ringspot virus (sTRSV)[178], called riboJ[169]. This self-cleaving structure is followed by a short hairpin (23 bp) that helps to expose the RBS sequence[179]. Its capability to insulate gene expression from the identity of the driving promoter has been shown in *E. coli*[36,169,180]. Furthermore, utilisation of RiboJ can increase the output activity of a promoter[180]. This strategy was chosen to simplify the standardisation of the mRNA 5' UTR without the need to express additional burdensome exogenous proteins like in the use of CRISPR-based insulation[174].

The performance of our standardised promoter design was monitored by measuring the expression of a fluorescent reporter gene. A test plasmid was built with a *sfgfp* reporter gene preceded by a UTR-RBS linker, e.g. UTR A-RBSc which encodes a medium strength RBS (Figure 2.2B). The tested promoters (P_{xxx_BASIC}) were assembled into a simple expression cassette using BASIC DNA assembly[62].

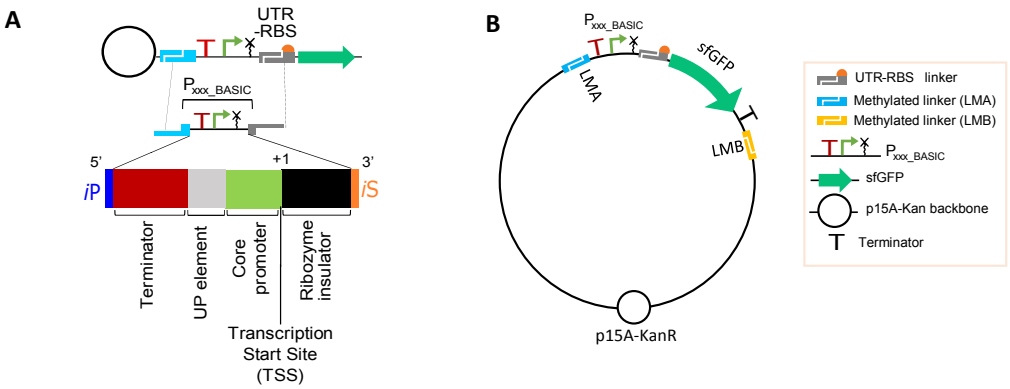


Figure 2.2 Overview of a standardised promoter architecture and testing plasmid. (A) Schematic of key components in standardised promoter design: terminator is shown in red, UP-element is shown in light grey, core promoter is shown in green, ribozyme insulator is shown in black. The BASIC prefix (*i*P) and suffix (*i*S) are shown are flanked to BASIC methylated linker (light blue) in the 5' end and UTR-RBS linker (dark grey) in the 3' end. (B) A map of a plasmid expressing sfGFP driven by various designs of tested promoter (P_{xxx_BASIC}).

2.3 The use of a self-cleaving ribozyme and UTR-RBS linker enhances output activity and modularity of the standardised promoter design

Prior to evaluating the standardised promoter architecture, we sought to validate the downstream sequence context of the reporter expression. Specifically, we were interested to determine the effects of the RiboJ and 5′ UTR arising from the BASIC linker on gene expression.

Four alternative promoter designs were created (Table 2.1, Figure 2.3A). Design 1 represented a promoter without RiboJ or the UTR region of the BASIC linker. Design 2 was constructed without RiboJ but included the UTR region. Design 3 included RiboJ but omitted the UTR region. Design 4 contained both the RiboJ and UTR region. It is worth noting that the construction of the expression cassette without a UTR-RBS linker, such as Design 1 and 3, was less modular than the one with a UTR-RBS linker (Design 2 and 4). To generate these constructs, the UTR sequence was removed by PCR-based site-directed mutagenesis (Section 6.2.4), leaving only the RBS sequence. Briefly, standardised promoters in Design 1 were derived from Design 2 whereas the standardised promoters in Design 3 were generated from Design 4.

Table 2.1 Summary of the initial design of the standardised promoter.

Design	Terminator	UP element	Core promoter	Ribozyme	UTR	RBS
1	L3S3P11*	UP A	P_{J23xxx}	No	No UTR	RBSc
2 [a]	L3S3P11*	UP A	P_{J23xxx}	No	UTR A	RBSc
3	L3S3P11*	UP A	P_{J23xxx}	RiboJ	No UTR	RBSc
4 [a]	L3S3P11*	UP A	P_{J23xxx}	RiboJ	UTR A	RBSc

[a] Denoted as highly desirable for modularity and easy implementation to BASIC DNA assembly framework[62]. * Mutated nucleotide at position 45 to remove *Bsa*I recognition site[36,176].

In each promoter design, four core promoters (P_{J23xxx}) from the Anderson library with different strengths (P_{J23119}, P_{J23111}, P_{J23104}, and P_{J23101}) were tested. In total, 16 different transcriptional units (TUs) were obtained by combining four core promoters and four promoter designs. These promoter sets were assembled with a *sfgfp* reporter gene in a p15A backbone and were propagated in *E. coli* DH5α cells.

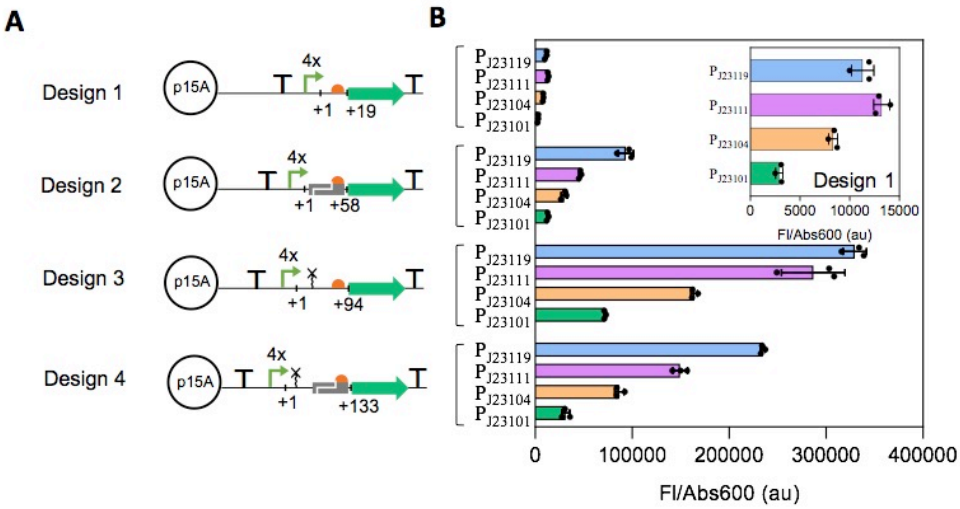


Figure 2.3 Effect of downstream sequences on promoter activity. (A) Schematic of the different promoter designs used. The numbers indicate the position of the TSS (+1) and translational start codon. (B) Functional characterisation result of different promoter architecture showing the response of different designs in Figure 2.3A. The engineered strains were grown in LB medium at 30°C and 600 rpm. The relative promoter activity is defined as Fl/Abs600 taken from cultures at 8 h of incubation. Subsequently, the Fl/Abs600 value was calculated by comparing the normalised read-out of sfGFP and Abs600 from a microplate reader-based measurement. Inset shows the promoter sets within Design 1. Data are shown with error bars for the mean ± standard deviation (SD) of triplicate measurements (black dots).

Notably, Design 1 resulted in the lowest output at all promoter strengths across different promoter designs. A short distance between TSS and translational start codon significantly reduced gene expression compared to the other designs. Moreover, lacking a ribozyme insulator in Design 1 leads to a variation of 5' UTR sequence that interferes the output gene expression as previously investigated[169,170].

The addition of UTR-RBS linker downstream core promoter in Design 2 significantly increased the output expression of sfGFP (Figure 2.3B, compare Design 1 and Design 2). It should be noted that the UTR-RBS linker has been designed to exclude a complex secondary structure. This result suggests that the presence of UTR-RBS in the designed promoter not only improved the modularity and compatibility of the standardised promoter, but also enhanced the mRNA stability[181,182].

In general, the inclusion of the RiboJ gave a substantial increase in expression (Figure 2.3B, Design 1 vs Design 3, Design 2 vs Design 4). The presence of both RiboJ and the UTR-RBS linker gave a significantly higher expression than UTR-RBS alone, but slightly lower than RiboJ alone (Figure 2.3B, compare Design 3 and Design 4). One possible explanation for the reduction compared to just RiboJ is that there may be some interaction between the RiboJ and the UTR

sequence, altering the secondary structure of 5′ UTR of mRNA and therefore affecting the output activity of the designed promoter. Altogether, although Design 4 yielded lower expression of sfGFP in comparison with Design 3, Design 4 can be easily reused through BASIC DNA assembly. Design 4 (with a ribozyme insulator and UTR-RBS linker) was therefore used for further experiments.

2.4 Evaluating output promoter activity in response to other core elements in the standardised promoter design

To further dissect the specific contribution of each element in Design 4, we systematically assessed the promoter activity in response to varying three other major components in the design: terminator, UP element, and ribozyme insulator. Two constitutive core promoters, P_{J23119} and P_{J23101} were used to create different promoter strengths ($P_{J23119} > P_{J23101}$). Similar to the previous tests, this set of promoters was assembled in a p15A plasmid backbone and drove expression of sfGFP (see Figure 2.2B).

2.4.1 Impact of terminator on the activity of the standardised promoters

As mentioned previously, a terminator has an important role in the standardised promoter design, limiting transcriptional read-through from any previous transcriptional units[176,183,184]. Here, the intention was to test the impact of the terminator on the promoter output across two core promoter sequences. To test the impact of different terminators in Design 4, five terminators with varying length and termination strength were chosen (Table 2.2). These terminators were selected from a set of engineered terminators that have been previously studied in *E. coli*[176] and chosen due to their reduced recombination activity[176].

Table 2.2 Selected synthetic terminators used in the standardised promoter design.

Terminator	Length (bp)	Termination strength[a]	Sequence (5'-> 3')	Source
L3S2P21	61	382.13	CTCGGTACCAAATTCCAGAAAAGAGG CCTCCCGAAAGGGGGGCCTTTTTTCGT TTTGGTCC	Chen *et al.*, 2013
L3S2P11	57	261.55	CTCGGTACCAAATTCCAGAAAAGAGA CGCTTTCGAGCGTCTTTTTTCGTTTTGG TCC	Chen *et al.*, 2013
L3S1P13	51	177.94	GACGAACAATAAGGCCTCCCTAACGG GGGGCCTTTTTTATTGATAACAAAA	Chen *et al.*, 2013
L3S2P55	57	255.66	CTCGGTACCAAAGACGAACAATAAGA CGCTGAAAAGCGTCTTTTTTCGTTTTG GTCC	Chen *et al.*, 2013
L3S3P11*	47	172.60	CCAATTATTGAACACCCTTCGGGGTGT TTTTTTGTTTCTGGTCTACC	Chen *et al.*, 2013, Nielsen *et al.*, 2016

[a] Predicted termination strength reproduced from ref [176]. * nucleotide at position 45 is mutated to remove *Bsa*I recognition site[36,176].

Following the chosen terminators, UP element A was used upstream of the core promoter, and RiboJ was placed downstream (Figure 2.4A). The assembled plasmids were propagated in *E. coli* DH5α cells. The promoter activity was evaluated via continuous plate reader measurement (Section 6.3.1). For each core promoter, the output activity (Fl/Abs600) of standardised promoter remained relatively constant for all of the tested terminators (Figure 2.4B).

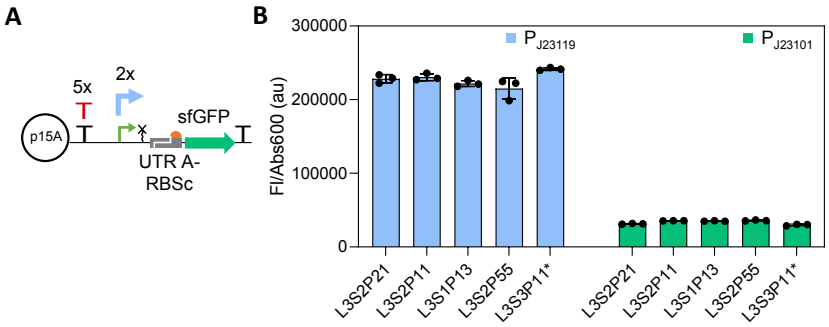


Figure 2.4 Impact of changing various terminators towards the activity of standardised promoters. (A) Schematic design of expression cassette used with five different terminators and two promoters. (B) The promoter activity of P_{J23119} and P_{J23101} in response towards variant of terminators. The promoter activity (Fl/Abs600) was evaluated by cultivating the engineered strains in LB medium at 30°C with continuous shaking at 600 rpm for 6 h. The Fl/Abs600 value was calculated by comparing the normalised read-out of sfGFP and Abs600 from a microplate reader-based measurement. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

2.4.2 Impact of different UP elements on the activity of the standardised promoters

An UP element was included in the standardised promoter design to insulate the promoter activity from the interaction with the previous adjacent sequence. This is intended to produce more consistent promoter output compared to that lacking the UP element. It has been shown in several studies that UP elements help insulate gene expression from a variety of upstream sequences[168,172]. Here, in addition to varying the terminator sequence, the effect of the length and sequence of the UP element on promoter activity was further investigated.

Three strategies were employed to generate and select variants of UP elements to test. First, UP elements with 45 bp length (UP A and C) were computationally designed using R2oDNA Designer[34]. This strategy aimed to reduce secondary structure and potential cross-talk with host-genome, plasmid backbone, and other bioparts. Second, an UP element with a 60 bp length (UP D) was constructed from a known insulating sequence[168]. Third, two UP elements with length of 20 bp (UP B and E) were produced by truncating UP A and UP D, respectively. This truncation created a shorter DNA spacer and aimed to reduce the probability of homologous recombination[164,165] if multiple promoters are used concurrently in a design. Lastly, a set of promoter designs without an UP element (No UP) were created as a control. A summary of the designed UP elements is shown in Table 2.3.

Table 2.3 Selected UP-elements used in the standardised promoter design.

UP element	Length (bp)	Sequence[a] (5'-> 3')	Source
No UP	0	-	This study
UP A	45	ATCTCGTTGTGATAATAGACCTGAAGTGCCTACTC TGGAAAATCT	This study
UP B	20	GTGCCTACTCTGGAAAATCT	This study
UP C	45	TAGTCGGAAGCGTGTTTATCTGCTATCCTACAATA CCTGAAGAAT	This study
UP D	60	CACAGCTAACACCACGTCGTCCCTATCTGCTGCCCT AGGTCTATGAGTGGTTGCTGGATAAC	Davis *et al*., 2010
UP E	20	ATGAGTGGTTGCTGGATAAC	This study

[a] DNA sequence in underlined highlights residue (20 bp) of truncated sequence: UP A was truncated to create UP B and UP D was truncated to create UP E.

Each UP element (Table 2.3) was combined with five different terminators (Table 2.2) and a RiboJ insulator was used downstream of two core promoters (P_{J23119} and P_{J23101}). In total, 60 promoter designs were constructed and tested. Following the previous strategy, this set of

designs drove expression of a sfGFP reporter gene using the UTR A-RBSc linker (Figure 2.5A). The assembled plasmids were tested in DH5α cells.

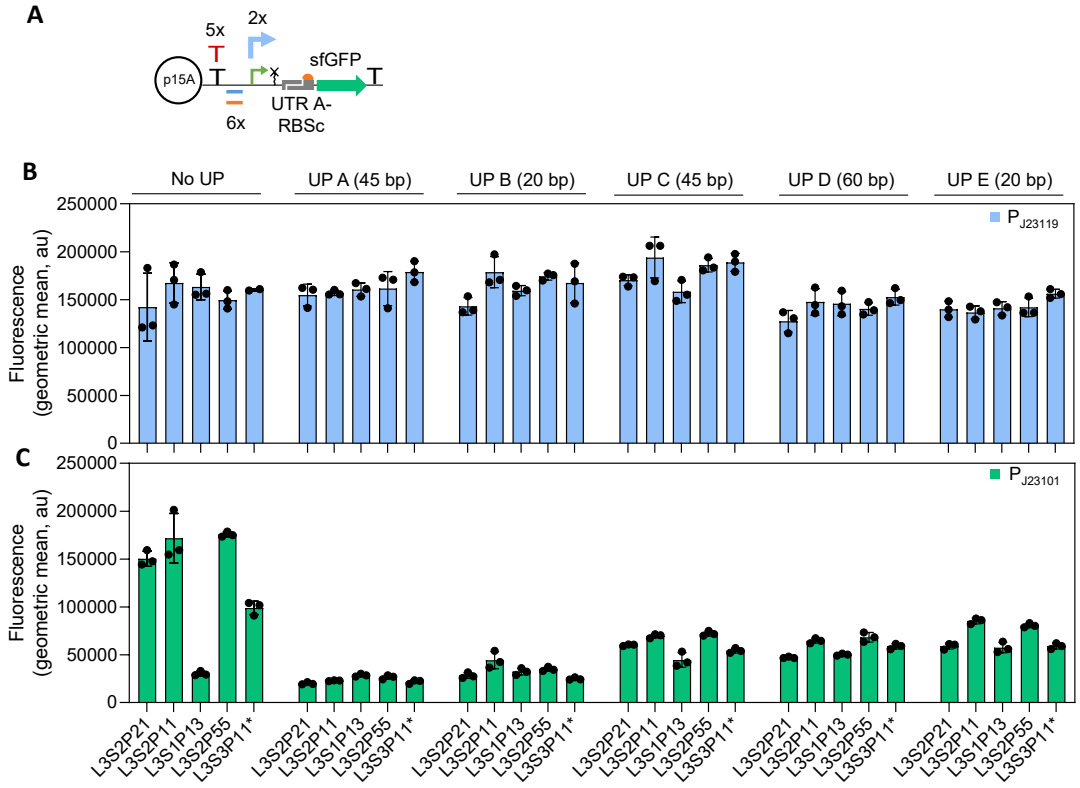


Figure 2.5 Effect of variation of upstream elements towards the performance of standardised promoters. (A) Schematic design of expression cassette used with five different terminators, six variants of UP elements, and two constitutive promoters. (B) The promoter activity of P_{J23119} and (C) The promoter activity of P_{J23101} in response towards a variant of terminators and UP elements. The engineered strains were grown in LB medium at 30°C with continuous shaking at 600 rpm. The sfGFP fluorescence was measured at exponential phase (6 h) by flow cytometry. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

The presence of different UP elements did not seem to affect the output activity of P_{J23119} across different terminators (Figures 2.5B). However, promoter designs without an UP element (No UP) showed fluctuation in activity for promoter P_{J23101} across various terminators (Figure 2.5C). This effect was observed in the use of a medium promoter (P_{J23101}) rather than a strong promoter (P_{J23119}). The use of a strong promoter may result in maximum gene expression due to saturation of promoter activity. Therefore, inconsistent output activity due to the different upstream elements could not be observed in the use of a strong promoter.

Insertion of UP elements (UP A-E) resulted in a relatively constant expression level for both P_{J23119} and P_{J23101}. These results demonstrated the important role of UP element in the

promoter architecture, which is to insulate promoter activity from variation in upstream DNA contexts. Furthermore, we found that sequence composition rather than length of UP elements affects the output level of the tested constitutive promoters.

2.4.3 Impact of ribozyme choice on the activity of the standardised promoters

Next we explored the impact of different ribozymes downstream of the core promoter. In addition to a natural ribozyme (i.e. RiboJ), three engineered ribozymes (RiboJ10, RiboJ51, and RiboJ53) were tested in the standardised promoter design (Table 2.4). The engineered ribozymes were derived from RiboJ with a modification in the sTRSV scaffold[36]. The length of all four ribozyme variants is 75 bp.

These ribozymes were selected by considering their cleavage efficiency and insulator functionality as previously described[36,169]. The cleavage activity was determined using a Rapid Amplification of Complementary DNA End (5'-RACE) method[36,169]. Briefly, the complementary DNA (cDNA) was synthesised from mRNA by reverse-transcription for PCR amplification. The ratio between cleaved and total cDNA from acrylamide gel was calculated to measure the cleavage activity. The insulation functionality was measured by comparing the GFP expression ratio of GFP and cI-GFP under P_{LlacO1} and P_{Tac} upon IPTG induction with various concentrations[36,169]. If the slopes are approximately constant, especially in the presence of a ribozyme insulator, the two reporters are proportionally expressed at different promoter activities. A set of ribozymes with a cleavage efficiency of 1.0 and various functional insulation values were chosen (Table 2.4).

Table 2.4 Selected ribozyme insulators used in the standardised promoter design.

Ribozyme insulator	Cleavage efficiency[a]	Insulator functionality[b]		Sequence (5'-> 3')	Source
		Slope[c] $P_{LlacO-1}$	Slope[c] P_{Tac}		
RiboJ	1.0	1.5	1.5	AGCTGTC‖ACCGGATGTG CTTTCCGGTCTGATGAGT CCGTGAGGACGAAACAG CCTCTACAAATAATTTTGT TTAA	Lou *et al.*, 2012, Nielsen *et al.*, 2016
RiboJ10	1.0	1.4	1.6	AGCGCTC‖AACGGGTGTG CTTCCCGTTCTGATGAGT CCGTGAGGACGAAAGCG CCTCTACAAATAATTTTGT TTAA	Nielsen *et al.*, 2016
RiboJ51	1.0	1.8	1.5	AGTAGTC‖ACCGGCTGTG CTTGCCGGTCTGATGAGC CTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGT TTAA	Nielsen *et al.*, 2016
RiboJ53	1.0	2.1	1.8	AGCGGTC‖AACGCATGTG CTTTGCGTTCTGATGAGA CAGTGATGTCGAAACCGC CTCTACAAATAATTTTGTT TAA	Nielsen *et al.*, 2016

The value of [a] cleavage efficiency and [b] insulator functionality were reproduced from ref[36]. [c] This value is calculated based on the GFP expression ratio of GFP and cI-GFP under P_{LlacO1} and P_{Tac} upon IPTG induction with various concentrations. ‖ indicates the cleavage site. The underlined sequence shows a conserved sequence of an additional hairpin.

To investigate the impact of ribozyme choice on the promoter activity, two constitutive promoters P_{J23119} and P_{J23101} were used. Terminator L3S3P11* and UP B (20 bp) were employed in the upstream region of these core promoters. In the downstream region of these promoters, four ribozymes were placed and tested (Figure 2.6A).

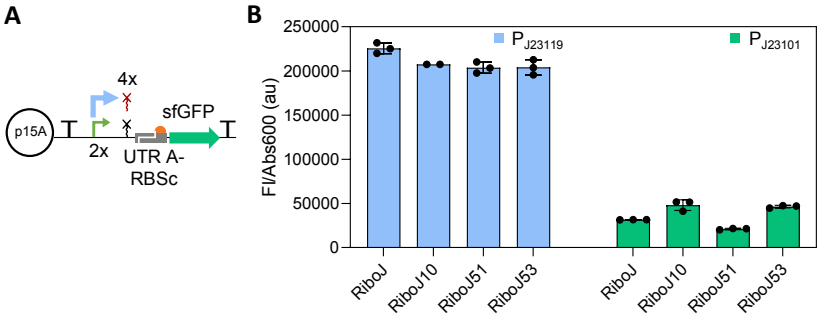


Figure 2.6 Impact of variant ribozyme insulators towards output activity. (A) Schematic design of expression cassette used with four different ribozyme insulators and two promoters. (B) The promoter activity of P_{J23119} and P_{J23101} in response towards a variant of ribozyme insulators. *E. coli* DH5α cells were transformed with the assembled plasmids and grown in LB medium at 30°C for 6 h incubation. The Fl/Abs600 value was calculated by comparing the normalised read-out of sfGFP and Abs600 from a microplate reader-based measurement. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

Different ribozymes in the designs resulted in a relatively consistent output for core promoter P_{J23119} (Figure 2.6B). In contrast, distribution of up to 2-fold was observed for sfGFP output for core promoter P_{J23101}. Furthermore, this variation may be caused by each ribozyme having different sequences after the cleavage site, which could trigger conformational changes in the 5' region of the mRNA and affect other processes during gene expression. Overall, we found that both upstream and downstream regions surrounding the core promoters determined the output activity of the standardised promoters.

2.5 Selection of three alternative standardised promoter groups

In the previous sections, the standardised promoter design was tested in a single gene expression system. The availability of standardised bacterial promoters together with a modular assembly approach can potentially be exploited to build larger synthetic circuits with sophisticated functions[36,53,54,185]. These circuits often require multi-gene expression. However, reusing the same standardised architecture of promoters in a design may lead to homologous recombination[163]. The frequency of homologous recombination is dependent on the genetic host strain, homology, and length of repetitive sequence[164,165]. There is a linear correlation of recombination frequency and homology sequence for a contiguous sequence with length more than 27 bp[164].

To overcome this, diversification of standardised promoters within the conserved architecture (Design 4, Figure 2.3) was investigated next. Three groups of standardised promoter designs were developed. This set was termed $P_{xxx\_BASIC(y)}$, where "xxx" represented core promoters of interest and "y" indicated standardised promoter groups (Table 2.5). Variants of core elements in standardised promoter groups were selected from previous characterisation results (Section 2.4).

Table 2.5 Groups of the standardised promoter.

Group promoter	Terminator	UP element	Core promoter	Ribozyme insulator
$P_{xxx\_BASIC(1)}$	L3S2P21	UP B	P_{xxx}	RiboJ
$P_{xxx\_BASIC(2)}$	L3S2P11	UP B	P_{xxx}	RiboJ10
$P_{xxx\_BASIC(3)}$	L3S1P13	UP B	P_{xxx}	RiboJ51

Considering the structure of the standardised promoter outlined in Section 2.2 and evaluated in Section 2.4, the key components with a sequence length > 27 bp are the terminator and ribozyme. Both components are 47-75 bp long, sufficient to trigger homologous recombination[164]. Therefore, the type of terminators and ribozyme insulators need to be varied for each standardised promoter group. To do this, three different terminators were selected: L3S2P21, L3S2P11, and L3S1P13; and three ribozymes variants: RiboJ, RiboJ10, and Ribo51. Furthermore, three different groups of standardised architecture were defined by varying the selected terminators and ribozymes, whereas the UP element was conserved. The UP element can be selected from UP element variants with a shorter length (Table 2.3). All designs shared the same 20 bp UP element (UP B). Subsequently, the minimal core promoter was inserted in between the UP B element and ribozyme insulators. With this underlying structure, a different type of promoter can be easily generated by substituting the core promoter to one of interest.

2.6 Expanding the library of bacterial constitutive promoters through the implementation of standardised architectures

The diverse standardised promoter groups are also beneficial for balancing gene expression in constructs containing multiple transcriptional units arrangement, such as those commonly found in biosynthetic pathways[78,104]. In this section, the standardised design was focused on constitutive promoters. Eight constitutive Anderson promoters (Table 2.6) were selected from the iGEM repository (http://parts.igem.org).

Table 2.6 List of core constitutive Anderson promoters.

Core promoter (P_{J23xxx})	**Sequence[a] (5'-> 3')**
P_{J23119}	TTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC
P_{J23111}	TTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC
P_{J23104}	TTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGC
P_{J23101}	TTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGC
P_{J23108}	CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC
P_{J23106}	TTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC
P_{J23105}	TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGC
P_{J23116}	TTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGC

[a] DNA sequence is underlined indicates the -35 and -10 boxes.

These minimal constitutive promoters (35 bp) were included in the core element of each standardised promoter group (Table 2.5). In total, 24 different promoter designs were created. These promoter collections were used to control sfGFP expression connected by the UTR A-RBSc in a p15A backbone (Figures 2.7A, 2.7B, and 2.7C). These assembled plasmids were propagated in *E. coli* DH5α cells and the promoter activity of each calculated. As shown in Figure 2.7, the order of constitutive promoter activity was maintained across different promoter groups: P_{J23119_BASIC} > P_{J23111_BASIC} > P_{J23104_BASIC} > P_{J23101_BASIC} > P_{J23108_BASIC} > P_{J23106_BASIC} > P_{J23105_BASIC} > P_{J23116_BASIC}. Additionally, the order of relative promoter activity for the same constitutive promoter sets was also maintained, $P_{xxx\_BASIC(2)}$ > $P_{xxx\_BASIC(1)}$ > $P_{xxx\_BASIC(3)}$. This library of standardised constitutive promoters with varying promoter strength could be used to balance constitutive gene expression in multi-gene systems and complex regulatory network.

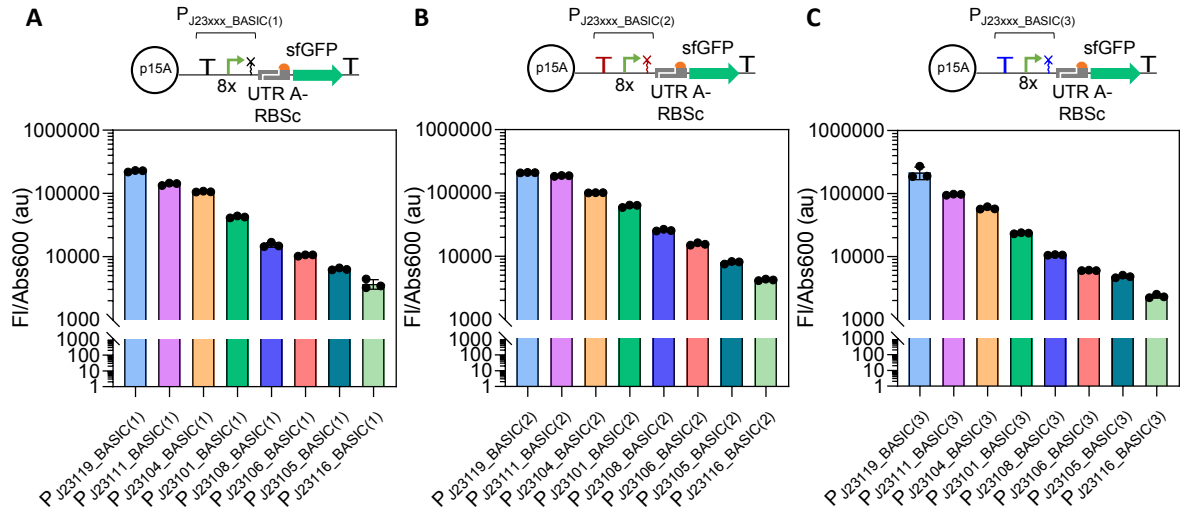


Figure 2.7 Characterisation of constitutive promoters within different standardised design groups for (A) $P_{J23xxx\_BASIC(1)}$, (B) $P_{J23xxx\_BASIC(2)}$, and (C) $P_{J23xxx\_BASIC(3)}$. Representative schematic design of corresponding expression cassette is shown with eight different promoters for each group. The engineered strains were cultured in LB medium at 30°C and 600 rpm. After 24 h incubation, the promoter activity (Fl/Abs600) were calculated. Y-axis showed log expression level. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

2.7 The inclusion of ribozyme insulator increases the output level of standardised inducible promoters

After a library of standardised constitutive promoters had been generated, the same standardised design was initially expected to be applicable for bacterial promoter sets that regulate inducible gene expression. It should be noted that these inducible promoters typically have a longer sequence than constitutive ones. In this case, the library of

standardised inducible promoters could be generated simply by inserting a core inducible promoter in the standardised design. As initial set, three frequently employed inducible systems were chosen: Ara, Tac, and Tet-based inducible systems (Table 2.7). The Ara-inducible system represents an activator-based system while Tac and Tet-inducible systems were selected as examples of repressor-based systems. These core inducible promoter sequences were adapted from previous work[36].

Table 2.7 List of inducible promoters and cognate regulator cassettes used in the initial design of standardised promoter.

Core promoter	Source of core promoter	Regulator cassette	Source of regulator cassette	Inducer
P_{BAD}	Nielsen *et al*., 2016	P_{J23105}-AraC	Nielsen *et al*., 2016	Ara
P_{Tac}	Nielsen *et al*., 2016	P_{LacI}-LacI	Nielsen *et al*., 2016	IPTG
P_{Tet}	Nielsen *et al*., 2016	P_{J23116}-TetR	This study	aTc

Before creating this standardised inducible promoter library, the design was firstly evaluated in the context of different core inducible promoters and the effect of other elements. Specifically, the effect of the ribozyme insulator (i.e. RiboJ) was tested. Following the earlier designs outlined in Sections 2.2 and 2.3, two alternatives were tested by varying the presence and absence of RiboJ downstream of the core inducible promoters. Immediately downstream of the core inducible promoters, UTR A-RBSc was used to connect each design to a *sfgfp* gene reporter (Figures 2.8A, 2.8B, and 2.8C).

Typically, gene regulation via an inducible promoter requires expression of its corresponding Transcription Factor (TF). Therefore, the cognate TF (i.e. AraC, LacI, and TetR) of each promoter set was expressed under a constitutive promoter. For example, AraC was driven by a constitutive promoter P_{J23105}, LacI repressor was expressed under a constitutive promoter P_{LacI}, and TetR was constitutively expressed under P_{J23116}. The P_{J23105}-AraC and P_{LacI}-LacI expression cassettes were extracted from plasmid pAN3945[36] by Polymerase Chain Reaction (PCR) (Section 6.2.4) while P_{J23116}-TetR was chemically synthesised with a medium RBS and ordered from Integrated DNA Technology. All TF cassettes were designed as BASIC bioparts.

The inducible promoter and its regulator cassette were then combined in a single plasmid system with a pMB1(~15-20 copy per cell). The assembled plasmids were propagated in *E. coli* DH5α cells. For all inducible promoters tested, the usage of RiboJ improved output promoter activity by an average of 2-fold (Figures 2.8A, 2.8B, and 2.8C). Presumably, this

improvement was caused by an increase in mRNA stability as previously reported[180,181] (Section 2.3). Moreover, this result confirmed that previously standardised promoter designs can also be applied to generate a library of standardised inducible promoters.

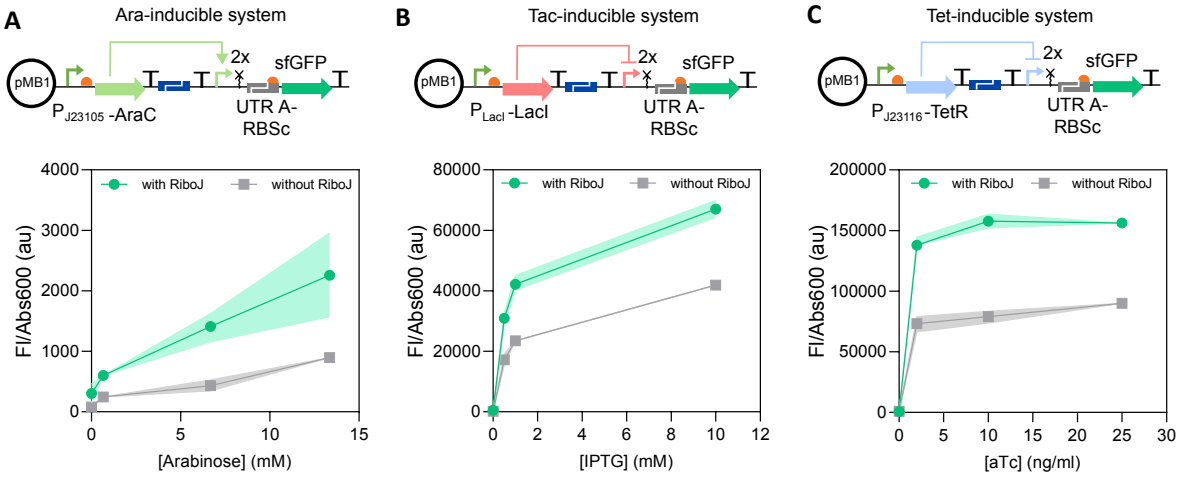


Figure 2.8 Dose responses of tested inducible systems: (A) Ara-inducible system, (B) Tac-inducible system, and (C) Tet-inducible system. Representative schematic design of corresponding expression cassette is shown with the presence and absence of RiboJ insulator downstream of the core inducible promoter. The engineered cells were grown in LB medium at 30°C and 600 rpm. After 1 h outgrowth, the expression of sfGFP was induced with addition of corresponding inducer molecules: L-arabinose (Ara), isopropyl β-d-1-thiogalactopyranoside (IPTG), and anhydrotetracycline (aTc) (Table 2.7). At 24 h post-induction, the promoter activity (Fl/Abs600) was calculated. Lines and points show the mean from three independent replicates with shaded area showing ± SD.

2.8 The sensitivity of the standardised inducible promoters is modulated by the abundance of its cognate regulatory protein

Using the defined standardised structure of the inducible system, a new variant of standardised inducible promoters can be easily generated by substituting the core inducible promoter of interest. Before expanding these variants, it is important to understand the induction response of a standardised promoter when used in different plasmid copy numbers and regulated by different levels of TF. As an example, the Tet-inducible system based on a repressor-regulated system was selected.

Three different expression systems were generated with varied plasmid copy numbers and promoter strengths driving the TF for the Tet-inducible system (Figure 2.9A). The first expression cassette used the initial TetR cassette with a weak promoter (P_{J23116}) to regulate the standardised Tet-inducible promoter ($P_{Tet\_BASIC(1)}$) in a medium plasmid copy number (p15A,

~10 copy per cell). The second system was constructed with the similar TetR expression cassette in a higher plasmid copy number (pSB1C3, ~100-300 copy per cell). In the third system, the promoter driving expression of the TetR cassette was replaced by P_{J23105} in the same plasmid backbone (pSB1C3). It should be noted that the relative promoter strength of P_{J23105} was expected to be stronger than P_{J23116} (see Figure 2.7). The modification of promoter in this transcriptional unit was performed using site-directed mutagenesis of the TetR cassette as a BASIC biopart (see Section 6.2.4). All designed systems drove expression of sfGFP and were preceded by UTR A-RBSa carrying a strong RBS. Following the previous protocol, the assembled plasmids were transformed in *E. coli* DH5α cells.

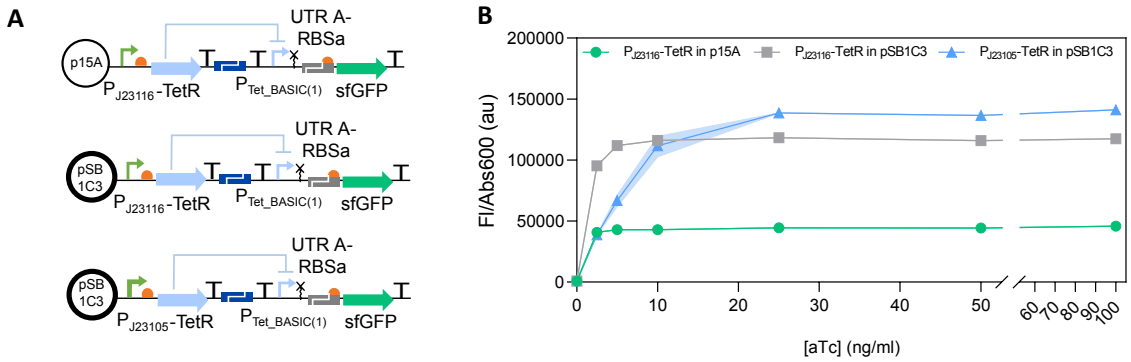


Figure 2.9 The response of the Tet-inducible system with the modulated abundance of its regulatory cassette. (A) The schematic design depicts three expression systems constructed in the different copies of plasmid and TetR. (B) Dose responses of tested systems as a function of inducer (aTc) concentration (0-100 ng/ml). The engineered cells were grown in LB medium at 30°C and 600 rpm. The expression of sfGFP was turned on with addition of aTc at different concentrations after 1 h of outgrowth. The promoter activity (Fl/Abs600) were calculated in stationary phase (24 h post-induction). Lines and points show the mean from three independent replicates with shaded area showing ± SD.

The output promoter activity as a function of inducer concentration was modulated by plasmid copy numbers and promoter strengths controlling TF in the system (Figure 2.9B). At the same induction level, the use of P_{J23116}-TetR and regulated promoter in a higher plasmid copy number resulted in a greater response of on- and off-states. The pattern of induction response was closely matched that on the medium plasmid copy number.

A slightly shifted induction curve was observed when a stronger promoter controlled expression of TetR (P_{J23105}-TetR), especially with aTc concentration below 10 ng/ml. One possible reason is the binding saturation of regulatory protein and operator site tunes sensitivity of the standardised inducible promoter. This graded response further confirmed the previously reported finding regarding titration of TF concentration in a synthetic Tet-

controlled expression system[186]. In comparison to P_{J23116}-TetR, the designed system with P_{J23105}-TetR showed a greater sfGFP output at higher concentration of inducer. A further investigation is required to elucidate this behaviour discrepancy.

Furthermore, this section revealed that the standardised inducible promoter is portable in plasmid, with different copy numbers. Additionally, the sensitivity of the designed inducible system can be modulated by changing the abundance of its transcription factor, i.e. modifying the strength of promoter driving the corresponding transcription factor.

2.9 Multi-gene expression constructs using different standardised inducible promoter sets exhibit a comparable induction response

In addition to transcription factor abundance, the performance of the standardised Tet-inducible promoters was then tested when used in multi-transcriptional units (TUs). In this regard, multiple operator sites in the Tet promoters are targeted by the same intracellular concentration of TetR. The induction responses of constructs with a single and double output system were compared.

In a single expression system, the previously constructed plasmid consisted of only sfGFP expression driven by $P_{Tet\_BASIC(1)}$ and UTR A-RBSa was used. To enable multiple Tet promoters in the double expression systems, two alternative standardised Tet promoters, $P_{Tet\_BASIC(1)}$ and $P_{Tet\_BASIC(2)}$ were generated from different groups of the standardised promoters (Table 2.5). These standardised inducible promoters were followed by different UTR-RBS linkers to create two independent transcriptional units with sfGFP and Blue Fluorescent Protein (BFP) as fluorescent outputs (Figure 2.10A). Briefly, the $P_{Tet\_BASIC(1)}$ was used to drive sfGFP expression and was preceded by UTR A-RBSa, whereas $P_{Tet\_BASIC(2)}$ was used to control BFP expression and assembled with the UTR C-RBSa.

A comparable induction curve was obtained in both systems (Figure 2.10B). At the same induction level, the sfGFP expression in the single output system was slightly higher than for the double output system. It could be argued that the shared resources needed to express

the double expression system may have caused a reduction in sfGFP output. However, it should be noted that the double output system carries two independent TetR-binding sites. Since the TF is targeting two different binding sites, it required a higher concentration of inducer to achieve similar output. For example, the sfGFP expression in a single output system with 5 ng/ml is similar to that expressed in a double output system with 10 ng/ml aTc. In the double plasmid system, the relatively matched response curves of sfGFP and BFP indicated an equal induction response across the same inducible promoters. This also demonstrated that the same promoter within different groups of standardised promoters can regulate multiple TUs concurrently with a comparable response.

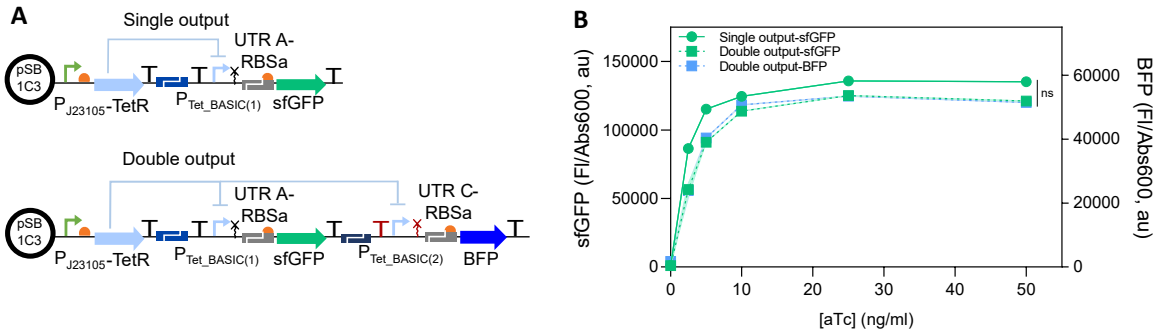


Figure 2.10 The performance of Tet-inducible promoter constructed in the multi-gene system. (A) Schematic design of genetic circuits with single output (top panel) and double output (bottom panel) system. (B) Dose responses of the designed systems (Figure 2.10B) as a function of inducer (aTc) concentration (0-50 ng/ml). These constructs were expressed in DH5α cells and grown in LB medium at 30°C and 600 rpm. The aTc inducer at different concentrations was added after 1 h of outgrowth to activate the expression of sfGFP and BFP. At 24 h post-induction, the promoter activity (Fl/Abs600) was calculated. The sfGFP output is shown in Y-left axis while BFP output is shown in Y-right axis. Lines and points show the mean from three independent replicates with shaded area showing ± SD. Statistically significant differences determined using Student's t-test (ns represents not significant).

2.10 A set of inducible promoters display compatibility in a host strain with genome-integrated regulator cassettes (*E. coli* DH10b_marionette)

To regulate gene expression, the standardised inducible promoters require expression of their cognate TFs. These TFs can be included in the plasmid construction (Sections 2.7, 2.8, and 2.9). In more complex circuits with multiple inducible systems, such strategy leads to large and burdensome regulatory cassettes. To overcome this, a set of standardised inducible promoters was characterised in a host strain containing genome-integrated TFs. Recently, Meyer and colleagues have developed *E. coli* marionette strains containing 12 regulatory proteins embedded in the genome[80]. The corresponding regulatory cassettes were clustered

into three operons driven by three variants of constitutive promoters. These modified strains have also been previously shown to independently control gene expression with low background expression, high dynamic range, increased sensitivity, and reduced cross-talk[80].

To further explore the potential implementation of standardised inducible promoters in these marionette strains, seven inducible promoters covering activator-regulated (P_{BAD}, P_{Cin}, and P_{LuxB}) and repressor-regulated (P_{PhlF}, P_{VanCC}, P_{Tet}, and P_{CymRC}) systems were characterised. These chemically inducible promoters were selected and adapted from previous works by considering their cognate TFs in the host-genome[36,80]. Like the standardised constitutive promoter sets (Section 2.6), these core inducible promoters were incorporated into three different promoter groups (Table 2.5) to expand the library of standardised inducible promoters available. In total, 21 standardised inducible promoters were constructed for functional characterisation with the addition of a corresponding molecule inducer to the growth medium. A summary of the selected core inducible promoters is provided in Table 2.8.

Table 2.8 List of core inducible promoters.

Core promoter (P_{xxx})	Type	Inducer[a]	Length (bp)	Sequence (5'-> 3')	Source
P_{PhlF}	Repressor-regulated	DAPG	66	CGACGTACGGTGGAATCTGATTCGTT ACCAATTGACATGATACGAAACGTAC CGTATCGTTAAGGT	Meyer *et al.*, 2018
P_{VanCC}	Repressor-regulated	Van	50	ATTGGATCCAATTGACAGCTAGCTCAG TCCTAGGTACCATTGGATCCAAT	Meyer *et al.*, 2018
P_{LuxB}	Activator-regulated	3OC6-HSL	55	ACCTGTAGGATCGTACAGGTTTACGC AAGAAAATGGTTTGTTACAGTCGAAT AAA	Meyer *et al.*, 2018
P_{BAD}	Activator-regulated	Ara	289	AGAAACCAATTGTCCATATTGCATCAG ACATTGCCGTCACTGCGTCTTTTACTG GCTCTTCTCGCTAACCAAACCGGTAAC CCCGCTTATTAAAAGCATTCTGTAACA AAGCGGGACCAAAGCCATGACAAAAA CGCGTAACAAAAGTGTCTATAATCAC GGCAGAAAAGTCCACATTGATTATTTG CACGGCGTCACACTTTGCTATGCCATA GCATTTTTATCCATAAGATTAGCGGAT CCTACCTGACGCTTTTTATCGCAACTCT CTACTGTTTCTCCATACCCG	Nielsen *et al.*, 2016
P_{Tet}	Repressor-regulated	aTc	74	TACTCCACCGTTGGCTTTTTTCCCTATC AGTGATAGAGATTGACATCCCTATCA GTGATAGAGATAATGAGCAC	Nielsen *et al.*, 2016
P_{CymRC}	Repressor-regulated	Cuma	90	AACAAACAGACAATCTGGTCTGTTTGT ATTATGGAAAATTTTTCTGTATAATAG ATTCAACAAACAGACAATCTGGTCTGT TTGTATTAT	Meyer *et al.*, 2018
P_{Cin}	Activator-regulated	3OHC14-AHL	226	CCCTTTGTGCGTCCAAACGGACGCAC GGCGCTCTAAAGCGGGTCGCGATCTT TCAGATTCGCTCCTCGCGCTTTCAGTC TTTGTTTTGGCGCATGTCGTTATCGCA AAACCGCTGCACACTTTTGCGCGACAT GCTCTGATCCCCCTCATCTGGGGGGG CCTATCTGAGGGAATTTCCGATCCGGC TCGCCTGAACCATTCTGCTTTCCACGA ACTTGAAAACGCT	Meyer *et al.*, 2018

[a] DAPG: 2,4-diacetylphophloroglucinol; Van: vanillic acid; 3OC6-HSL: 3-oxohexanoyl-homoserine lactone; Ara: L-arabinose; aTc: anhydrotetracycline; Cuma: cuminic acid; and 3OHC14-AHL: 3-hydroxytetradecanoyl-homoserine lactone.

To characterise this library, each standardised inducible promoter was used to drive expression of the sfGFP reporter gene using UTR A-RBSc and assembled in a p15A backbone (Figure 2.11A). The assembled plasmid sets were then propagated in one of the *E. coli* marionette strains, in particularly the marionette cloning strain (DH10b_marionette, K-12 strain). The output promoter activity for each inducible system was then compared across different standardised promoter groups.

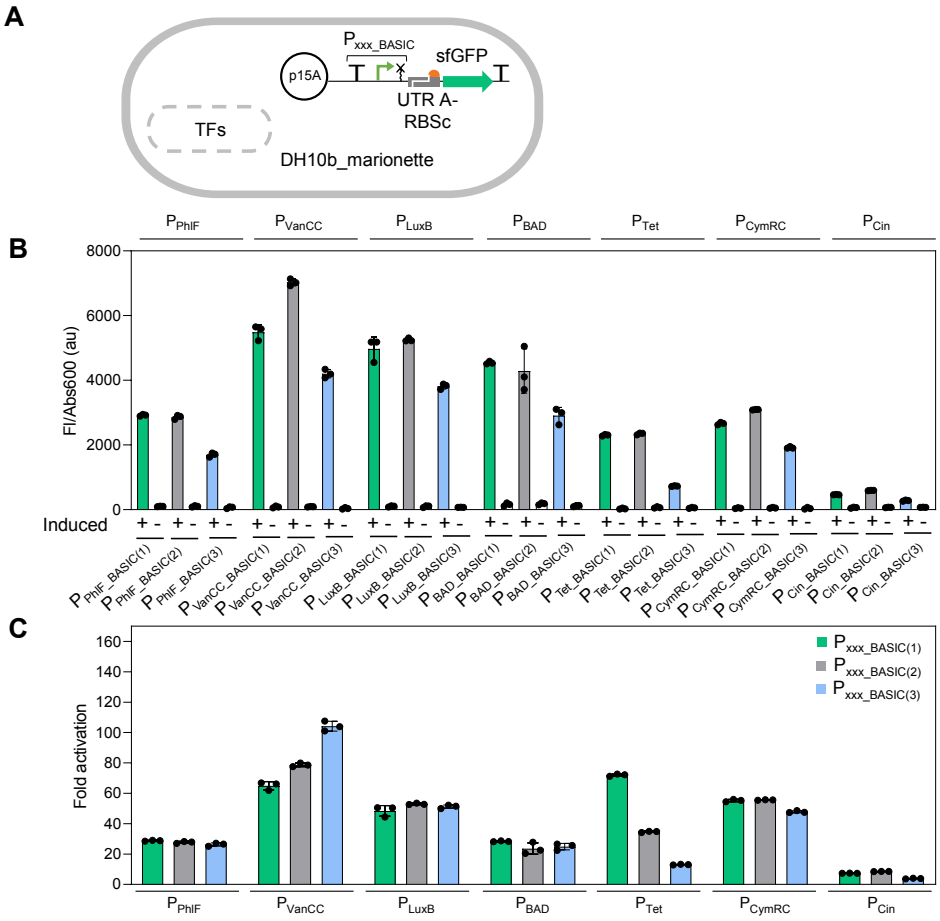


Figure 2.11 The characterisation of a set of standardised inducible promoters in *E. coli* DH10b_marionette strains. (A) Schematic design of genetic circuits with a variety of tested inducible promoter (P_{xxx_BASIC}). The assembled plasmid contains promoter and reporter gene only while the TFs are encoded in the genome. (B) The functional characterisation result and (C) Fold activation of the standardised inducible promoters in a given inducer concentration. The engineered cells were grown in LB medium at 30°C and 600 rpm. After 1 h of outgrowth, a set of inducers were added: 25 µM DAPG, 100 µM cuminic acid, 10 µg/ml aTc, 100 µM vanilic acid, 10 µM 3OC6-HSL, 10 µM 3OHC14-AHL, and 4 mM arabinose. After 24 h post-induction, the promoter activity (Fl/Abs600) were calculated. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

All the standardised inducible promoter sets were functional in DH10b_marionette strains (Figure 2.11B). Low basal expression of promoter activity was obtained without addition of any inducers, indicating tight control of the inducible system even when regulated by transcriptional factors embedded in the host genome. At a given inducer concentration, a distribution of promoter activity was observed across different standardised promoter groups. Typically, the order of output promoter activity in each inducible promoter sets was $P_{xxx\_BASIC(2)} > P_{xxx\_BASIC(1)} > P_{xxx\_BASIC(3)}$. This pattern seems to be consistent with previously observed results in the standardised constitutive promoters (Section 2.6 and Figure 2.7).

To gain a better understanding of promoter behaviour, fold activation for each set of inducible promoters were calculated by dividing the output activity of each promoter for induced and uninduced states (Figure 2.11C). Interestingly, the fold activation of P_{PhlF}, P_{LuxB}, P_{BAD}, P_{CymRC}, and P_{Cin} remained relatively constant across different standardised promoter groups, with P_{VanCC} and P_{Tet} the exception. However, further experiments are required to elucidate this effect.

2.11 Standardised inducible promoter sets are portable to a different marionette strain (*E. coli* BL21(DE3)_marionette)

Having shown the compatibility of the standardised inducible promoter set in DH10b_marionette strains, we next evaluated their performance in a different marionette strains to demonstrate their portability. A marionette strain derived from the protease-deficient *E. coli* (BL21(DE3)_marionette, B strain) was transformed with four constructed plasmids (Figure 2.12A), which used activator-based regulation (P_{LuxB}, and P_{BAD}) and repressor-based regulation (P_{PhlF}, and P_{VanCC}).

All the standardised inducible promoter sets were also functional in BL21(DE3)_marionette strains (Figure 2.12B). The response of the standardised inducible promoter set in BL21(DE3)_marionette strains showed a correlation to those in DH10b_marionette strains (Figure 2.12C). Additionally, a low fold activation was also observed in both marionette strains for the activator-based promoters (i.e. P_{LuxB}, and P_{BAD}), indicating a slow induction response as previously observed through time-course induction[80]. Meanwhile, a difference in the fold activation across standardised promoter groups was observed for P_{PhlF} and P_{VanCC} (Figure 2.12C). It has been reported that these core promoters are host and growth phase-dependent[80].

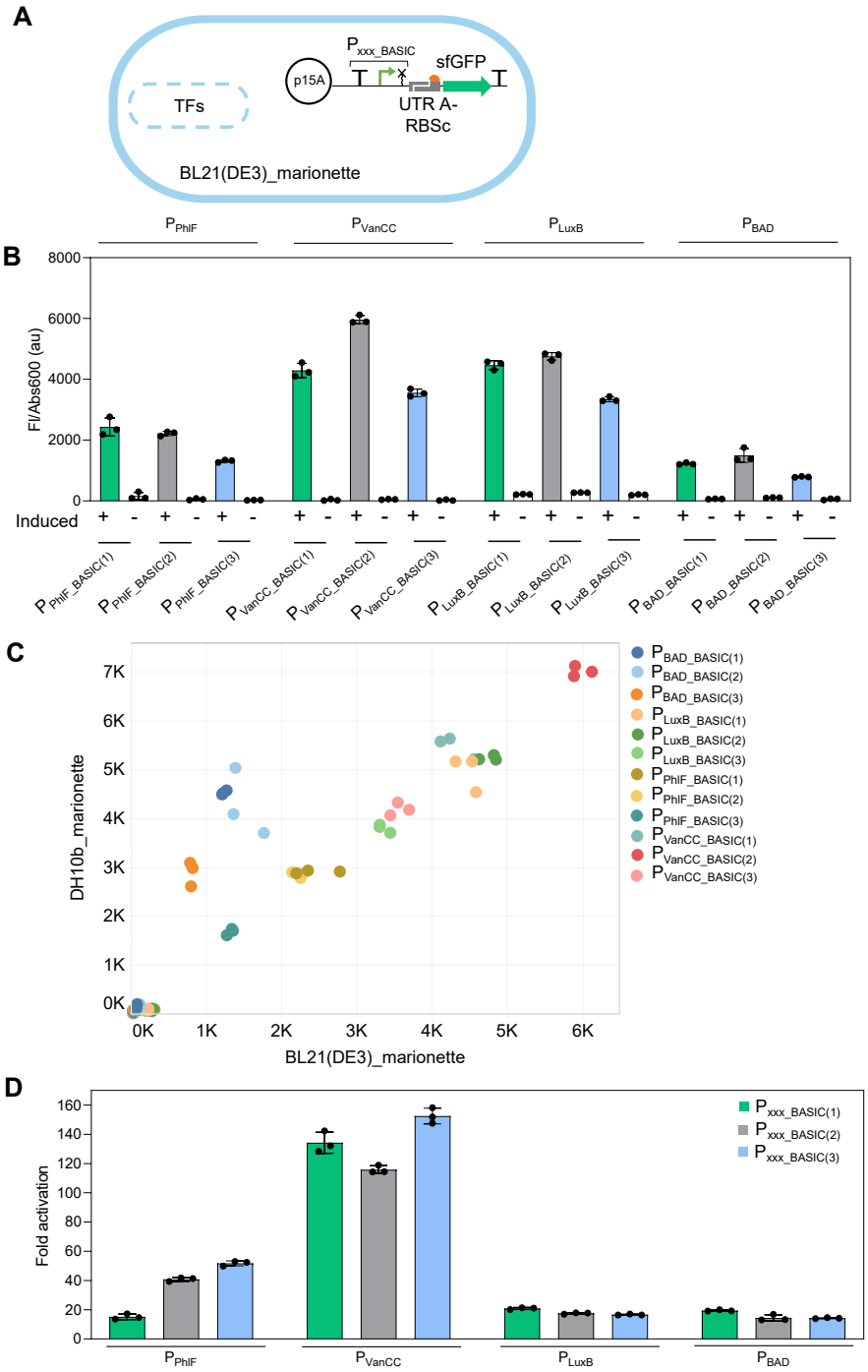


Figure 2.12 The characterisation of a set of standardised inducible promoters in *E. coli* BL21(DE3)_marionette strains. (A) Schematic design of genetic circuits with a variety of tested inducible promoter (P_{xxx_BASIC}). The assembled plasmid contains promoter and reporter gene only while the TFs are encoded in the genome. (B) The functional characterisation result and (C) Correlation of inducible promoters' activities in *E. coli* BL21(DE3)_marionette and DH10b_marionette strains is shown. (D) Fold activation of standardised inducible promoter in a given inducer concentration. The modified cells carrying these constructs were cultured in LB medium at 30°C and 600 rpm. After 1 h outgrowth, a set of inducers were added into growth medium: 25 µM DAPG, 100 µM vanillic acid, 10 µM 3OC6-HSL, and 4 mM arabinose. After 24 h post-induction, the promoter activity (Fl/Abs600) was calculated. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

The results in this section further confirmed that the standardised inducible promoters are compatible with genome-integrated TFs, with their performance largely affected by the genetic host context.

2.12 The use of standardised inducible promoters enables multiplexed and independent control of multi-gene expression in marionette strains

The individual characterisation of standardised inducible promoters in two host backgrounds of marionette strains (Sections 2.10 and 2.11) provides relevant information to guide implementation of gene expression from multiple Transcriptional Units (TUs). In this section, the three standardised inducible promoters within three different standardised promoter groups were further tested: $P_{Tet\_BASIC(1)}$, $P_{VanCC\_BASIC(2)}$, and $P_{CymRC\_BASIC(3)}$. To create a functional TU, each of those promoters was connected to various UTR-RBS linkers and reporter genes (Figure 2.13A). Briefly, (i) the $P_{Tet\_BASIC(1)}$ was employed to drive sfGFP expression and preceded by UTR A-RBSc, (ii) the $P_{VanCC\_BASIC(2)}$ was used to control BFP expression and assembled with UTR B-RBSc, and (iii) the $P_{CymRC\_BASIC(3)}$ was deployed to control Monomeric Cherry Fluorescent Protein (mCherry) expression and connected with UTR C-RBSc. These expression cassettes were further joined together using BASIC neutral linkers (L1 and L3) in a p15A backbone. Subsequently, these assembled plasmids were placed in *E. coli* DH10b and *E. coli* BL21(DE3)_marionette strains (Figure 2.13B).

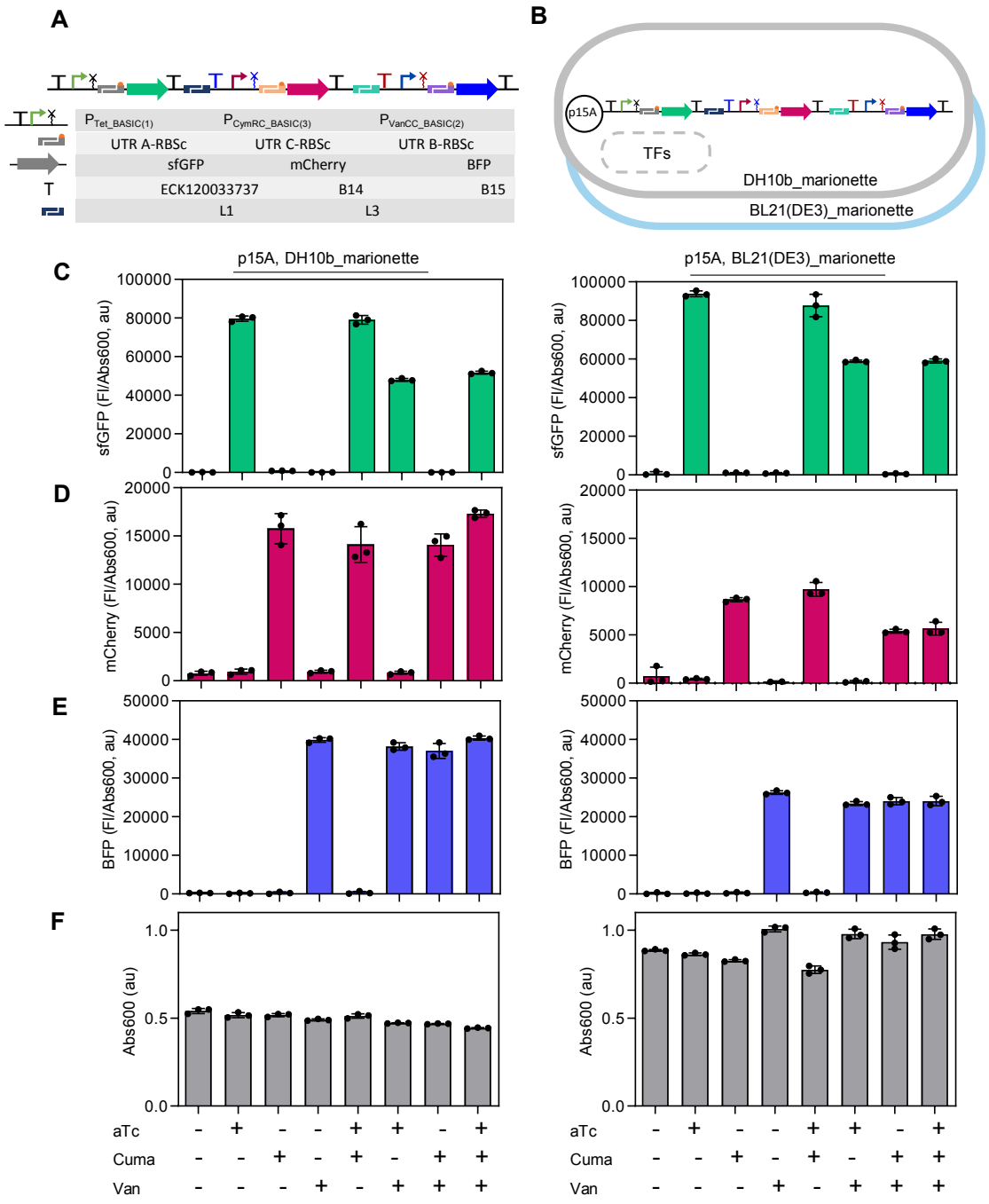


Figure 2.13 Multiplexed and independent regulation of a multi-gene expression system in different marionette strains. (A) Schematic design of genetic circuits with three output fluorescent reporters assembled in a p15A backbone. Each genetic element is indicated. (B) The expression of circuits is independently controlled by genome-integrated TFs in DH10b and BL21(DE3)_marionette strains. The output of (C) sfGFP signal, (D) mCherry signal, (E) BFP signal, and (F) Abs600. Both engineered strains were grown in LB medium at 34°C and 600 rpm. To induce expression of fluorescent proteins, a combination of appropriate inducers was added into growth medium after 1 h of outgrowth: 100 µM cuminic acid, 10 µg/ml aTc, and 100 µM vanillic acid. After 6 h post-induction, the fluorescent signal and Abs600 were measured and the FI/Abs600 output was calculated. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

The functional characterisation results showed an expected pattern of gene activation in response to the addition of cognate inducers (Figures 2.13C, 2.13D, and 2.13E). Higher gene expression in BL21(DE3)_marionette strain was only observed for sfGFP, while the other

genes showed higher expression in DH10b_marionette strains. This can be explained by the performance of the corresponding inducible promoters in both marionette strains. For instance, the activity of $P_{VanCC\_BASIC(2)}$ was higher in DH10b_marionette strain than BL21(DE3)_marionette strain (see Figures 2.11B and 2.12B). A slight leaky expression of mCherry driven by $P_{CymRC\_BASIC(3)}$ was observed in both engineered strains (Figure 2.13D). In contrast, its characterisation in Section 2.10 showed a low background expression with high fold activation (>40-fold). Reduced expression of fluorescent proteins was noticed in the presence of multiple inducers (i.e. aTc and Van). This may be caused by cross-reactivity of each inducer and inducible system as highlighted in a previous study[80]. Interestingly, this independent multi-gene system assembled in a medium plasmid copy number (a p15A backbone) and multiple marionette strains did not significantly impair growth (Figure 2.13F).

2.13 Standardised inducible promoters in a high copy plasmid exhibit relatively tight control of multi-gene expression

The copy number of a TF and its target promoter had an important role in modulating the response of the standardised inducible promoters (Sections 2.8 and 2.9). With the use of marionette strains, the abundance of genome-integrated regulatory proteins is maintained at a low level and occupies only 1.3% of the host cells proteome[80]. This abundance of TFs has been shown to tightly control multi-gene expression in medium copy number plasmids (Section 2.12). We next examined the ability of a marionette strain (i.e. DH10b_marionette strains) to regulate multi-gene expression from a higher copy number plasmid. A three-transcriptional unit system previously constructed (Figure 2.13A or 2.14A) was placed in two different plasmids with different copy numbers: a medium copy number plasmid (p15A, ~10 copy per cell) and a high copy number plasmid (pMB1, ~15-20 copy per cell). The assembled plasmids were expressed in *E. coli* DH10b_marionette strains (Figure 2.14B).

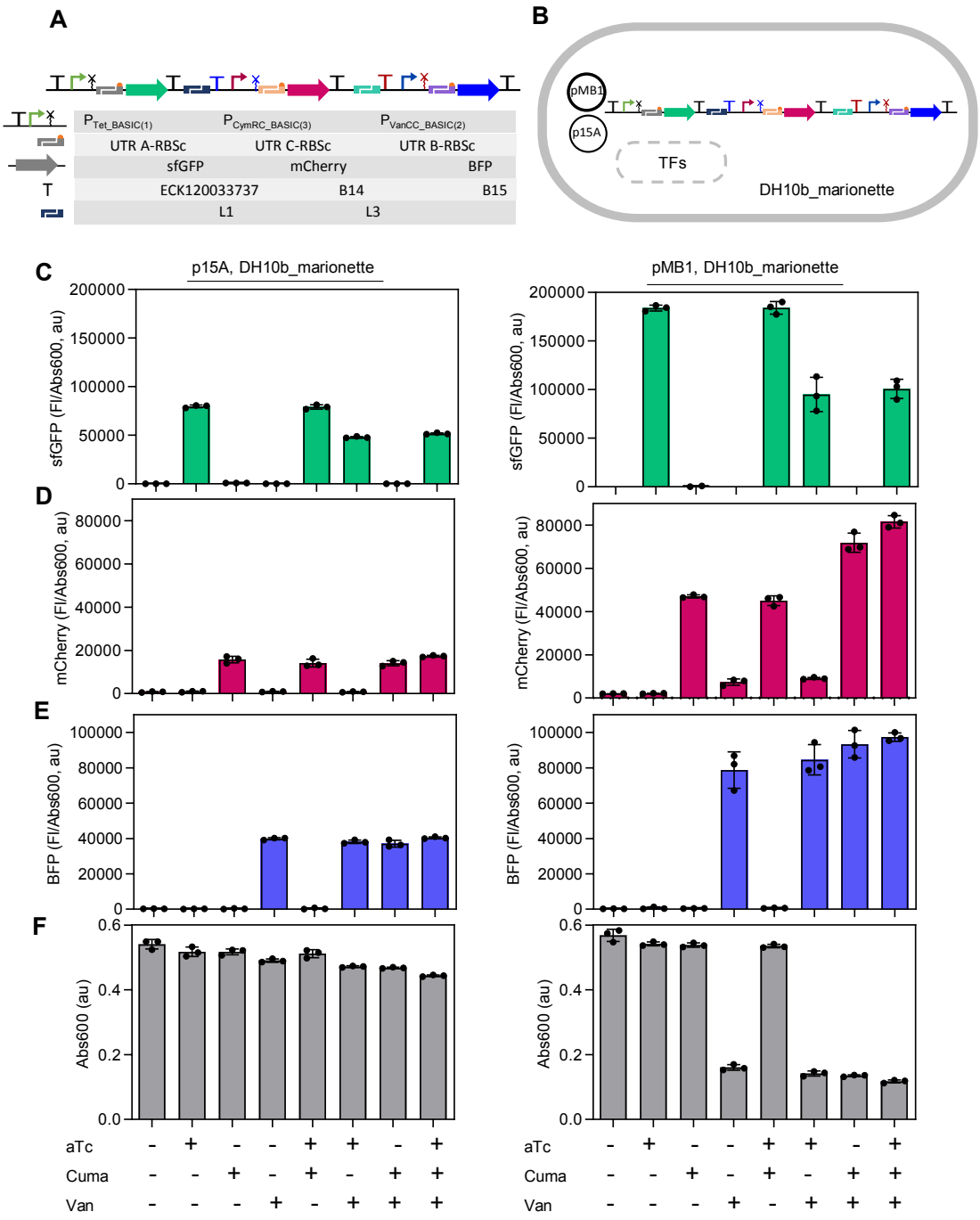


Figure 2.14 Multiplexed and independent regulation of the multi-gene expression system in different plasmid copy numbers. (A) Schematic design of genetic circuits with three output fluorescent reporters. Each genetic element is indicated. (B) The expression of the circuit is assembled in a p15A and pMB1 backbones. These assembled plasmids are controlled by genome-integrated TFs, especially DH10b_marionette strains. The output of (C) sfGFP signal, (D) mCherry signal, (E) BFP signal, and (F) Abs600. The engineered DH10b_marionette strains were grown in LB medium at 34°C and 600 rpm. To induce expression of fluorescent proteins, a combination of appropriate inducers was added after 1 h outgrowth into growth medium: 100 µM cuminic acid, 10 µg/ml aTc, and 100 µM vanillic acid. After 6 h post-induction, the fluorescent signal and Abs600 were measured and the FI/Abs600 output was calculated. Data are shown with error bars for the mean ± SD of triplicate measurements (black dots).

Consistent with the previous observation (Section 2.12), the activation responses of the reporter genes were in accordance with the encoded logic (Figures 2.14C, 2.14D, and 2.14E).

Stronger expression was observed for high copy number plasmids. A low background of gene expression was also observed. This result shows that the standardised inducible promoter system in a high plasmid copy number can be tightly regulated by genome-encoded TFs. Marked growth retardation was observed in the strains with BFP expression (Figure 2.14F). It should be noted that the BFP used in this construct is an enhanced monomeric BFP, mTagBFP[187]. The over-expression of this reporter protein at a high level may pose a toxicity issue in the DH10b_marionette strains.

Our results show that for single and multi-gene regulation using standardised inducible promoter sets, our bioparts can be tightly controlled, are compatible and portable between strains. By using strains embedding TFs in their genomes, the implementation of multi-gene regulation using standardised inducible promoter set is simplified. It also removes the need to include large regulatory cassettes in the construction of the synthetic circuits.

2.14 Summary

The promoter is an essential regulatory element required to initiate gene expression in genetic circuits. Due to its importance, remarkable efforts have been devoted to eliminating context-dependency of a promoter from adjacent DNA sequences. The work to standardise, modularise, and characterise bacterial promoters was expanded in this chapter. The structure of the 5′ and 3′ elements surrounding of core promoter were first identified and systematically assessed. By considering output activity of tested promoters and their flexible integration into a modular assembly method (i.e. BASIC DNA assembly[62]), the selected components of our standardised promoters were a terminator, UP element, core promoter, and ribozyme insulator. The variation of DNA sequences upstream and downstream of core promoters had a major contribution towards the consistency of their activity. Three alternative groups of standardised promoter design were produced varying terminators and ribozyme insulators while retaining the same UP element (UP B with 20 bp) for a collection of constitutive and inducible promoters in *E. coli*. In total, 24 constitutive promoters and 21 inducible promoters were standardised and characterised. The induction response of standardised inducible promoters was modulated by the copy number of transcription factor

and regulated promoter. Furthermore, these inducible promoters were then implemented and characterised in cloning (DH10b) and production (BL21(DE3)) marionette strains which are equipped with cognate transcriptional factors integrated into *E. coli* genome[80]. The standardised inducible promoter sets were used to control expression of three fluorescent reporters in multiple transcriptional units with different plasmid backbones. This further demonstrated compatibility, portability, and tight control of the standardised inducible promoter libraries. Taken together, the collection of bacterial promoters developed in this chapter provide a valuable, standardised, and modular biomolecular regulatory toolbox at the transcriptional level for bacterial synthetic circuits. In addition, this set of promoters can be readily integrated into a modular assembly framework, greatly supporting their use for rapid prototyping of modular synthetic genetic circuits with various designs.

2.15 Declaration

In this chapter, selection of key elements of the standardised promoter architecture including terminator, UP element, core promoter (constitutive and inducible sets) and ribozyme insulator was discussed together with Dr Marko Storch. The UP elements with 45 bp length (UP A and C) were designed using R2oDNA Designer[34] by Dr Marko Storch. The initial regulator cassette for Tet inducible system, P_{J23116}-TetR was designed and created as a BASIC biopart by Dr Marko Storch.

Chapter 3

Design and characterisation of modular Artificial RNA interference (mARi) system for robust gene regulation at a post-transcriptional level

Portions of this chapter are reproduced from:

Dwijayanti, A., Storch, M., Stan, G.-B. & Baldwin, G. S. A modular RNA interference system for gene regulation and robust dynamic feedback control. *bioRxiv* 2019.12.12.873844 (2019). doi:10.1101/2019.12.12.873844

3.1 Introduction

In the previous chapter, the investigation of core elements to create a standardised transcriptional regulator was outlined. The developed design of a standardised promoter was then applied to constitutive and inducible promoter sets. This set of standardised promoters serves as a valuable bacterial regulatory toolbox that allowing easy implementation of interchangeable biomolecular controllers within a modular assembly framework. In addition to regulation at a transcriptional level, the development of bacterial controller toolbox is continued in this chapter with a focus on modular gene regulation at a post-transcriptional level.

RNA-RNA interactions are known to provide rapid signal propagation and faster degradation rates compared to those of proteins[99–101]. These reasons have drawn interests in the use of RNA-based regulators as a versatile tool for controlling synthetic circuits, especially the use of *trans*-encoded sRNA. In this chapter, the modular *trans*-encoded sRNA is employed for down-regulating the expression of target GOIs.

The majority of gene regulation at a post-transcriptional level relies on RNA-RNA interaction at specific regions of a transcript[102–104,106–108]. The previous approach to implement targeted sRNA, especially in the TIR was exemplified in the use of native sRNAs and their cognate targets i.e. *rpoS*, *hns*, *sodB*, and *ompC*[188,189]. These targets were obtained as a leader sequence from natural RNAi systems and fused to a reporter gene (i.e. GFP and mCherry). Therefore, the performance of the native RNAi system can be quantified from the fluorescence output.

In addition to this, sRNA has highly composable structure that consists of two functional domains: a seed sequence and sRNA scaffold (see Section 1.6.3 and Figure 3.1A). Hence, a new sRNA can be re-programmed to target any RNA sequences beyond its natural targets by the modification of the seed sequence while retaining the native sRNA scaffold[102,107,188]. We term this new type of sRNA as artificial RNAi system. Several rules and guides have been developed to improve the rational design of a set of novel RNAi systems[106,107,190,191]. This

strategy has been implemented to create artificial RNAi systems with bespoke designs, where the sRNA is targeted to bind the first 24 bp of the GOIs[192,193].

Recently, artificial sRNA has also been employed to specifically bind a set of standardised target sequences that are inserted upstream of GOIs[104]. In that work, six orthogonal targets and corresponding sRNAs are used in multi-gene systems, especially to optimise a metabolic pathway (β-carotene) and genetic circuit (XNOR logic gate). However, the insertion of target sites upstream of each GOI may impede the implementation and adoption of the RNAi-based controllers for broader applications.

To overcome this issue, the development of a modular Artificial RNAi (mARi)-based controller is presented here. The mARi-based regulator is developed using a modular design approach to improve its modularity. Specifically, the mARi-based controllers were created as reusable bioparts. As previously described in Section 1.5, the modular assembly process using BASIC DNA assembly[62] is driven by the use of standardised linkers between interchangeable DNA parts (Figures 3.1A and 3.1B). A UTR-RBS linker acts as a connector between a standardised promoter (see Chapter 2) and GOIs in an assembled expression cassette (Figure 3.1B). The transcription of this expression cassette results in an mRNA containing the standardised 5′ UTR-RBS sequence in its TIR. Thus, a standardised 5′ UTR-RBS sequence can be utilised as a modular target for post-transcriptional regulation by mARi. RNA-RNA interactions can be used to cause gene silencing mediated by recruiting Hfq (Figure 3.1C).

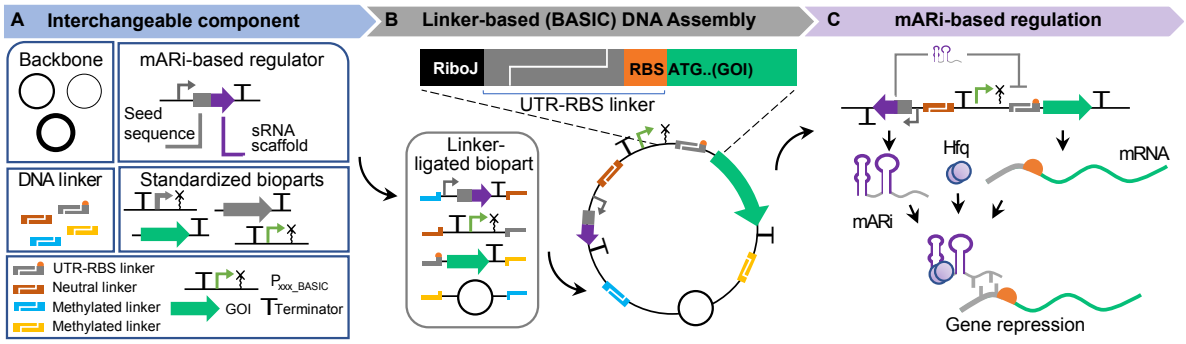


Figure 3.1 The concept of mARi-based gene regulation integrated into the modular design approach. (A) The mARi-based regulation is constructed from a library of interchangeable parts: standardised backbones, bioparts, reusable DNA linkers, and a mARi-based regulator. The transcriptional unit of the mARi-based regulator is composed of highly modular components: a promoter, seed sequence, sRNA scaffold, and terminator. (B) The standardised backbones, bioparts, and mARi-based regulator are then ligated into orthogonal DNA linkers that determine the assembly order. The linker-ligated parts were then assembled in one-pot assembly to create a functional expression unit. In this functional expression cassette, a UTR-RBS BASIC linker was used to connect a standardised promoter and GOI downstream. (C) The transcription of the mRNA of interest contains this UTR-RBS BASIC linker in the 5′ end and can be targeted by the transcribed mARi in the system. Thus, led to repression of gene expression mediated by Hfq recruitment.

The integration of mARi-based regulation into a modular DNA assembly method offers a simple yet powerful strategy for implementing RNAi regulatory systems. Once these reusable UTR-RBS linkers are used in the assembly process, specific controlling gene expression can be obtained in the presence of cognate mARi expression. Given the established mARis design, their post-transcriptional regulation properties were then characterised in various genetic contexts, including transcript ratio, molecular copy number, spatial organisation, host strain, and growth phase. The mARi variants were further expanded and their potential implementation were exploited for simultaneous and multiplexed gene regulation in a multi-gene system with different gene arrangements.

3.2 The Hfq-dependent MicC sRNA as a case study for developing modular Artificial RNA interference (mARi) system

To implement the concept of modular gene regulation at the post-transcriptional level, *trans*-encoded sRNAs were chosen as a basis for the gene silencing approach. As aforementioned in Section 1.6.3, *trans*-encoded sRNAs require the Hfq chaperone[105,194]. This chaperone protein is highly conserved and abundantly available in a wide range of organisms[120,123–125,195,196]. The abundance of this native machinery also makes *trans*-acting sRNAs as an attractive yet portable regulator. Moreover, the use of the Hfq binding site to recruit the Hfq

chaperone has previously been shown to increase the silencing performance of artificial RNAi-based gene repression systems[190]. An Hfq-dependent sRNAs with an Hfq binding site was therefore employed in the development of mARi. One of the intensively-studied Hfq-dependent RNAis, MicC was chosen from sRNA scaffolds naturally found in *E. coli*[121] and *Salmonella enterica*[108]. The simplified structure of MicC targeting OmpC is illustrated in Figure 3.2A.

The nature of the multi-targeting of MicC sRNA (see Section 1.6.3) has inspired repurposing of MicC sRNA to target any mRNA sequence of interests. Because of its flexible design, the seed sequences can be modified and directly fused to the 5' end of the MicC sRNA scaffold while preserving the native MicC scaffold[106,107,192]. This general strategy was used to develop the mARi design (Figure 3.2B). These core structures are followed by a native Rho-independent transcriptional terminator, T1/TE. The mARi design is then preceded by an interchangeable promoter located upstream of the seed sequences to create a functional expression cassette (Figure 3.2C).

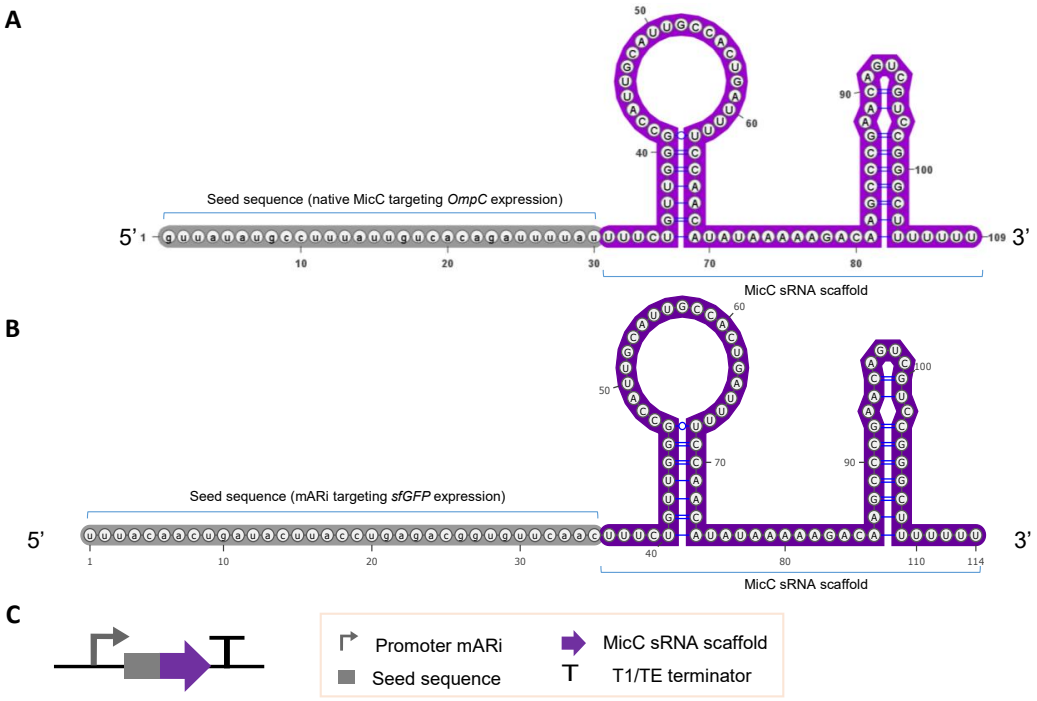


Figure 3.2 Design of mARi-based regulator. (A) Secondary structure of *Salmonella* MicC sRNA targeting OmpC (adapted from Pfeiffer *et al.*, 2009 with permission (see Appendix A)). (B) Secondary structure of a representative mARi with a 35 bp seed sequence. (C) Schematic of a mARi transcription unit consisting of a promoter, seed sequence, MicC sRNA scaffold, and Rho-independent terminator. Natural seed sequence or modified seed sequence shown in grey and MicC sRNA scaffold is coloured purple.

3.3 Selection of modular target site positions within mRNAs of interest

In silencing its target, mARi-based regulation relies on specific base-pairing interactions between a seed sequence of mARi and a target sequence in the mRNA of interest. The next key step in developing mARi is the identification of target sequence positions within an mRNA target.

In this work, the mARi was designed to target the TIR of mRNA. The silencing mechanism is thought by interfering with translation initiation in this region (see Section 1.6.3). It has also been hypothesised that the AU-rich regions at the 5' end are preferred for Hfq binding[105]. Secondly, the target located in this region has been shown to exhibit a low off-target effect in *trans*-encoded RNAi-based regulation[197], which is important for regulator specificity/orthogonality. Thirdly, this target site location offers potential modularity since it is independent of different GOIs. Ultimately, targeting mARi to bind the TIR of mRNA could potentially improve the modularity in the 5' UTR through the use of reusable DNA linkers in a modular design framework BASIC DNA assembly[62].

To realise the modular concept of gene expression and regulation, a functional unit of gene expression was then created as a target for mARi-based gene repression. This expression cassette was built using BASIC DNA assembly[62] and consisted of these following interchangeable bioparts: (i) a standardised constitutive promoter with medium strength ($P_{J23101\_BASIC(1)}$), (ii) a *sfgfp* gene as a fluorescent reporter, and (iii) a p15A backbone (Figure 3.3A). A UTR-RBS linker with medium RBS strength i.e. UTR A-RBSc, was used to connect the standardised constitutive promoter (Section 2.6) and *sfgfp* reporter gene (Figure 3.3B). The use of the UTR-RBS linker in this expression cassette provided a standardised 5' UTR sequence which could serve as an ideal target for modular gene repression.

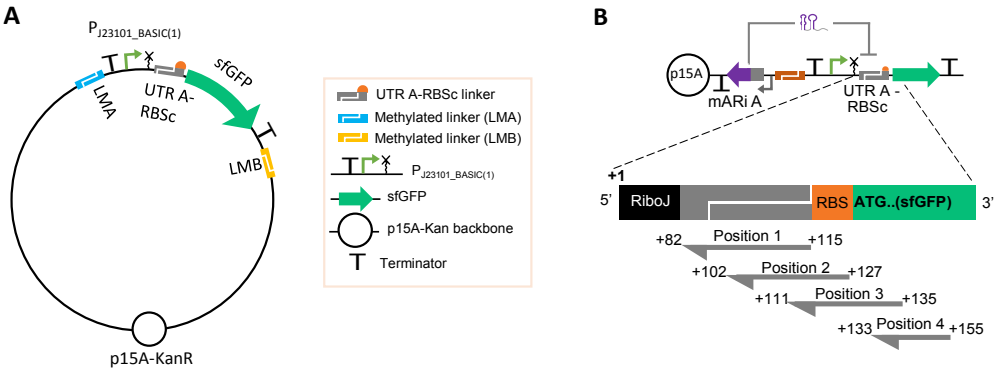


Figure 3.3 The construction of a sfGFP expression cassette used as target repression for mARi-based regulation. (A) A map of plasmid expresses sfGFP under the control of a standardised constitutive promoter, $P_{J23101\_BASIC(1)}$ and UTR A-RBSc in a p15A backbone. (B) The TIR of the sfGFP cassette was divided into four positions of target sites: Position 1 specific to 5' UTR upstream RBS; Position 2 covers 5'UTR and RBS; Position 3 covers 5'UTR, RBS, and start codon; and Position 4 specific to the early codons of *sfgfp*. A summary of mARi variants containing seed sequence that complement to different position of target sites is provided in Table 3.1.

The TIR of this expression cassette was subdivided into four different positions to facilitate the modularity of mARi (Figure 3.3B and Table 3.1). Position 1 consists of a specific sequence upstream of the RBS sequence, thus compatible with all RBS types and GOIs. This target site location has been investigated as a native location of MicC sRNA targeting OmpC[121] and recently used as an optimum target site location for artificial RNAi-mediated gene regulation[104]. A sequence that covers upstream of the RBS sequence and the RBS sequence itself was indicated as Position 2. Position 3 contains the sequence upstream the RBS sequence, the RBS sequence and the start codon. The mARi targeting this position can be used to regulate any gene of interest with ATG as a start codon. Position 1, 2 and 3 are independent of the GOIs context and specific to the UTR-RBS BASIC linker sequences. Positions 2 and 3 were included as up to 90% suppression of LacZ expression was achieved in this region[103]. Position 4 starts from the start codon (ATG) and covers up to 21 bp downstream of the proximal coding sequence. This position is specific to the early codons of the GOI (i.e. *sfgfp*) and independent of the different types of UTR-RBS linker sequences. This position has been reported as an effective recognition site for artificial RNAi-based regulations, in which the MicC sRNA scaffold was employed[106,107,192]. A summary of selected modular target sites is listed in Table 3.1.

Table 3.1 Summary of mARi variants targeting UTR A used to evaluate the different position of the target sites.

Parameter	mARi A targeting Position 1	mARi A targeting Position 2	mARi A targeting Position 3	mARi A targeting Position 4
Context dependency[a]				
Specific to the upstream region of UTR sequence	Yes*	Yes*	Yes*	No
Specific to the RBS sequence	No*	Yes	Yes	No*
Specific to the GOI sequence	No*	No*	No*	Yes
Seed sequence (5'→3')[b]	GTTGAACACCGTCTCAGGTAAGTATCAGTTGTAAA	GTATCAGTTGTAAAAAGAGGGGAAAT	GTAAAAAGAGGGGAAATAGTCCATG	ATGCGTAAAGGCGAAGAACTGTT
Length (bp)	35	26	25	23
GC content (%)	40	34.62	40	43.8
Free binding energy (kcal/mol)[c]	-55.2	-36.8	-36.7	-35.7
mARi sequence[d]	TTTACAACTGATACTTACCTGAGACGGTGTTCAACtttctgttgggccattgcattgccactgatttttccaacatataaaaagacaagcccgaacagtcgtccgggctttttttt	ATTTCCCCTCTTTTTACAACTGATACttctgttgggccattgcattgccactgattttccaacatataaaaagacaagccgaacagtcgtccgggctttttttt	CATGGACTATTTCCCCTCTTTTTACtttctgttgggccattgcattgccactgattttccaacatataaaaagacaagcccgaacagtcgtccgggctttttttt	AACAGTTCTTCGCCTTTACGCATtttctgttgggccattgcattgccactgattttccaacatataaaaagacaagcccgaacagtcgtccgggcttttttt

[a] The dependency of the target site position towards different contexts is highlighted. '*' denotes highly desirable for modularity. [b] Seed sequence of mARi cassettes is a reverse complementary sequence of cognate target sites within the UTR-RBS BASIC linker[62] used in a target expression cassette. [c] The free binding energy of the base-pairing region was calculated at 37˚C (see Section 6.4.3). [d] Seed sequences are shown in uppercase and MicC sRNA scaffold is shown in lowercase.

3.4 *In silico* evaluation of the designed mARi sequences

Having identified four positions of target sites, seed sequences were designed for the mARis. The seed sequence is a reverse complementary sequence of the cognate target site within the defined position of the mRNA of interest. The design of the seed sequence and full sequence of each mARi variant was computationally evaluated to satisfy the general requirements for their effective repression (Section 6.4.3). Firstly, the designed sequence for each mARi was made compatible with the BASIC DNA assembly[62], specifically avoiding the forbidden restriction site *Bsa*I. Secondly, to achieve effective repression and to have a high affinity binding between the seed and target sequence, the GC content and binding energy of the seed sequence were calculated. This varied in a range of 34.62 to 43.8% and -35.7 to -55.2

kcal/mol, respectively (Table 3.1). The binding energy has a positive correlation to binding affinity and observed repression capability[106,190]. The previous studies suggested the typical binding energy of -30 to -40 kcal/mol[106,107] for an effective repression activity. However, lower binding energy (less than -40 kcal/mol) has been experimentally validated and suggested to achieve higher repression due to a more thermodynamically favourable interaction[190].

In addition, the mARi variants have a seed sequence of 23-35 bp in length (Table 3.1). Although a target site with a minimum length of less than 20 bp[103,104] shows significant repression, a longer target site is preferable for increased silencing activity[190,191]. The longest seed sequence length is 35 bp as this is the limit before the seed sequence reaches the start of a BASIC biopart sequence, leading to a less modular regulation. It should be noted that the number of base-pairings also determines mARi specificity and target discrimination[190]. Shorter binding sites have a lower binding affinity and may unexpectly bind to other genes (causing off-target effects)[106,107], while a longer binding site may create a complex secondary structure which is less favourable for base-pairing with the mRNA target.

To test this, four chosen seed sequences were combined with the MicC sRNA scaffold, resulting in a set of full mARi sequences. The RNA-RNA interaction of these full mARi sequences and mRNA targets were then simulated using IntaRNA[198,199]. All mARi designs showed perfect complementary base-pairing towards mRNA of interest (i.e. *sfgfp*) in their seed sequences (Figure 3.4).

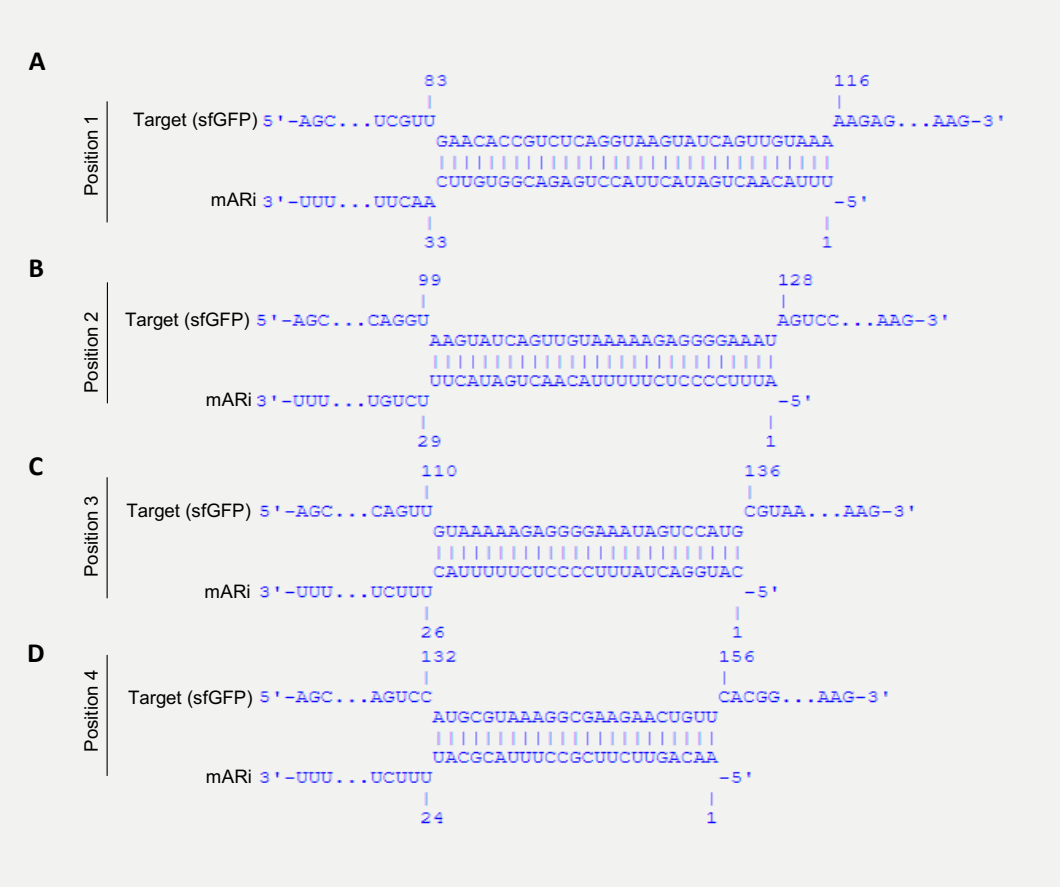


Figure 3.4 Simulation of RNA-RNA interaction between mRNA target (i.e. sfGFP) and mARi with a seed sequence in (A) Position 1, (B) Position 2, (C) Position 3, and (D) Position 4. The simulation was performed using IntaRNA[198,199]. The numbers show the relative position of bases from the TSS (+1).

3.5 Position of target sites within Translational Initiation Region (TIR) dictates the modularity and repression activity of mARi

Subsequently, a set of mARi targeting UTR A (hereafter, mARi A) at different positions were expressed under the control of a medium constitutive promoter P_{J23101} (Figures 3.5A and 3.5B). The aim of this test was to obtain a similar intracellular transcript ratio of mARi and mRNA. It should be noted that the promoter set driving mRNA expression is a standardised constitutive promoter (see Section 2.6) containing a ribozyme insulator (i.e. RiboJ) to provide a standardised 5' UTR. In contrast, the promoter controlling mARi expression is lacking any ribozyme insulator upstream of the seed sequence. This is intended to avoid extra bases in the 5' end that may reduce the repression activity of mARi. This mARi expression cassette was then placed in the opposite direction to the mRNA expression cassette and assembled on a single plasmid system with a p15A backbone. Apart from ensuring both mARi and mRNA are

present at the same copy number, clustering of these transcriptional units would increase the likelihood of base-pairing. Moreover, it is generally accepted that in bacterial systems mRNAs are translated soon after they are transcribed[200].

As a control to test the efficacy of mARi with four different target positions, a plasmid harbouring a sfGFP expression cassette without mARi expression was used. These expression plasmids were placed in *E. coli* DH5α cells. The mARi repression activity can be measured by the reduction of sfGFP expression compared to the control plasmid without mARi expression by flow cytometry measurement. Uniformity of representative expressing strains is provided in Appendix B. The mARi expression activity is presented as normalised fluorescence by dividing the fluorescence intensity of the sample with mARi by the sample without mARi expression.

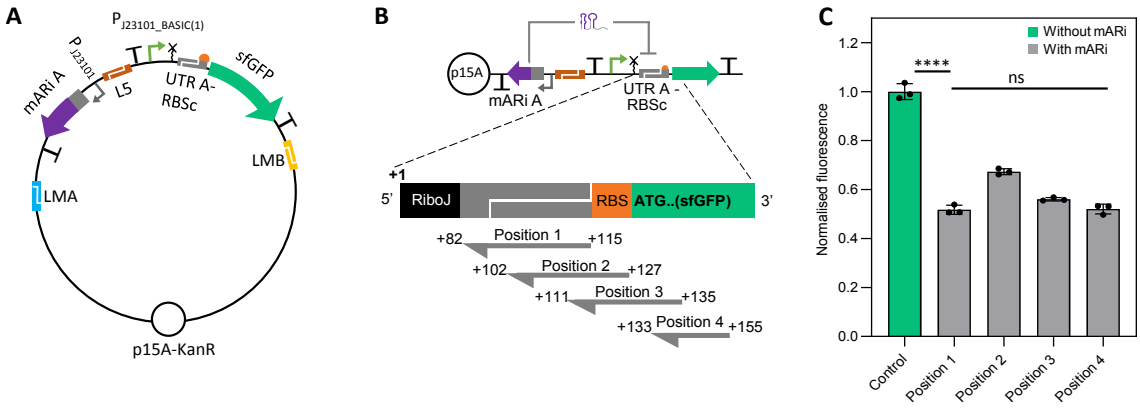


Figure 3.5 Characterisation of mARi-based gene repression targeting different position of target sites. (A) A plasmid map of mARi-based regulation constructed in a single plasmid system. (B) Schematic design of target sites selection of four different positions within TIR. (C) The repression activity of mARi A targeting these four different positions of target sites. The DH5α cells were then grown in LB media at 37°C and 600 rpm. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD. Statistically significant differences determined using Student's t-test (**** represents p<0.0001, ns represents not significant).

The initial testing of mARi expression targeting the TIR resulted in 33-50% repression in sfGFP (Figure 3.5C). As expected, target Position 1 with 35 bp in length and with the lowest free binding energy (-55.2 kcal/mol) showed significant repression activity. Significant repression was also observed in target Position 4, which is in agreement with previously reported results[106,192]. The repression activity in Position 4 did not have statistically significant difference to that in Position 1. Target Position 4 can be employed as a leader sequence when it is fused or inserted to a reporter gene or other GOIs. Such a strategy to insert target

sequence[104,201] is less desirable from an engineering perspective, especially in the modular approach since the bespoke design and optimisation are required for each new expression cassette. Aiming to improve the modularity of gene regulation at the post-transcriptional level, target Position 1 was chosen as an ideal modular target site. In addition, this target position in the 5′ UTR region is independent of both the RBS and GOI sequences, suggesting that the 5′ UTR upstream of the RBS sequence within UTR-RBS BASIC linker sequence[62] can be used as a target site for developing mARi further.

3.6 mARi exhibits relatively constant repression activity across a variety of RBS sequences

The mARi targeting the 5′ UTR upstream of the RBS sequences (target Position 1) was hypothesised to be independent from the GOI sequences and does not overlap with the RBS sequence. Additionally, the Hfq binding site would be located in the opposite direction of the RBS sequences when base-pairing occurs. This further facilitates the modularity of mARi making it independent of RBS sequence. To confirm this, the repression activity of mARi was experimentally tested when regulating sfGFP expression driven by different RBS sequences.

Altering the RBS sequence and strength in a genetic construct is one of the general strategies used for fine-tuning gene expression. In the BASIC assembly framework[62], the RBS sequence is embedded in the suffix-linker section of the UTR-RBS linker. Replacing the RBS sequences to tune gene expression can be easily performed by combining the existing prefix-linker section to a set of the suffix-linker sections with different RBS sequences (Figure 3.6A). Therefore, a set of sfGFP expression cassettes controlled by $P_{J23101\_BASIC(1)}$ and UTR A with five different RBS sequences were selected (Table 3.2). These expression cassettes were then characterised and compared to their predicted output. The mean relative expression strength covered in this library in a range from 0.18-1. Surprisingly, this observed mean relative expression is different from the predicted strength using RBS calculator v2.0[31,32] and EMOPEC[33] (Section 6.4.4).

Table 3.2 Comparison of experimental data and predicted of RBS performances within UTR-RBS BASIC linker sequences.

UTR-RBS linker	Sequence[a]	Mean relative expression	Predicted strength	
			RBS calculator v2.0[31,32]	EMOPEC[33]
UTR A-RBSa	ggctcgttgaacaccgtctcaggtaagta tcagttgtaa**aaagaggagaaa**tagtcc	1.00	715277.31	77.8%
UTR A-RBSb	ggctcgttgaacaccgtctcaggtaagta tcagttgtaa**atctaaggaggg**tagtcc	0.52	673992.76	80.5%
UTR A-RBSc	ggctcgttgaacaccgtctcaggtaagta tcagttgtaa**aaagaggggaaa**tagtcc	0.35	689020.88	79.5%
UTR A-RBSd	ggctcgttgaacaccgtctcaggtaagta tcagttgtaa**atcccaggaggg**tagtcc	0.23	17315.58	90.5%
UTR A-RBSe	ggctcgttgaacaccgtctcaggtaagta tcagttgtaa**atctcgggaggg**tagtcc	0.18	715277.31	77.3%

[a] DNA sequence colours correspond to upstream scar (blue) and downstream scar (orange). The DNA sequence in bold indicates the RBS sequence.

This series of test cassettes were next used to target repression of mARi using target Position 1 (Figure 3.6A). Following the previous assembly strategy, both mARi and mRNA expression cassettes were assembled in a single plasmid system with a p15A backbone.

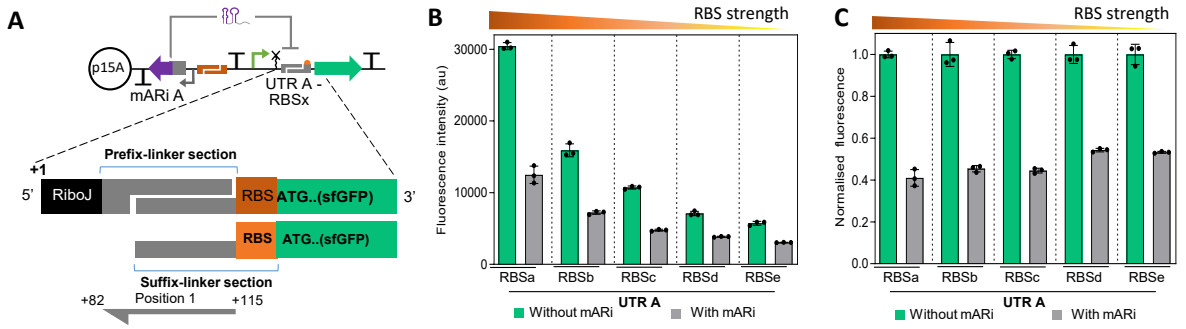


Figure 3.6 Characterisation of mARi-based regulation in silencing a library of expression cassettes with varied RBS strength. (A) A schematic diagram of expression cassettes used to characterise the repression activity of mARi against different RBS strengths. A library of RBSs embedded in UTR A (Table 3.2) was used to connect a standardised constitutive promoter with medium strength ($P_{J23101\_BASIC(1)}$) and *sfGFP* as a gene reporter. Position 1 as a target for mARi-based regulation was indicated. (B) The fluorescence intensity of sfGFP expression with varying RBSs in the presence and absence of mARi targeting Position 1. (C) The normalised fluorescence showing a relative constant repression activity of mARi with Position 1 across different RBS strength. The DH5α cells were then grown in LB media at 37°C and 600 rpm. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD.

All tested constructs displayed a reduction in sfGFP expression in the presence of mARi (Figure 3.6B). Notably, the normalised fluorescence showed relatively constant repression activity regardless of RBS strength (Figure 3.6C). This demonstrated that the mARi targeting Position 1 is modular and independent of the RBS used, which is important for the system to be used where different RBS contexts are required.

3.7 Transcript expression ratio modulates mARi efficacy

Using the established mARi design, the post-transcriptional regulation of mARi was next investigated in a number of typical genetic contexts. It should be noted that mARi relies on transient RNA-RNA interactions[202]. Two key determinants modulating the repression activity of this post-transcriptional regulator are (i) the degradation rate of RNAs complexes and (ii) the ratio of transcript abundance[202]. Although the degradation rate of mARi-mRNA target has been previously investigated and optimised (Sections 3.5 and 3.6), the effect of transcript ratio towards repression activity has not, and was examined in this section. Specifically, the concentration ratio of mARi transcripts relative to the mRNA transcript to achieve maximum repression activity was investigated.

Unlike previous work where inducible expression of transcripts was used for simulating the dynamics of transcript abundance[127,202,203], here we constitutively expressed both transcripts to provide a better approach of different transcript level in a steady state condition. As a proxy of transcript abundance at steady state, the estimation of promoter strength driving the regulator and target has been used[202]. A similar strategy was then used to control the transcript stoichiometry. A set of standardised constitutive promoters ($P_{J23xxx\_BASIC(1)}$) were used to express a *sfGFP* reporter gene using UTR A-RBSc in a p15A backbone (Figure 3.7A). These assembled plasmids were propagated in *E. coli* DH5α cells and the promoter activity calculated. Figure 3.7B shows the functional characterisation of five assembled constructs. These results were then used as a basis to infer the relative transcript expression ratio, which is calculated by dividing the strength of the promoter upstream mARi against the promoter expressing the mRNA target.

Due to the nature of RNAi regulation, an excess of regulator sRNA is preferred to create significant repression[127,202]. Therefore, a set of four stronger constitutive promoters were used to drive expression of mARi, whereas four weaker standardised constitutive promoters were used for mRNA expression. A matrix of calculated relative expression ratio values showed the relative abundance of mARi regulators against the mRNA targets (Figure 3.7C). The highest expression ratio was obtained for strong mARi expression and low mRNA

expression (Figure 3.7C, top right panel), whereas the lowest expression ratio was obtained for a low mARi expression and high mRNA expression (Figure 3.7C, bottom left panel).

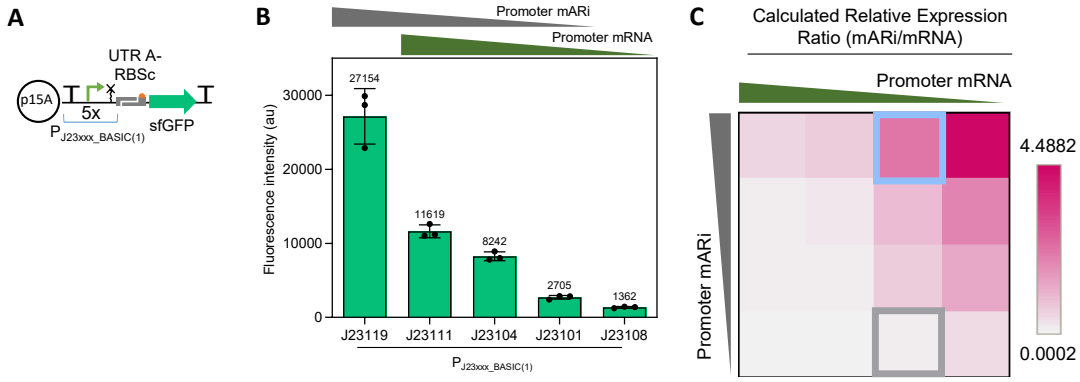


Figure 3.7 Promoter characterisation used to determine relative expression ratio of transcripts. (A) A schematic design of construct variants used to characterise standardised constitutive promoter set ($P_{J23xxx\_BASIC(1)}$). (B) Characterisation result of standardised constitutive promoter set. The DH5α cells were then grown in LB media at 37°C and 600 rpm. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD. (C) A heat-map of relative expression ratio calculated by dividing the activity of promoters driving mARi regulator over mRNA target from Figure 3.7B. Representatives of high and low expression ratios were highlighted in a blue and grey box, respectively.

In order to modulate the repression activity of mARi, different transcript expression ratios were then produced by varying the strength of both promoters driving mARi and mRNA (Figure 3.8A). Similar to the earlier described work, both mARi and mRNA expression cassettes were constructed in a single plasmid system with a p15A backbone (Figure 3.8A). A combinatorial and modular assembly from interchangeable bioparts in response to the expression ratio values in Figure 3.7C or 3.8B were then carried-out. In total, 16 constructs with different transcript ratio and four plasmids control without mARi were generated using the BASIC DNA method[62].

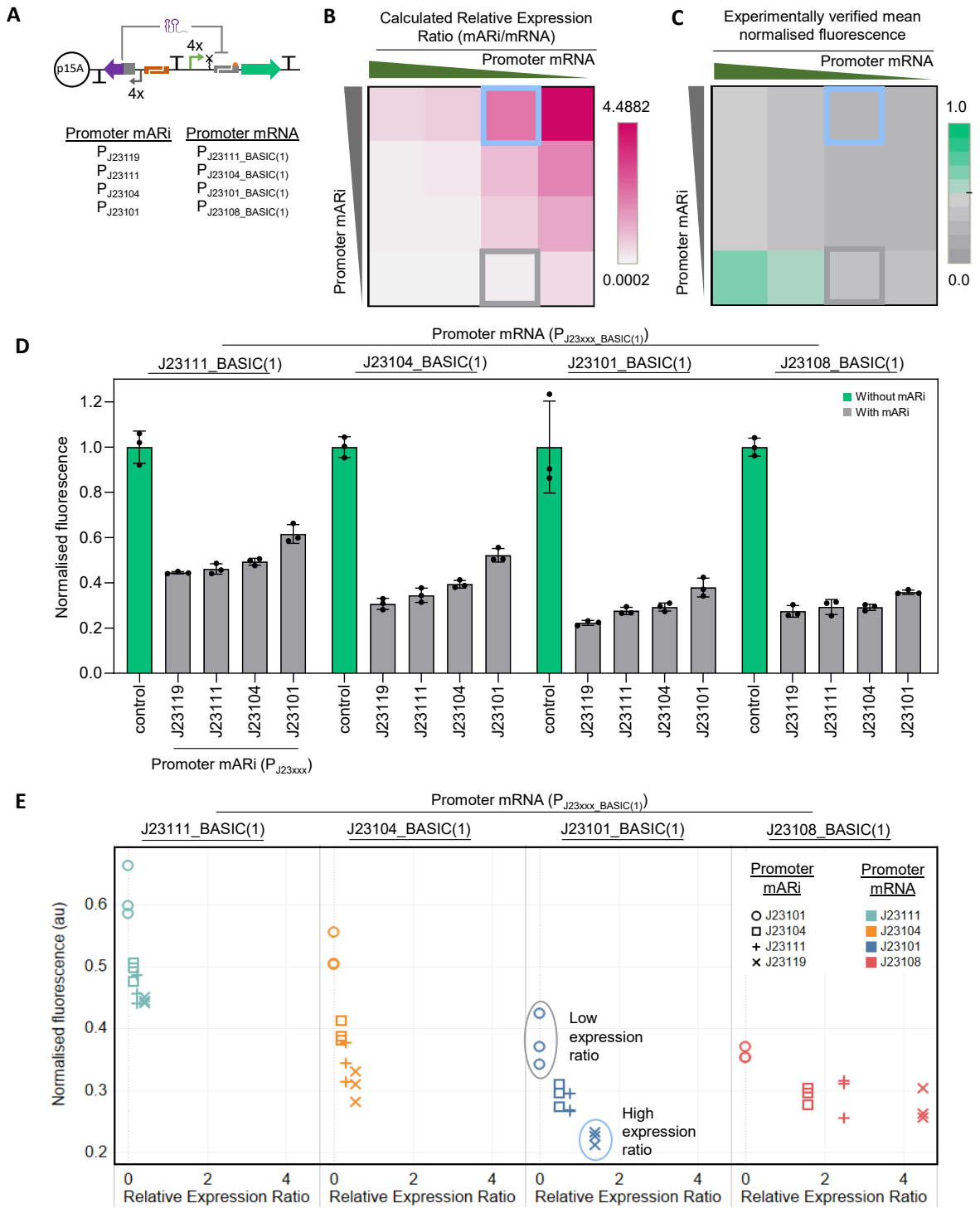


Figure 3.8 Tuning repression activity of mARi-based regulatory system through transcripts titration in a steady state expression. (A) A schematic diagram of mARi-based repression constructed in a single plasmid system. The selected promoters controlling mARi expression: P_{J23119}, P_{J23111}, P_{J23104}, and P_{J23101}; whereas promoters driving mRNA expression: $P_{J23111\_BASIC(1)}$, $P_{J23104\_BASIC(1)}$, $P_{J23101\_BASIC(1)}$, and $P_{J23108\_BASIC(1)}$. (B) A matrix of calculated relative expression ratio as a proxy of intracellular transcript ratios. (C) A heat-map showing the repression activity of mARi in response to different transcript ratios in Figure 3.8B. Representatives of high and low expression ratios were highlighted in a blue and grey box, respectively. The repression activity showed as mean normalised fluorescence towards control plasmid without mARi expression. (D) Normalised fluorescence data associated with the error bars in Figure 3.8C. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD. (E) Correlation of transcript expression ratios in Figure 3.8B and normalised fluorescence in Figure 3.8D. The high and low expression ratios were indicated and clustered. The variant of the promoter driving mARi and mRNA were indicated. In the promoter notations, x refers to the promoter number.

As shown in Figures 3.8C and 3.8D, the repression activity of mARi was modulated by both mRNA and mARi levels. In comparison to previous studies[101,127], the results obtained in this experiment provide additional insight via the two-dimensional dose response. The maximum repression was observed when an excess of mARi was present in comparison to its cognate mRNAs. On the other hand, a low expression ratio led to weak repression and partial degradation, presumably due to an insufficient amount of mARi available to inactivate mRNA target[190,202].

To gain a better understanding of the correlation between repression strength and relative transcript ratio, the measured normalised expression of sfGFP was plotted against the relative expression ratio of transcripts (Figure 3.8E). A positive correlation between repression activity and relative expression ratio was clearly shown. A graded response with relative expression ratio was also observed from the correlation graph (Figure 3.8E). A high expression ratio (Figures 3.8B and 3.8C, blue box) and low expression ratio (Figures 3.8B and 3.8C, grey box) were indicated and clustered in Figure 3.8E. Interestingly, the highest expression ratio did not result in a greater repression activity, indicating the possibility of saturated mARi.

These results confirmed that the repression activity of mARi is strongly correlated with transcript expression ratio. To create optimum gene silencing, a high expression ratio was required between mARi and the target mRNA. However, an excessive mARi expression led to slightly reduce efficacy of mARi, presumably due to binding saturation.

3.8 mARi effectiveness is further modulated by plasmid copy number

It should be noted that both mARi and the target mRNA were built in a single system for easy construction. Hence, the RNA abundance of the mARi-based regulator and target was also determined by the plasmid copy number used. Different concentrations of mARi and mRNA transcripts could further modulate the silencing activity in addition to previously investigated expression ratio (Section 3.7). To verify this hypothesis, mARi-based gene regulation systems with high and low relative expression ratios (Section 3.7) were assembled in three different plasmid copy numbers: low copy origin of replication (pSC101, ~5 copy per cell), the medium

copy origin of replication (p15A, ~10 copy per cell), and the higher copy origin of replication (pMB1, ~15-20 copy per cell) (Figure 3.9A). The assembled constructs were then propagated in *E. coli* DH5α cells.

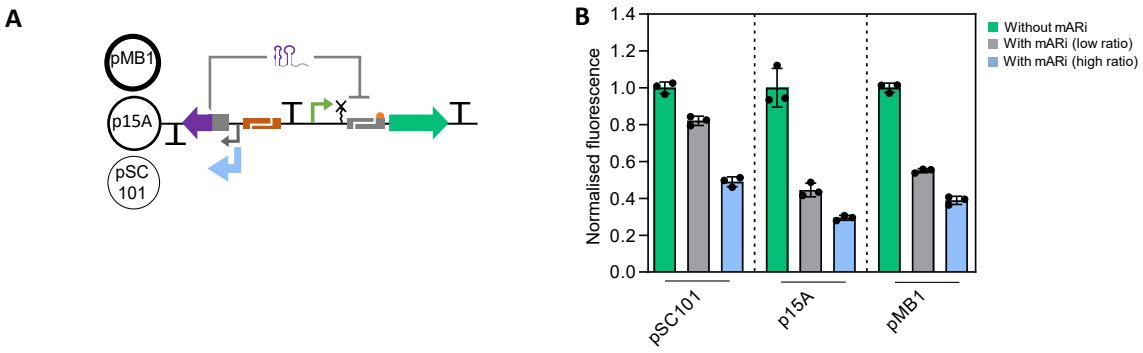


Figure 3.9 Characterisation of mARi-based regulation in a single plasmid system. (A) A schematic of mARi-based repression in a single plasmid system with a design variation of plasmid copy number and relative expression ratio of transcripts. The high and low expression ratios are shown by the use of a strong promoter (blue) and medium promoter (grey) for driving mARi. (B) The repression activity of mARi-based regulation in a single plasmid system with different plasmid copy numbers and relative expression ratios. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD.

The mARi expression either in high or low expression ratio reduced the sfGFP expression for all tested plasmid copy numbers, showing that the mARi-based regulatory system is functional in different plasmid copy numbers (Figure 3.9B). In line with previous results (Section 3.7), high expression ratios showed greater repression. Notably, the expression of mARi in a medium plasmid copy number p15A had the maximum repression activity. The use of higher plasmid copy number such as a pMB1 backbone resulted in slightly reduced inhibition compared to medium plasmid copy number. One possible explanation is due to the saturated binding of mARi-mRNA as previously observed in Section 3.7. Additionally, it should be noted that the different copy number of ColE1-based plasmids may also contribute to the reduced repression activity[70]. Interestingly, the use of mARi-based repression in a low plasmid copy number, pSC101 showed the lowest repression activity for both high and low expression ratios. This behaviour, presumably due to stochastic variation of low cellular transcript number for both mARi and mRNA. The data presented in this section revealed that mARi was functional in all type of tested plasmid backbones, despite the fact that its repression activity was further modulated by plasmid copy number.

3.9 mARi expression in a higher transcript ratio and double plasmid system resulted in slightly decreased repression activity

To further elucidate the saturated binding of mARi-mRNA, the relative transcript ratio was further increased by placing mARi cassette in a higher plasmid copy number than its cognate mRNA expression cassette. The sfGFP expression cassette driven by $P_{J23101\_BASIC(1)}$ was expressed in a p15A backbone, whereas two types of higher plasmid copy number: pMB1 and pUC were used as a backbone for mARi expression plasmids (Figure 3.10A). As a comparison, the estimated plasmid copy number of pMB1, pUC, and p15A are ~15-20, ~500-700, and ~10 copy per cell, respectively. To demonstrate the different transcript expression ratios in this double plasmid system, mARis were expressed under the control of a strong promoter (P_{J23119}) and medium promoter (P_{J23101}). It was expected that a higher expression ratio results in a stronger repression activity.

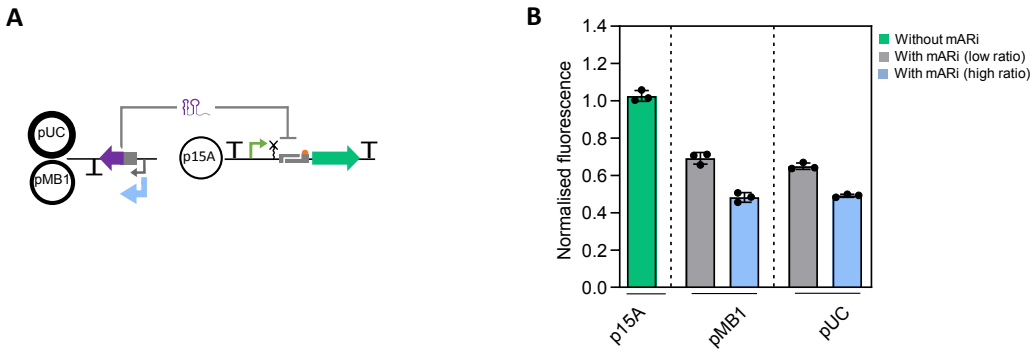


Figure 3.10 Characterisation of mARi-based regulation in a double plasmid system. (A) A schematic of mARi-based repression in a double plasmid system with a design variation of plasmid copy number and relative expression ratio of transcripts. The high and low expression ratios are shown by the use of a strong promoter (blue) and medium promoter (grey) for driving mARi. The mARi cassettes were cloned in a pMB1 and pUC backbones, whereas the mRNA cassette was assembled in a p15A backbone. (B) The repression activity of mARi-based regulation in a double plasmid system with different plasmid copy numbers and relative expression ratios. The sfGFP fluorescence measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD.

In this double plasmid system, the expression of mARi in pMB1 and pUC backbones reduced sfGFP expression in all tested case (Figure 3.10B). Consistent with the previous silencing behaviour (Sections 3.7 and 3.8), the higher transcript ratio demonstrated an increased repression activity than those observed in a low expression ratio. In this experimental design, the transcript expression ratio was described as a cumulative of plasmid copy number and transcription rate of mARi relative to mRNA. For the similar amount of mRNA target (in a p15A

backbone), the repression activity observed in the double plasmid system was lower than the single plasmid system (Section 3.8), even though it had an excess of mARi. One possible explanation is that high expression of mARi saturates the mARi-mRNA system as previously hypothesised in Sections 3.7 and 3.8.

In addition to the transcript expression ratio, the spatial organisation and heterogeneity of transcripts may also contribute to tune the post-transcriptional regulatory activity of mARi, as suggested in a prior study[204]. The fact that the RNAi-based regulator has a shorter half-life and stability than mRNA may further reduce the mARi efficacy when it is used in a double plasmid system. Compared to a single plasmid system, a double plasmid system has a broader transcript distribution.

Altogether, mARi expressed in a single plasmid system is preferable over double plasmid system because it displays a greater repression. Additionally, this strategy is beneficial for simplifying the system, especially when designing more complex genetic circuits and removes plasmid compatibility issues.

3.10 mARi expression exerts low cellular burden in a range of tested host strains

Further to the characterised post-transcriptional regulation properties examined in the previous sections, it is important to investigate the impact of the mARi system in different *E. coli* strains. Unlike the other biomolecular regulators (e.g. transcription factors, TALEs, zinc fingers, or CRISPRis), the functional mARi only requires transcriptional process without further translation. In addition, the seed sequence of mARi has been computationally validated to minimise unwanted interactions with the host backgrounds (discussed in Section 3.11). Thus, the expression of a short sequence for mARi would be expected to have a low cellular burden and not significantly impact host cell growth. However, the overexpression of mARi may consume shared post-transcriptional machinery, thereby indirectly affecting cell fitness.

The high transcript ratio for mARi in a p15A backbone (see Sections 3.7 and 3.8) and control plasmid without mARi expression were used to compare the effect of mARi overexpression on host fitness. Four different *E. coli* strains: DH5α (K-12 strain), DH10b (K-12 strain), BL21(DE3) (B strain), and BL21star(DE3) (B strain) were transformed by these test (with mARi expression) and control (without mARi expression) plasmids. To monitor the growth profiles over time, the engineered strains were grown in a plate reader with continues measurements taken (Section 6.3.1). As shown in Figure 3.11, the high expression of mARi did not affect bacterial growth across different host strains, indicating the mARi-based repression system has a low production cost, thus does not impact on host cells.

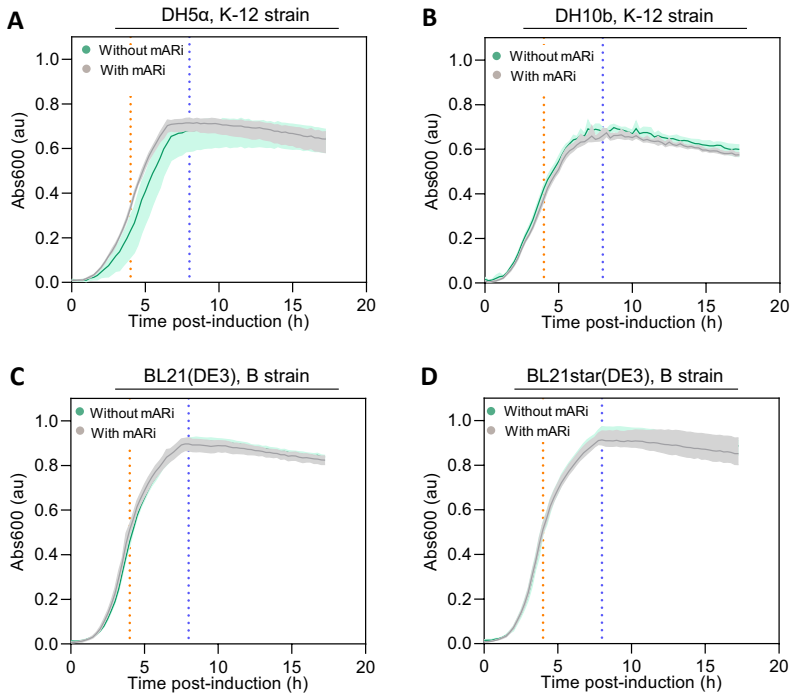


Figure 3.11 Growth profile of mARi expression strains with high expression ratio in a p15A backbone across four different tested host strains of *E. coli*: (A) DH5α (K-12 strain), (B) DH10b (K-12 strain), (C) BL21(DE3) (B strain), and (D) BL21star(DE3)) (B strain). Two different time points were selected to describe different growth phases: mid-exponential phase, 4 h (orange line) and early stationary phase, 8 h (blue line). The engineered strains were then grown in LB medium at 37°C and 600 rpm with continuous measurement in a microplate reader. Lines shows the mean from three independent replicates with shaded area showing ± SD.

3.11 mARi is portable between different host strains and growth phases

Having demonstrated the low cellular cost of mARi production in four *E. coli* strains including K-12 and B strains, the silencing activity of mARi was next evaluated. As a biomolecular controller, the performance of the regulator is often affected by the host genotype and growth[130,205]. In particular, mARi uses native stress-response regulatory machinery (Hfq chaperone and degradosome complex (i.e. RNase E) as helper proteins) for the RNAi-based regulatory system (Section 1.6.3). However, the genotype of the host strain has been known to affect the intracellular abundance of these components[205].

Functional characterisation taken at early stationary phase, 8 h (Figure 3.11, blue line), showed about 60-80% repression activity in both *E. coli* K-12 and B strains (Figure 3.12A). The silencing activity in the B strains was higher than in the K-12 strains for comparable plasmid copy numbers (Figure 3.12A). It was also previously shown in Section 3.8 that the inhibition activity of mARi was modulated by plasmid copy number in DH5α cells. Notably, this modulation effect was also observed in DH10b cells. This effect, however, was not clearly observed in B strains, suggesting the different level of the native regulatory machinery such as Hfq chaperone may play important roles in this modulation.

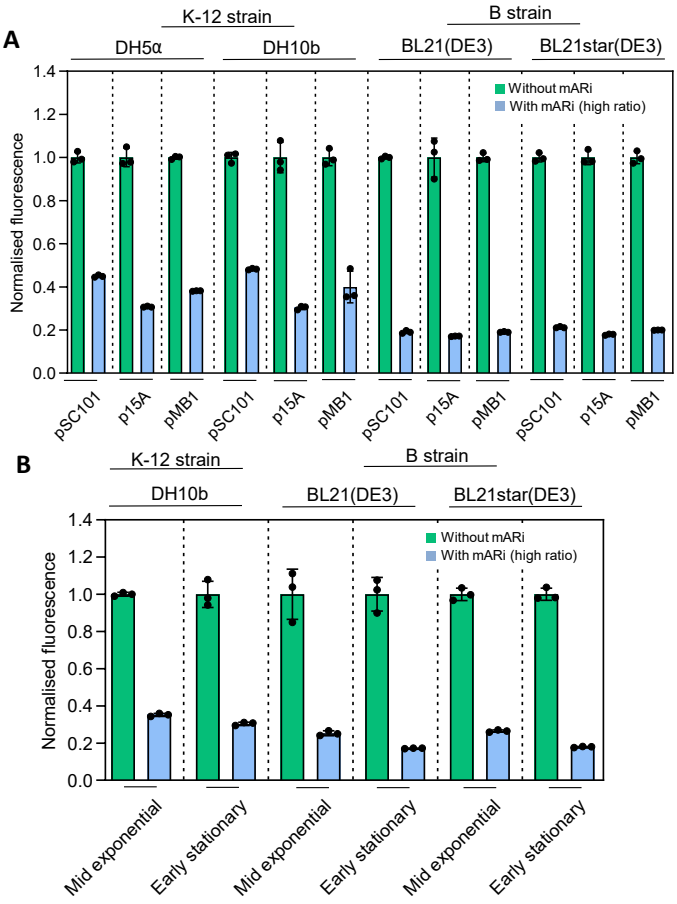


Figure 3.12 Characterisation of the mARi-based regulatory framework in different tested host strains. (A) The performance of high mARi expression in the different host strains of *E. coli* and plasmid copy numbers. Raw data associated with Figure 3.12A are provided in Appendix C. (B) The performance of high mARi expression in different growth phases. The engineered strains were grown in LB medium at 37°C and 600 rpm. Data were taken from a plate-based continuous measurement. Bars represent the average of triplicates data points (black dots) ± SD.

The intracellular concentration of these helper proteins are known to be dynamic in response to growth phase[130]. The intracellular abundance of the Hfq protein is constant during the rapid growth phase (exponential phase) and then its concentration reduces when *E. coli* cell reaches stationary phase[130,131,196]. To investigate the repression activity in response to this dynamic abundance, the mARi repression in the different growth phases of the *E. coli* strains was tested. Both K-12 and B strains with high expression ratio of mARi in a p15A backbone were monitored at different growth phases: mid-exponential phase, 4 h (Figure 3.11, orange line) and early stationary phase, 8 h (Figure 3.11, blue line). The mARi system was active across all phases of cell growth (Figure 3.12B). The *E. coli* K-12 strain (DH10b cells) showed relatively stable repression activity in both growth phases, whereas a slightly increased repression activity was observed in early stationary phases in the BL21(DE3) and BL21star(DE3) strains.

One of the natural mechanisms in RNAi-based gene repression is the sRNA-mRNA duplex degradation via RNase E[113,114,206] (Section 1.6.3). To test this silencing mechanism in the mARi-based regulation, two different production strains: *E. coli* BL21(DE3) and BL21star(DE3) strains were used. The *E. coli* BL21star(DE3) contains *rne*131 mutation lacking 477 aa of RNase E in the C terminal[207]. In that work, the *rne*131 mutation enhanced mRNA stabilisation by reducing RNase E activity. This mutant results in two-fold slower mRNA degradation compared to that in the wild-type strains[207].

There was no significant difference in repression activity observed when mARi was used in *E. coli* BL21(DE3) and RNase E-deficient *E. coli* strains BL21star(DE3) across different growth phases (Figure 3.12B). This result suggested a different mode of action in which RNase E-independent mechanism is employed for mARi-based repression. Although RNase E-dependent mechanism may lead to a more effective repression, it may also limit the flexibility of the designed sequence for artificial RNAi and target mRNA as an RNase cleavage site would need to be introduced into the design. In addition, this strategy may give higher basal degradation rates with RNase E degrading unpaired sRNA and mRNA transcripts. The mARi and UTR-RBS linker used in this study, therefore, have been designed from orthogonal and defined sequences that exclude RNase E cleavage sites. Hence, the alternative mode of action of gene repression which is independent from RNase E activity could be achieved. This is greatly beneficial for the flexibility implementing of mARi-based gene regulation in different genetic circuits and host strains, removing the need to consider the genotype of host cell.

In general, the mARi showed strong repression in all tested host strains across different plasmid copy numbers. The repression activity was relatively constant in *E. coli* K-12 strains, however, slightly increased in *E. coli* B strains. Further experimental investigation in B strains suggested that the mARi repression mechanism was probably independent of RNase E activity.

3.12 Expanding a set of orthogonal mARis through seed sequence modification

In the BASIC assembly framework[62], an UTR-RBS linker can be used to create multi-gene expression systems, such as independent expression cassettes and operon-based systems

(see Section 1.5). With its modularity, mARi-based regulation can be targeted to UTR-RBS linkers used in these multi-gene systems. This provides a simple way of specifically and simultaneously controlling multi-gene systems.

To enable this implementation, five different UTR-RBS linkers were firstly identified using R2oDNA Designer[34], whilst preserving the same medium strength RBS (RBSc). The selected 35 bp upstream of the RBS sequence were then evaluated their GC content and free binding energy of the target/seed sequences (Table 3.3). The GC content of the seed sequence in a range of 31-43% and resulted in free binding energy of -48.9 to -57.2 kcal/mol. All selected seed/target sequences showed low binding energy in which is favourable for RNA-RNA interaction[191,197] and in agreement with an effective repression activity in Section 3.3.

Table 3.3 A set of selected seed/target sequence pairs.

mARi-UTR pair	**Seed/Target sequence (35 bp)**	**GC content**	**Free binding energy (kcal/mol)**
mARi A /UTR A	TTTACAACTGATACTTACCTGAGACGGTGTTCAAC	40 %	-55.2
mARi B /UTR B	TTACAATAGATTTTACCGTCAGACCACGAGATACC	40 %	-54.9
mARi C/UTR C	TTTTCTGCTACCCTTATCTCAGCCAATAGTAACAC	40 %	-55.8
mARi D/UTR D	TTTACATAGAATACACAGCCGGGACAGGGTATAAC	42.86 %	-57.2
mARi E/UTR E	TTTACATATGTTTTATCGTCAAGACGCTGTATAAC	31.43 %	-48.9

Coordinated multi-gene regulation requires non-cross reacting (orthogonal) target/seed sequence pairs. In this context, the similarity of the sequence between target and the non-cognate target is an important factor[191,197]. Low sequence similarity of non-cognate pairs is desirable to achieve minimal cross-talk interaction and off-target effects[197]. To test this hypothesis, the identity scores of the UTR set were calculated using EMBOSS needle method (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) (Section 6.4.5). The target/seed sequence analysis resulted in about 48.6-65.8% sequence identity for non-cognate pairs (Figure 3.13). However, a high identity value (61.1-65.8%) was observed for UTR D and UTR E towards UTR A and UTR B. These high similarity scores may affect the repression activity and orthogonality of the designed mARi set.

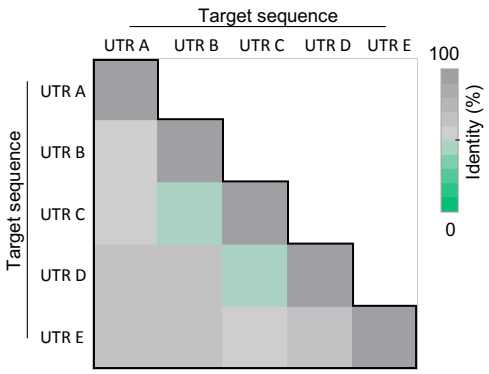


Figure 3.13 Calculation of target sequence identity similarities of 5 different UTRs. The predicted identity was computed using the EMBOSS needle method (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) (See Section 6.4.5).

Using the underlying design principles detailed in Section 3.4, a set of mARis was created by combining the selected seed sequence with the MicC scaffold. Furthermore, the possible off-target interaction of each sequence was assessed for three different *E. coli* genomes (*E. coli* MG16555, DH10b, and BL21(DE3)) using CopraRNA[199,208,209] (Section 6.4.6). The off-target prediction (Table 3.4) resulted in designs with a higher binding energy which makes them less favourable for non-cognate RNA-RNA interactions compared to the cognate mRNA target (Table 3.3).

Given the identity value of the target and non-target (Figure 3.13) as well as low scores for predicted off-target binding in the host genomes (Table 3.4), the designed mARi variants were then experimentally validated in a double plasmid system (Figure 3.14A). A set of mARi systems were cloned in a pMB1 backbone and driven by a strong constitutive promoter P_{J23119}. The cognate UTR-RBS linkers with a medium strength RBS, indicated as UTRx-RBSc, were used to connect the constitutive promoter $P_{J23101\_BASIC(1)}$ to a *sfgfp* reporter gene in a p15A backbone. *E. coli* DH5α cells were then co-transformed with these expression plasmids to examine the orthogonality of the designed mARi/UTR set.

Table 3.4 Off-target prediction of a set of orthogonal ARis towards three different *E. coli* genomes.

mARi variant	mARi sequence (5'→3')[a]	Rank	Energy (kcal/mol)	Gene annotation[b]
mARi A	TTTACAACTGATACTTA CCTGAGACGGTGTTCA ACtttctgttgggccattgcat tgccactgattttccaacatat aaaaagacaagcccgaaca gtcgtccgggctttttt	1	-27.95	RNA pyrophosphohydrolase
		2	-28.08	YhcH/YigK/YiaL family protein
		3	-25.69	16S rRNA (cytosine(1407)-C(5))-methyltransferase RsmF
mARi B	TTACAATAGATTTTACC GTCAGACCACGAGATA CCtttctgttgggccattgcat tgccactgattttccaacatat aaaaagacaagcccgaaca gtcgtccgggctttttt	1	-33.85	DUF386 family protein cupin superfamily
		2	-25.10	16S rRNA m(5)C1407 methyltransferase SAM-dependent
		3	-19.07	DUF554 family putative inner membrane protein
mARi C	TTTTCTGCTACCCTTAT CTCAGCCAATAGTAAC ACtttctgttgggccattgcat tgccactgattttccaacatat aaaaagacaagcccgaaca gtcgtccgggctttttt	1	-27.42	NADH:ubiquinone oxidoreductase membrane subunit J
		2	-27.04	Sensory histidine kinase in two-component system with NarP
		3	-27.03	DUF386 family protein cupin superfamily
mARi D	TTTACATAGAATACAC AGCCGGGACAGGGTA TAACtttctgttgggccattg cattgccactgattttccaaca tataaaaagacaagcccgaa cagtcgtccgggctttttt	1	-27.06	16S rRNA m(5)C1407 methyltransferase SAM-dependent
		2	-26.28	DUF386 family protein cupin superfamily
		3	-23.62	O-antigen capsule forming protein-tyrosine-phosphatase;Etk-P dephosphorylase
mARi E	TTTACATATGTTTTATC GTCAAGACGCTGTATA ACtttctgttgggccattgcat tgccactgattttccaacatat aaaaagacaagcccgaaca gtcgtccgggctttttt	1	-28.14	DUF386 family protein cupin superfamily
		2	-23.97	Enterobactin/ferrienterobactin esterase
		3	-21.20	16S rRNA m(5)C1407 methyltransferase SAM-dependent

[a] Seed sequences are shown in uppercase and MicC sRNA scaffold is shown in lowercase. [b] The off-target effect of the full sequence of mARis was predicted using CopraRNA[199,208,209] towards NC_000913 (E. coli MG1655), NC_010473 (E. coli DH10b), NC_012892 and NC_012971 (E. coli BL21(DE3)) as host references (see Section 6.4.6).

The expression of sfGFP was significantly repressed in the presence of the cognate mARi systems (Figures 3.14A and 3.14B), confirming their specificity with reduced cross-reactivity between each pair. Nevertheless, there is an increased expression in UTR D-sfGFP when mARi A and mARi B are used. One possible reason for this is that the high similarity score of UTR D towards both UTR A (62.9%) and UTR B (61.1%) and the distribution of similarity as shown in Figure 3.14D, may contribute to these unexpected interactions. However, further experimental investigations are required to explore this hypothesis. Overall, the resulting

mARi set generated in this section displayed specific targeting with minimal cross-talk interaction, greatly expanding the set of available mARi systems.

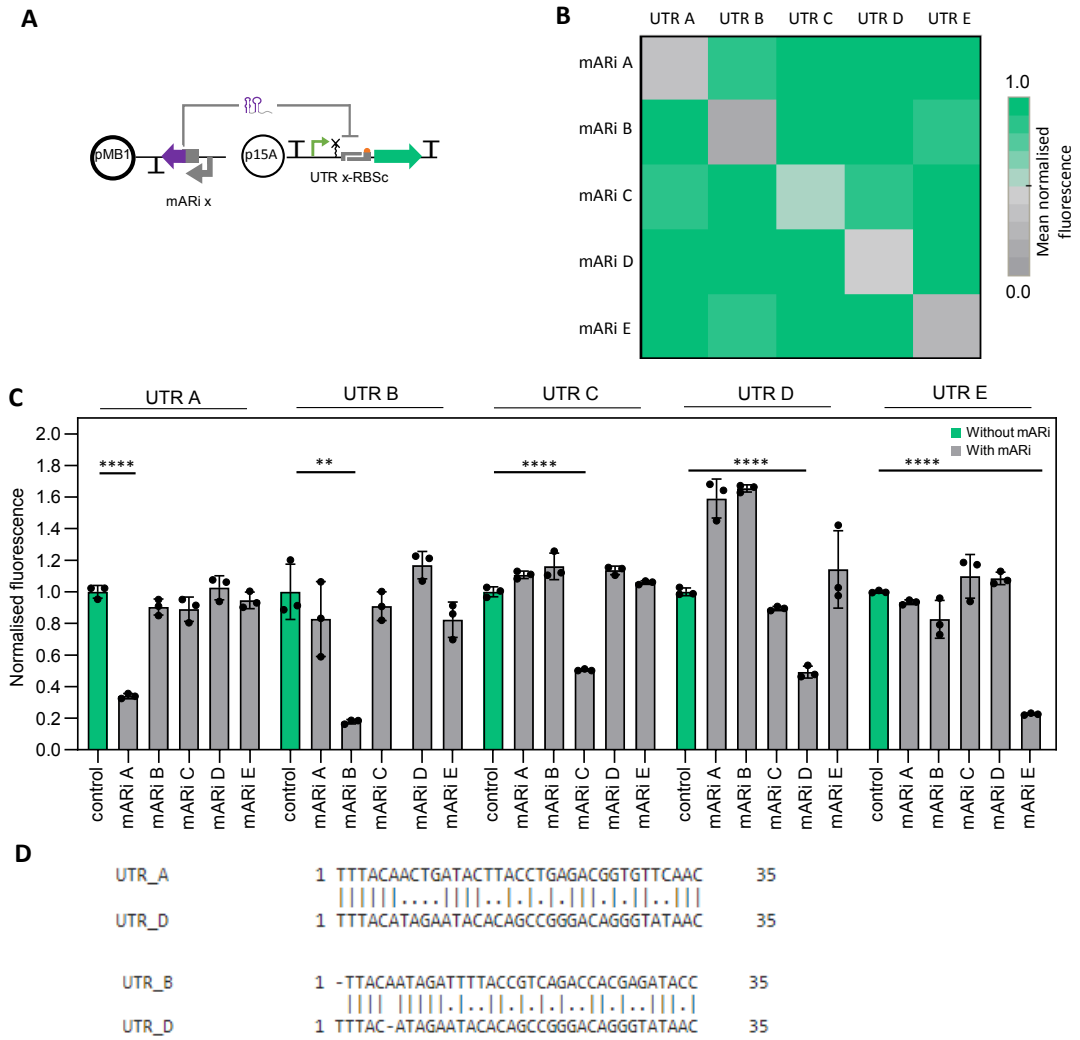


Figure 3.14 Orthogonality test of mARi/UTR set. (A) A schematic design of a high expression ratio of mARi built in a double plasmid system to evaluate the orthogonality of mARi/UTR set. (B) A heat-map showing the specificity of mARis in repressing their cognate UTR targets only. (C) Normalised fluorescence data associated with the error bars in Figure 3.14B. The sfGFP fluorescence was measured at exponential phase (6 h) by flow cytometry. Bars represent the average of triplicates data points (black dots) ± SD. (D) The pairwise alignment showing similarity between UTR A-UTR D (top panel) and UTR B-UTR D (bottom panel). Statistically significant differences determined using Student's t-test (**** represents p<0.0001, ** represents p<0.01).

3.13 mARi-based regulatory systems facilitate multiplexed and simultaneous gene regulation in a multi-gene system

To explore the potential of multiplexed regulation using our mARi set generated in Section 3.12, multiple transcriptional units expressing sfGFP and mCherry as reporters were next constructed. The expression of these reporters was driven by a medium strength promoter

$P_{J23101\_BASIC(1)}$ in a p15A backbone. As an mRNA expression cassette, GOIs for *sfgfp* and *mCherry* were preceded by UTR A-RBSc and UTR B-RBSc, respectively (Figure 3.15A). It is worth noting that the assembled dual transcriptional units produce two independent transcripts. To perform the multiplexed repression of *sfgfp* and *mCherry*, a strong expression of two mARis (mARi A and/or mARi B) was achieved by using P_{J23119} in a pMB1 backbone (Figure 3.15B). *E. coli* DH5α cells were then co-transformed with both reporter and mARi plasmids.

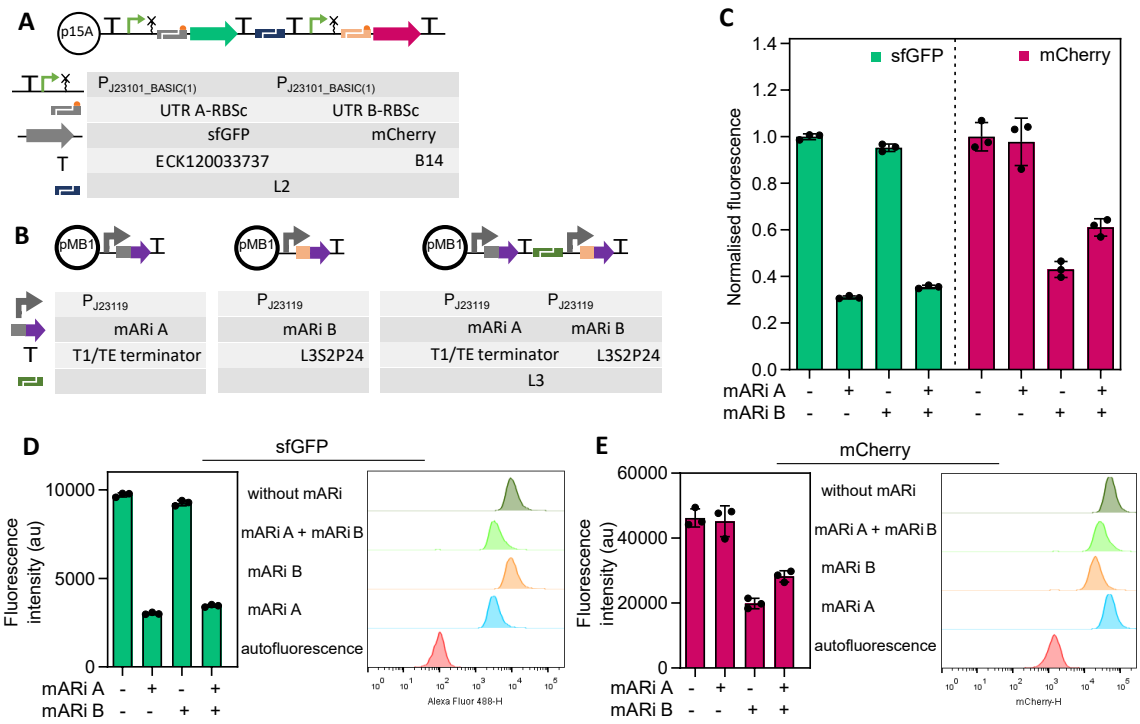


Figure 3.15 Multiplexed and simultaneous repression of dual transcription units using mARi-regulatory framework. (A) A schematic of dual transcription units of reporters resulted in two independent transcripts. The UTR A drives the expression of sfGFP and UTR B drives the expression of mCherry. The sfGFP and mCherry expressions were controlled by $P_{J23101\_BASIC(1)}$ in single plasmid system with a p15A backbone. (B) A schematic of mARi expression cassettes (mARi A and/or mARi B) were driven by P_{J23119} in a pMB1 backbone. (C) The performance of mARi-based repression showing independent, multiplexed, and simultaneous regulation of dual independent transcripts. The sfGFP and mCherry fluorescence measured at exponential phase (6 h) by flow cytometry. The fluorescence intensity (left panel) and representative of flow cytometry measurement (right panel) for (D) sfGFP expression and (E) mCherry expression correspond to Figure 3.15C. Bars represent the average of triplicates data points (black dots) ± SD.

Specific and independent repression by each mARi towards its cognate target was observed, with UTR A-*sfgfp* being selectively targeted by mARi A and UTR B-*mCherry* being targeted by mARi B (Figures 3.15C, 3.15D, and 3.15E). However, a slightly reduced repression of sfGFP and mCherry was found when mARis were simultaneously used to target multiple transcripts. This may due to the limited availability of the shared Hfq chaperone and the high expression of

the multiple mARi systems. The results from these revealed the capability of mARis to specifically and simultaneously controlling gene expression from independent transcriptional units.

3.14 mARi-based regulatory systems facilitate multiplexed and simultaneous gene regulation in operon system

Apart from creating a single expression cassette, the reusable UTR-RBS linkers can also be functionalised to create an operon system with alternative gene arrangements (Section 1.5). The multiplexed regulation was next tested by the use of the orthogonal mARis in two different operon systems. These operon systems produce two different reporters from a single mRNA. The first operon system was constructed with UTR A-RBSc driving the expression of sfGFP, whereas upstream, UTR B-RBSc drives expression of mCherry (Figure 3.16A). In this operon structure, a medium strength promoter P_{J23101_BASIC} was used to drive the expression of a single transcript containing *sfgfp* and *mCherry* reporters. Similar to the multi-transcriptional unit design, the operon cassette was then cloned into a p15A backbone. This single transcript was subjected to multiple and conditional mARi regulations. A strong expression of mARis (mARi A and/or mARi B) were expressed under the control of P_{J23119} in a pMB1 backbone (Figure 3.16B).

Unlike for multiple transcriptional units (Figure 3.15), the operon system produces in a single mRNA transcript containing two reporters and mARi target sites. With both genes operating on a single transcript, regulation of the upstream *sfgfp* follows a simple correlation to the cognate mARi A. A comparable repression was observed in the multi-transcriptional unit (Figures 3.15C and 3.15D) and operon systems (Figures 3.16C and 3.16D). In addition, the transcription and translation processes occur simultaneously in prokaryotics[200]. Figure 3.16C shows that the targeted repression for the first gene (*sfgfp*) in slightly affected by expression of the second gene (*mCherry*). Despite being affected by the sfGFP repression, the expression of the downstream mCherry is most strongly attenuated by its cognate mARi B (Figures 3.16C and 3.16E). Note that the silencing effects being additive in the mCherry expression, presumably due to double mARi expression.

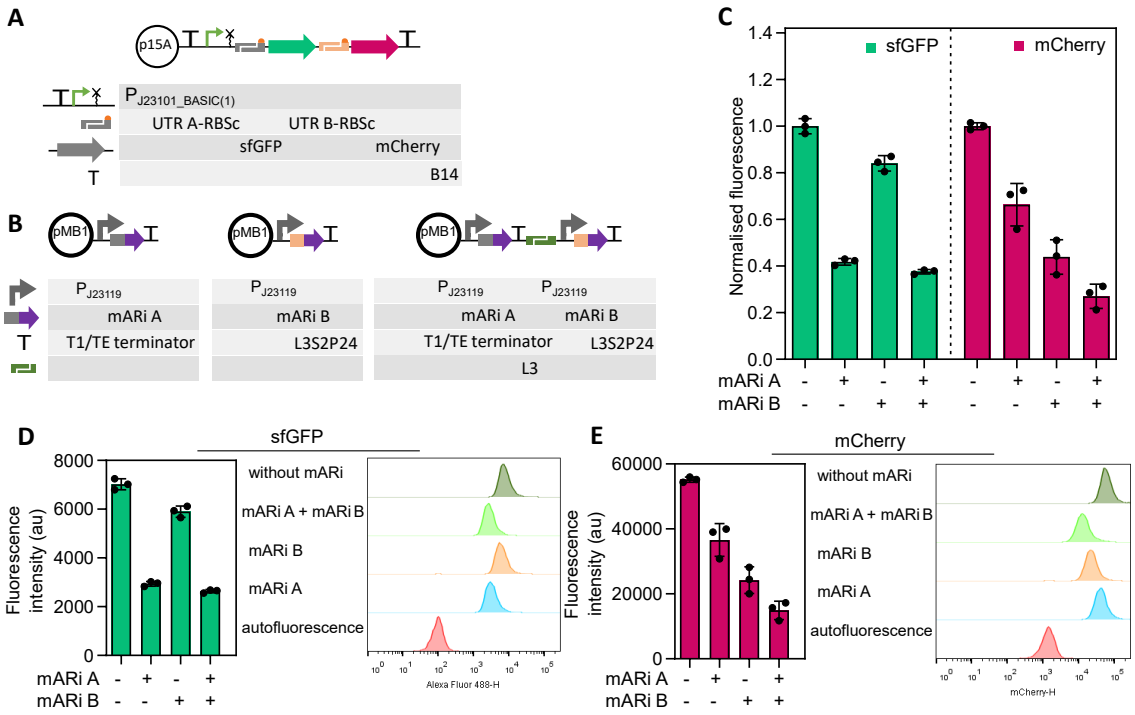


Figure 3.16 Multiplexed and simultaneous repression of an operon system using mARi-regulatory framework. (A) A schematic of dual reporters constructed in an operon system resulting in a single transcript. The UTR A drives the expression of sfGFP and UTR B drives the expression of mCherry. The sfGFP and mCherry expressions were controlled by P_{J23101_BASIC} in single plasmid system with a p15A backbone. (B) A schematic of mARi expression cassettes (mARi A and/or mARi B) were driven by P_{J23119} in a pMB1 backbone. (C) The performance of mARi-based repression showing multiplexed and simultaneous regulation of dual reporters in a single transcript. The sfGFP and mCherry fluorescence measured at exponential phase (6 h) by flow cytometry. The fluorescence intensity (left panel) and representative of flow cytometry measurement (right panel) for (D) sfGFP expression and (E) mCherry expression correspond to Figure 3.16.C. Bars represent the average of triplicates data points (black dots) ± SD.

To further investigate multiplexed regulation in the operon system, the orthogonal mARis were used to regulate an operon system with an inverted configuration from the first operon constructed (Figure 3.16A). The second operon was built by placing UTR B driving expression of sfGFP and UTR A expressing mCherry under the control of a constitutive promoter, $P_{J23101\_BASIC(1)}$ (Figure 3.17A). The same mARi set (mARi A and/or mARi B) was used to regulate the second operon (Figure 3.17B). While targeted repression was observed, the repression activity was also affected by the configuration of the UTRs and genes combination. The UTR B-sfGFP expression was slightly upregulated when the downstream UTR A-*mCherry* gene was repressed in the presence of mARi A. However, there is no clear explanation of this unexpected behaviour.

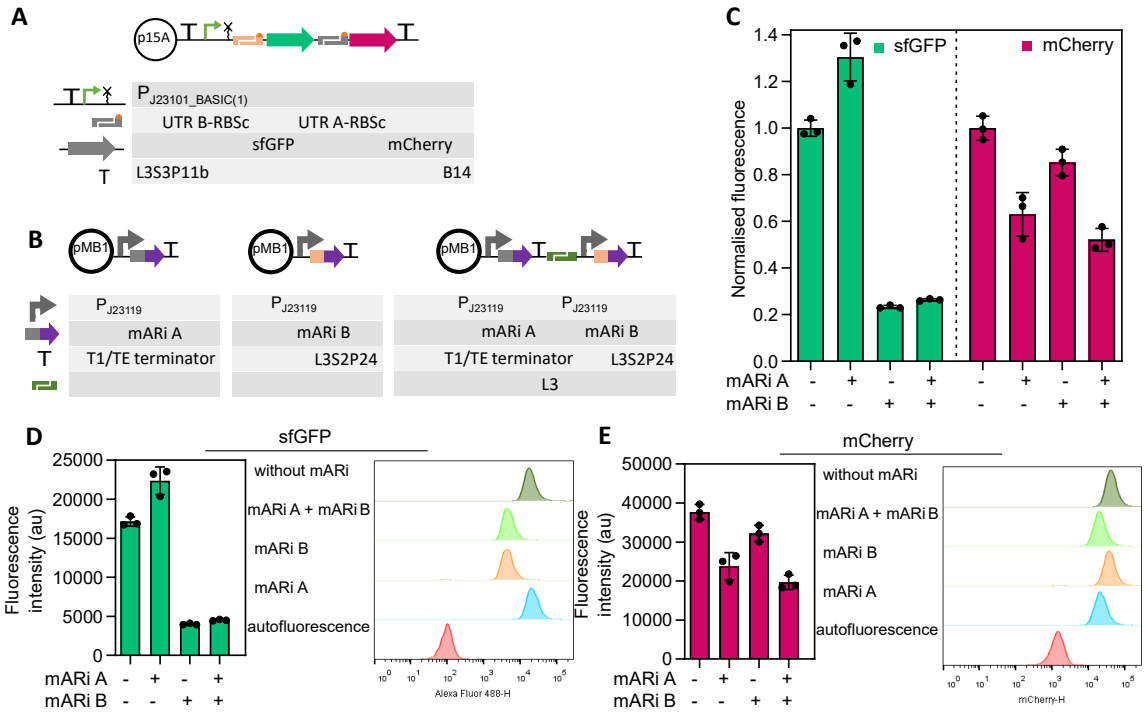


Figure 3.17 Multiplexed and simultaneous repression of an operon system using mARi-regulatory framework. (A) A schematic of dual reporters constructed in an operon system resulting in a single transcript. The UTR B drives the expression of sfGFP and UTR A drives the expression of mCherry. The sfGFP and mCherry expressions were controlled by $P_{J23101\_BASIC(1)}$ in single plasmid system with a p15A backbone. (B) A schematic of mARi expression cassettes (mARi A and/or mARi B) were driven by P_{J23119} in a pMB1 backbone. (C) The performance of mARi-based repression showing multiplexed and simultaneous regulation of dual reporters in a single transcript. The sfGFP and mCherry fluorescence measured at exponential phase (6 h) by flow cytometry. The fluorescence intensity (left panel) and representative of flow cytometry measurement (right panel) for (D) sfGFP expression and (E) mCherry expression correspond to Figure 3.17C. Bars represent the average of triplicates data points (black dots) ± SD.

3.15 Summary

In this chapter, an mARi-based regulatory system has been developed based on a natural Hfq-dependent sRNA scaffold, MicC. The mARi system controls gene expression at the post-transcriptional level by specifically targeting the standardised UTR sequences of the UTR-RBS linkers used during BASIC DNA assembly[62]. The target site, located upstream of the RBS showed effective repression while being independent of the RBS and GOI contexts. Therefore, mARi system can modularly applied to control gene expression when cognate UTR-RBS linkers are used in the construction of expression cassettes. The repression activity of mARis is further tuned by their relative expression ratio, molecular copy number, and spatial organisation of transcripts. In general, a greater repression activity was seen for high expression of mARi to mRNA, although overexpressed mARi may lead to saturating effects. As a modular and portable system, mARi-based regulation can be implemented in a single and

double plasmid-based configuration. This modular regulation system has also been shown to be portable and robust across host strains, including cloning strains (DH5α and DH10b) and expression strains (BL21(DE3) and BL21star(DE3)) and throughout different growth phases. The mARi design is expandable, so that mARi variants can be generated by modifying the seed sequences while retaining the MicC sRNA scaffold. Computational analysis of mARi system showed their specific interaction towards the only cognate target with minimal cross-interactions into non-cognate pairs and host backgrounds. The functional characterisation further validated the orthogonality of this set, and demonstrates the potential of mARi variants for multiplexed and simultaneous gene regulation. As an exemplar, two mARi systems (mARi A and/or mARi B) were employed for multi-gene regulation, covering multiple transcriptional units and an operon expression cassette. The results showed that multi-gene expression can independently and simultaneously be repressed by mARi systems. A key feature of this RNA-based regulation is that it can be readily integrated into the modular assembly framework, greatly easing the implementation f new biological systems.

Chapter 4

An adaptive and robust mARi-based negative feedback controller improves production of stress-inducing proteins

Portions of this chapter are reproduced from:

Dwijayanti, A., Storch, M., Stan, G.-B. & Baldwin, G. S. A modular RNA interference system for gene regulation and robust dynamic feedback control. *bioRxiv* 2019.12.12.873844 (2019). doi:10.1101/2019.12.12.873844

4.1 Introduction

An mARi-based gene regulator was developed and characterised for a range of tested genetic contexts in Chapter 3. This post-transcriptional regulator was also shown to be orthogonal and independent gene regulation in multi-gene systems with different configurations. Moreover, its integration into BASIC assembly framework[62] provides easy implementation in synthetic circuits for various potential applications. For example, mARi can be utilised as an effector implementing a negative feedback system. This implementation is further supported by the nature of RNA-RNA interaction behind the silencing mechanism of mARi, so that dynamic and rapid signal propagation can be achieved. Owing to this attribute, a feedback controller implemented using mARi can dynamically tune the level of an mRNA of interest under specific conditions. In this chapter, this strategy is pursued as an alternative solution to effectively alleviate translational resources limitation during overexpression of stress-inducing proteins.

It is generally accepted that overexpression of recombinant proteins, both heterologous and endogenous, require gene expression machinery which is also naturally used by host cells to maintain their growth[82,153,210,211]. As such, the production of recombinant proteins can significantly impact both transcriptional and translational resources, compete with host maintenance, and limit production yield (Figure 4.1A). To this extent, the availability of translational resources, rather than transcriptional resources, has been identified as a major limiting factor in *E. coli*[212–214]. This hypothesis has been highlighted through a number of mathematical modelling studies[2,215,216]. In addition to this computational work, the limitation in ribosomal proteins has also been shown through *in vivo* and *in vitro* experiments, with the use of strong RBSs known to significantly reduce cellular capacity and impact growth[82,211,217].

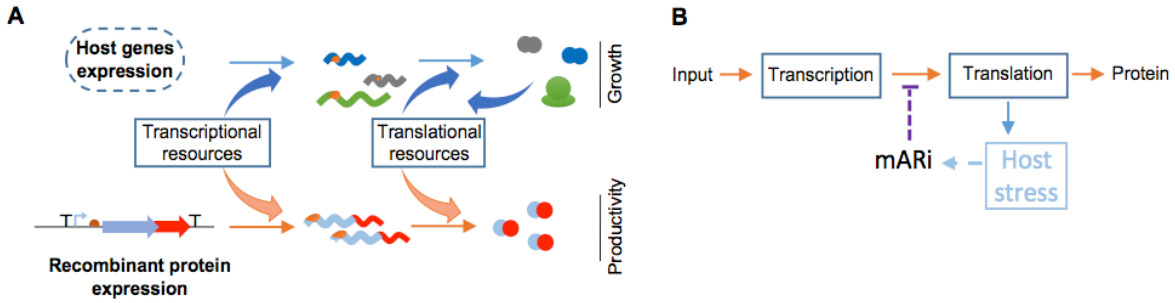


Figure 4.1 The general concept of shared resource allocation and a strategy to modulate it using negative feedback control. (A) The schematic of shared cellular resources (transcriptional and translational resources) utilised between cell fitness maintenance and maximal production of recombinant protein. (B) Block diagram of mARi-based dynamic negative feedback in modulating translational resource allocation. The induction (input) of the recombinant gene leads to transcription and translation process to produce a protein as output. The limitation of translational resources triggers host-stress response then used as input for the mARi production. Once mARi is transcribed, the mARi-based dynamic negative feedback modulates the level of mRNA recombinant at post-transcriptional level upon the translational stress.

To overcome this bottleneck in protein expression, a traditional approach is to optimise the design of a genetic circuits by substituting bioparts to tune gene expression[37,157,218]. Using a modular assembly method, variants of a genetic circuit can be produced, with combinatorial designs easily built from interchangeable bioparts[37]. In a plasmid-borne system, an efficient design could be obtained by combining different types of plasmid copy number, promoter strength, RBS strength, and employing different codon optimisation schemes[78,211]. Additionally, mathematical modelling has been used to guide the search for optimal genetic designs[211,219,220]. More specifically, modification of RBS strength has been used to rationally design genetic circuits with reduced experimental effort[104,220]. The RBS strength can be predicted with established computational methods[31–33]. However, the results in Table 3.2 showed the discrepancy between the predicted RBS strength of a set of RBS parts and their actual experimental characterisation, casting doubt on the ability to precisely tune gene expression levels without experiments.

In addition to static control, a forward engineering strategy using dynamic control offers a more practical approach to regulate gene expression. A number of negative feedback systems have been widely adopted in engineering biology to create dynamic and robust control of genetic circuits able to accommodate different perturbations[161,192,221,222]. As such, dynamic control relies on the ability to sense the changing conditions or specific circumstances and respond to it quickly[78]. In particular, negative feedback has been recently introduced to robustly adjust gene expression in response to expression burden, a condition where the

expression of heterologous proteins adds extra burden on a cell and significantly diminishes cell growth[82]. In that study, the authors identified native promoters related to a heat-shock response and used CRISPRi/gRNA to create a feedback control loop. This negative feedback control helped regulate cellular burden, helping to maintain cell fitness and improved the total protein production.

Instead of targeting the transcriptional level[82], in this chapter an alternative approach to alleviate the cellular burden is proposed by regulating gene expression at the post-transcriptional level (Figure 4.1B). This strategy is implemented through the previously designed and characterised mARi-based regulation (see Chapter 3) combined with a well-defined generic stress-responsive promoter, P_{htpG1}[82].

Firstly, the response of P_{htpG1} activation was characterised in the presence of overexpression of a stress-causing protein (i.e. VioB-mCherry) in the different genetic designs. Then, the P_{htpG1} coupled to a mARi-based gene controller was tested to dynamically modulate translational resource allocation and improve the yield production of VioB-mCherry. The robustness of the feedback system was examined in different host strains and external perturbations such as different temperatures, media, and production volumes. Finally, the modularity and portability of the P_{htpG1}-mARi were further explored to improve the production of a different protein of interest, LacZ-mCherry.

4.2 System design

To facilitate the modularity and easy implementation of the negative feedback system, we used the BASIC DNA assembly method[62]. The use of this approach allows us to easily swap components by the use of interchangeable bioparts (Section 1.6) that have been characterised in the previous chapters (see Chapters 2 and 3). Using this technique, a circuit with negative feedback system was built to demonstrate the concept of stress-driven negative feedback control mediated by mARi.

The negative feedback system was designed as two expression cassettes. The first expression cassette is an inducible cassette to express a recombinant protein (Figure 4.2A). A heterologous GOI such as *vioB* with 3 kb length was chosen as one of the tested recombinant proteins. The *vioB* gene taken from *Chromobacterium violaceum* encodes an enzyme, which is responsible for catalysing the dimerisation of indole-3-pyruvic acid imine (IPA) in violacein biosynthesis[223]. Since VioB does not directly interact with intermediate metabolites in the native *E. coli* metabolism, the production of this enzyme, is perceived to infer a cost on gene expression rather than metabolic burden[2]. The use of the protein-fusion VioB-mCherry has been shown to reduce *in vivo* cellular capacity in previous studies[82,211]. In this case, the codon optimised *vioB* was fused to the coding sequence of *mCherry* at its C terminus[211]. This fusion protein enables an easy and real-time monitoring system of protein production. This *vioB-mCherry* fusion results in 3.7 kb of long gene. Additionally, the length of the transcript has been known to play a central role in ribosome recruitment and allocation, affecting the efficiency of mRNA translation[224]. Theoretical studies have also suggested that long transcripts are highly affected by the changes in ribosome availability. Further, experimental investigations in a cell-free system and *in vivo* have confirmed that the different lengths of transcript reduce cellular capacity[217].

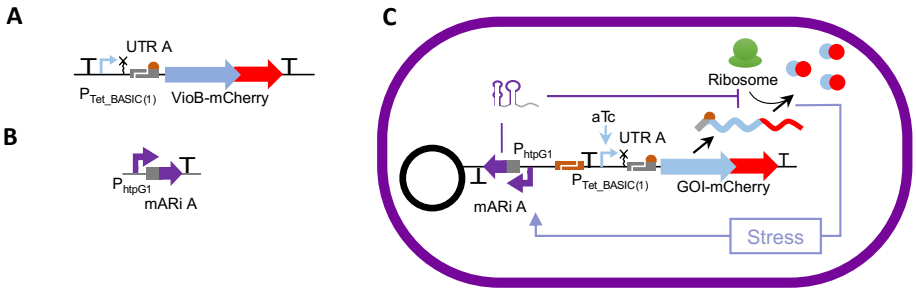


Figure 4.2 A schematic of a closed-loop circuits negative feedback system. (A) The recombinant protein was composed of Gene of Interest (GOI) i.e. VioB (3 kb) fused to mCherry. The expression of recombinant protein was driven by a standardised Tet inducible promoter, $P_{Tet\_BASIC(1)}$ and UTR A. (B) An adaptive feedback controller cassette built of a combination of a stress-responsive promoter, $P_{htpG1,}$ and mARi-based regulator. (C) A closed loop of negative feedback system constructed of inducible expression cassette and adaptive feedback controller in a single plasmid system.

In this inducible expression cassette (Figure 4.2A), the transcription of *vioB-mCherry* is driven by $P_{Tet\_BASIC(1)}$, an inducible promoter based on Tet-based regulation (see Sections 2.7, 2.8 and 2.9). This promoter is responsive to a tetracycline analogue, aTc, that prevents binding of Tet repressor protein (TetR) to its operator site. In particular, the TetR driven by P_{J23105} was used to regulate $P_{Tet\_BASIC(1)}$. The characterisation results for $P_{Tet\_BASIC(1)}$ show a high fold induction

in the presence of aTc (Sections 2.7, 2.8 and 2.9). To create a functional expression cassette, $P_{Tet\_BASIC(1)}$ and *vioB-mCherry* were connected using UTR A, one of the UTR-RBS linkers used in a BASIC DNA assembly[62]. Furthermore, this standardised UTR sequence has been shown to provide a modular target for mARi-based gene regulation (see Chapter 3).

The second transcriptional unit is an adaptive feedback controller that acts to specifically sense host-stress and respond to it accordingly. The biomolecular feedback controller is encoded as a genetic sensor and actuator (Figure 4.2B). The sensor functions to immediately detect the host-stress response imparted by the overexpression of the inducible construct. A stress-responsive promoter, P_{htpG1} was chosen as a stress detector on the basis of recently published work[82] and its characterisation result discussed in the following sections (Section 4.3 and 4.4). For creating a functional controller, the P_{htpG1} was then combined with an mARi-based regulator. Similar to the previously developed design of the mARi-based controller (Section 3.2), P_{htpG1} was placed directly upstream of the seed sequences without any additional bases. Since UTR A was used to assemble the inducible expression cassette, the cognate mARi A that targets UTR A was then employed to regulate the expression of VioB-mCherry. The post-transcriptional properties of this actuator have been characterised in diverse genetic contexts (see Chapter 3). Ultimately, mARi A has been shown to display a relatively constant activity for a range of RBSs (Section 3.6), is compatible with a set of plasmid backbones (Section 3.8), and is portable across different host strains and growth phases (Section 3.11).

A closed loop negative feedback system was then created in the presence of an inducible expression cassette and an adaptive feedback controller (Figure 4.2C). Both transcriptional units were then assembled in a single plasmid system and placed in the opposite direction to avoid transcriptional read-through. Moreover, this genetic configuration would increase the proximity of target and regulator, so that an increase of mARi efficacy could be obtained as previously observed in Sections 3.8 and 3.9.

4.3 Overexpression of VioB-mCherry results in activation of host-stress response

One of the key strategies in the establishment of the designed adaptive feedback is its capability to rapidly detect changes of native cellular response, i.e. host-stress response when inducible constructs are strongly expressed. Previously, the overexpression of VioB-mCherry has been shown to trigger a host-stress response in *E. coli* DH10b and MG1655 cells[82]. Upon overexpression of VioB-mCherry, native promoters related to heat-shock response σ factor (σ^{32}) were up-regulated. Among the native promoters responsive to host-stress responses, the *htpG1* promoter, P_{htpG1} was shown to display the best fold induction[82]. The P_{htpG1} promoter is also responsible for controlling expression of the native heat shock protein in *E. coli*, namely HtpG, that is commonly present under environmental stress conditions[83].

The P_{htpG1} has been shown to strongly and quickly respond to VioB-mCherry-induced expression[82]. Further to this finding, the P_{htpG1} was used as a host-stress sensor for detecting and quantifying cellular stress levels *in vivo* and in real-time[82]. In particular, the P_{htpG1} is employed to drive an sfGFP expression in a ColE1 plasmid backbone. This plasmid is referred to as P_{htpG1}-sfGFP plasmid (Figure 4.3A). Here, the P_{htpG1}-sfGFP plasmid[82] was then used to investigate the different stress levels observed in response to VioB-mCherry overexpression.

Three VioB-mCherry expression plasmids, were constructed to represent a variety of VioB-mCherry expression levels and host-stress responses by differing RBS strength and backbone copy numbers (Figure 4.3B). Two different RBSs were selected from the previous functional characterisation results in Table 3.2. Specifically, RBSa (mean relative fluorescence = 1) and RBSc (mean relative fluorescence = 0.35) were used to represent a strong and medium RBS strength, respectively. In the first expression plasmid, the UTR A-RBSa, with a strong RBS was used to drive VioB-mCherry expression and cloned in a medium copy origin of replication (p15A, ~10 copy per cell). The second expression plasmid was constructed with a similar expression cassette in a high plasmid copy number (pSB1C3, 100-300 copy per cell). Whereas, the third expression plasmid used the UTR A-RBSc carrying a medium RBS in a pSB1C3 backbone. With mCherry reporter fused to VioB, the protein production can easily be

approximated as a red fluorescent output. Both P_{htpG1}-sfGFP plasmid[82] and VioB-mCherry expression plasmid were propagated into BL21(DE3) strains (Figure 4.3C).

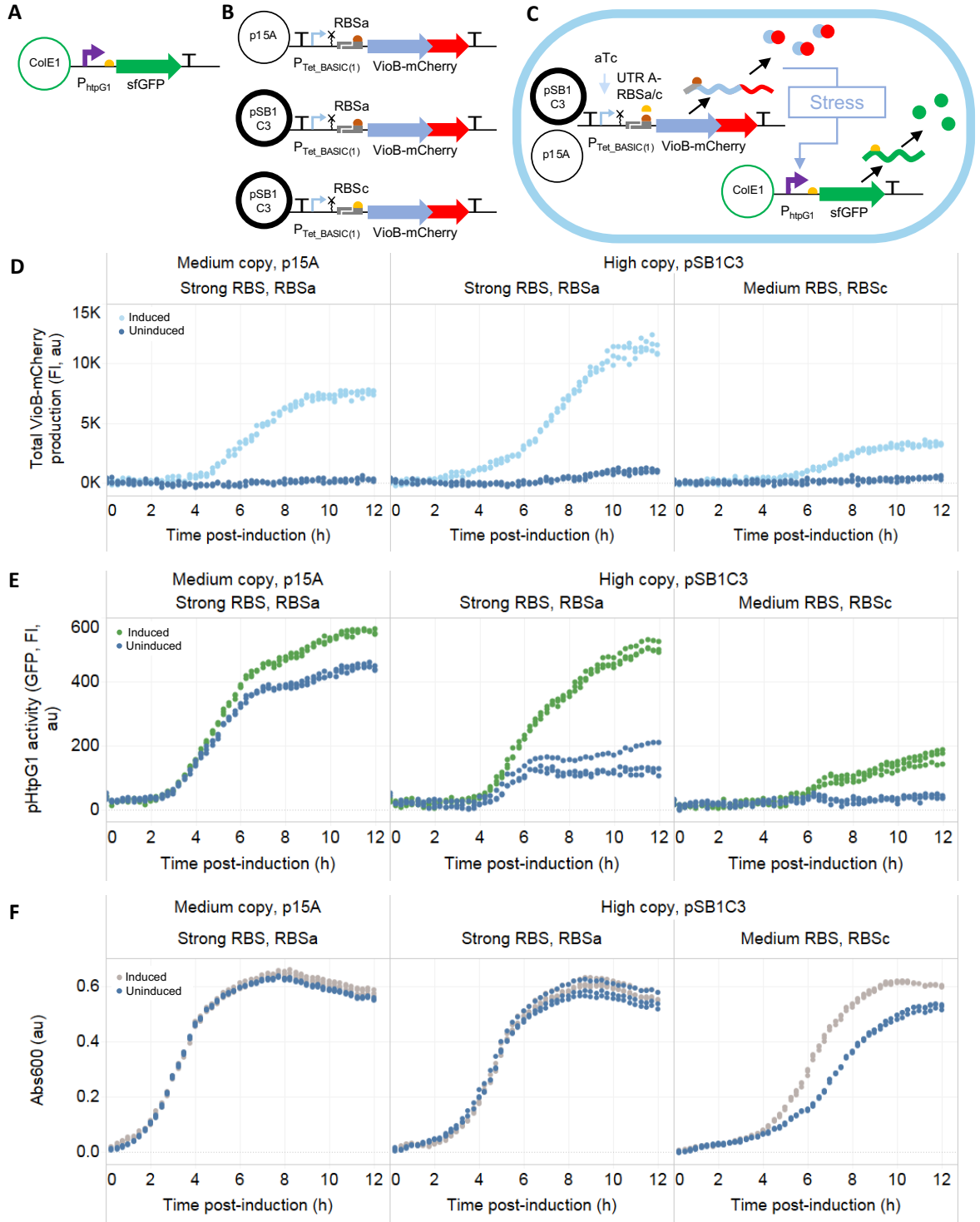


Figure 4.3 Characterisation of a stress-sensitive promoter, P_{htpG1} over the time of VioB-mCherry induction. (A) A schematic of P_{htpG1}-sfGFP plasmid in a ColE1 backbone[82] used as a reporter of the host-stress level. (B) A schematic of variant VioB-mCherry expression plasmids in a high plasmid copy number (pSB1C3) and a medium plasmid copy number (p15A) backbone. A strong RBS (RBSa) and a medium RBS (RBSc) were used in the VioB-mCherry expression cassette. (C) The system was constructed in a double plasmid system and propagated into BL21(DE3) strains. (D) The total VioB-mCherry production, (E) P_{htpG1} activity readout, and (F) growth profile over the time of induction. aTc inducer at 0 (uninduced) and 5 (induced) ng/ml was used to induce the expression of VioB-mCherry. Dots represent independent data point of three independent replicates.

A variety of VioB-mCherry outputs from different genetic designs over 12 h post-induction are shown in Figure 4.3D. For the same promoter and a comparable inducer concentration, a higher output level of VioB-mCherry expression was observed when using a stronger RBS in both plasmids. Interestingly, production strains using a medium RBS did not result in the same level output as a strong RBS in a medium copy plasmid, even though it is in a high copy number plasmid. This result conflicted with previous results, where the use of a strong RBS in a medium plasmid backbone resulted in an equivalent output as an expression cassette with a medium RBS in a high copy number plasmid[211].

As predicted, VioB-mCherry overexpression triggered P_{htpG1} activation for all genetic designs (Figure 5.3E). The strong RBS used to express VioB-mCherry caused higher activation of P_{htpG1} than a medium RBS. Surprisingly, a high baseline of the P_{htpG1} activation was observed in a p15A backbone despite low expression of VioB-mCherry without aTc induction. One possible reason is the backbone for VioB-mCherry expression in this design (p15A) has a similar initiation and regulation of replication system to the host-stress reporter plasmid (ColE1)[225], hence may affect the output of P_{htpG1}-sfGFP plasmid.

It was previously reported that overexpression of VioB-mCherry significantly induced expression burden and diminished cell growth[82,211,217]. By contrast, the overexpression of VioB-mCherry constructed with a strong RBS in a high copy plasmid did not impact the growth of the engineered strains (Figure 4.3F). Another growth anomaly was also observed in the expression of VioB-mCherry using a medium RBS in a high plasmid copy number. Compared to without induction, a slightly increased growth was found in the construct with aTc induction. There is no clear explanation for these phenomena and further investigation is required to elucidate this interesting effect.

4.4 Evaluation of the *htpG1* promoter (P_{htpG1}) upon different induction of a stressful construct

The mARi-based negative feedback was intended to alleviate translational stress by dynamically modulating the level of mRNA of interest at the post-transcriptional level. To verify that the activation of P_{htpG1} is mainly caused by translational stress, i.e. in the use of strong RBS, the P_{htpG1} response was evaluated against a variety of VioB-mCherry transcript levels. The effect of different transcript levels was examined by varying the concentration of aTc inducer added to the growing culture of selected genetic designs. Due to the low basal expression of P_{htpG1} (see Section 4.3), the VioB-mCherry expression controlled by a strong and medium RBSs in a pSB1C3 were employed in this section (Figure 4.4A). The expression of VioB-mCherry, P_{htpG1} activity, and growth profile were then examined under different induction levels (0-100 ng/ml of aTc inducer) at 12 h post-induction.

Consistent with previous results, the strong RBS had a higher VioB-mCherry production compared to the medium RBS for the equivalent concentration of aTc (Figure 4.4B). For the stronger RBS, maximum expression was achieved at a lower aTc concentration. Furthermore, the medium RBS had a maximum expression output at higher aTc concentrations. These shifted dose-response curves were in agreement with recently developed theoretical work[226].

In agreement with the construct-induced expression, the P_{htpG1} activity was positively correlated to the VioB-mCherry output at all induction levels (Figure 4.4C). In spite of this, the RBS strength dictated the magnitude of P_{htpG1} response. This further confirmed that RBS strength, rather than the amount of mRNA and plasmid copy number, substantially contributed to activating a host-stress response.

The maximum aTc induction for the construct with a strong RBS and high copy plasmid, did not impair the growth of engineered strains (Figure 4.4D). Notably, P_{htpG1} was strongly activated as seen in Figures 4.3E and 4.4C. This paradox suggested that the existing translational stress, somehow, limited protein production and induced the host-stress response even though a growth defect was not observed.

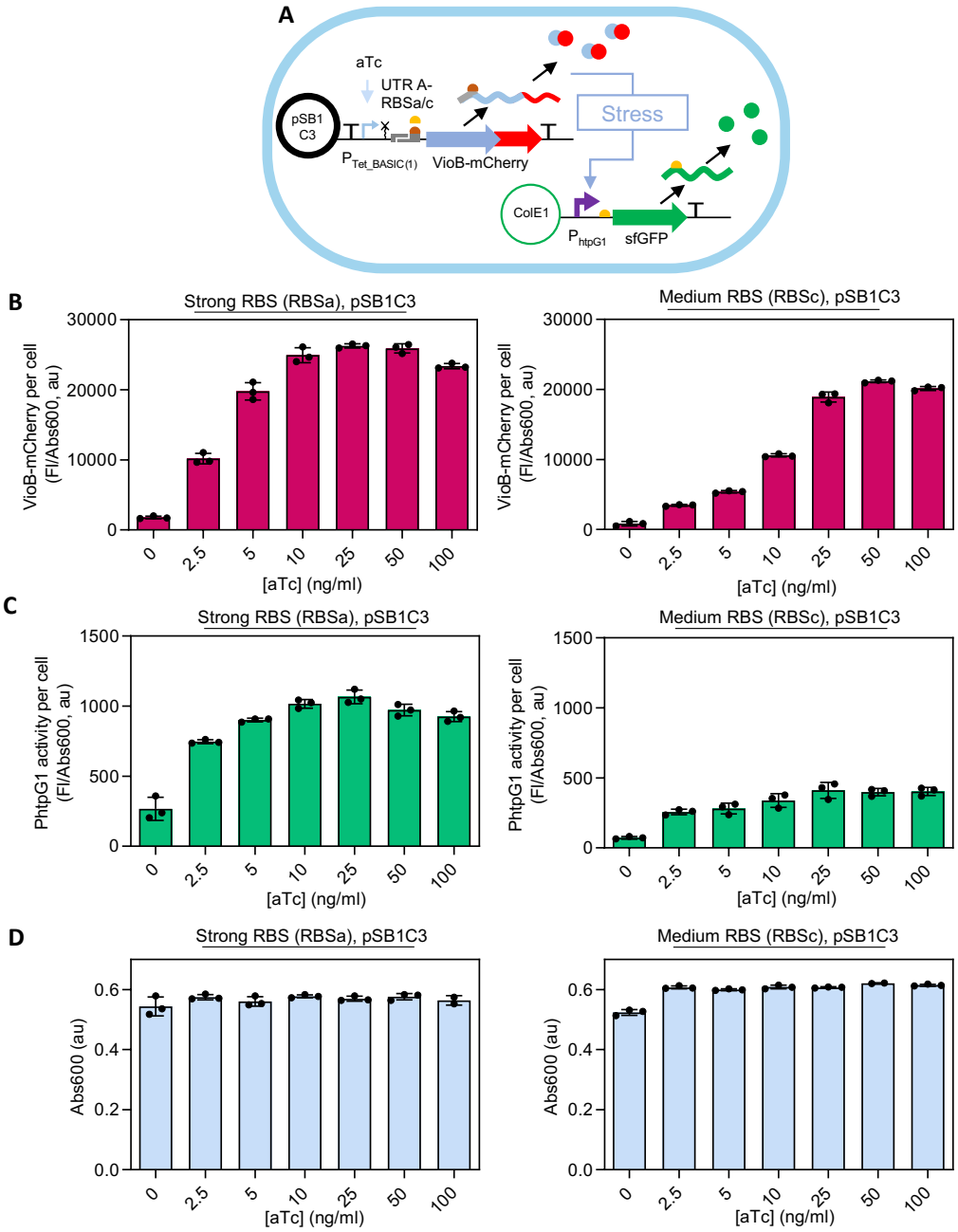


Figure 4.4 Characterisation of a stress-sensitive promoter, P_{htpG1} with varied inducer (aTc) concentration. (A) A schematic of a system used to evaluate the native cellular stress-response towards VioB-mCherry overexpression. The system was constructed in a double plasmid system where the VioB-mCherry expression cassette harbouring a strong RBS (RBSa) and a medium RBS (RBSc) in a high plasmid copy number (pSB1C3). As a reporter of host-stress level, the previously built plasmid contains a stress-responsive promoter, P_{htpG1} drives the expression of sfGFP in a ColE1 backbone[82] was used. (B) The VioB-mCherry per cell, (C) P_{htpG1} activity per cell, and (D) Cell density at 12 h post-induction. A strong RBS, RBSa (left panel) and a medium RBS, RBSc (right panel) were used in the VioB-mCherry expression cassette. aTc inducer at different concentration (0-100 ng/ml) was used to induce the expression of VioB-mCherry. Black dots represent three data points and bars represent the average of triplicates data points ± SD.

The magnitude of host-stress level, rather than growth retardation, is a more important factor to ensure the strength of the feedback control. As P_{htpG1} was used to transcribe mARi, the output of this promoter determined the amount of mARi produced. To effectively repress the mRNA of interest, a high concentration of mARi is required to tune the feedback strength (see

Section 3.7). In this regard, a high feedback strength is also crucial for the robustness of the negative feedback system[221].

4.5 The use of mARi-based negative feedback improves VioB-mCherry production and extends exponential growth

After inspecting the activation response of P_{htpG1} upon VioB-mCherry induction, this stress-responsive promoter was used to drive the transcription of a mARi regulator as described earlier in Section 4.2. To this end, the VioB-mCherry expression driven by a strong RBS (RBSa) in a pSB1C3 backbone was chosen to test the feedback system.

This adaptive biomolecular feedback controller was then deployed to create a closed feedback loop in the designed circuits expressing VioB-mCherry (Figure 4.5A). Since the mRNA of interest contains UTR A, the cognate mARi (i.e. mARi A) was used to dynamically regulate the intracellular concentration of *vioB-mCherry* transcripts. Due to nature of the interaction of the RNA species, P_{htpG1}-mARi was expected to provide rapid signal propagation as well as reversible and dynamic repression of the *vioB-mCherry* transcript. This leads to greater efficiency in modulating and balancing translational resource allocation. In turn, maximum production and robust growth could simultaneously be achieved over the time of induction.

As a control, an inducible expression cassette without P_{htpG1}-mARi was used. These regulated and unregulated circuits were first used in *E. coli* BL21(DE3) cells. The total production of VioB-mCherry could be estimated from the red fluorescent signal and the cell growth was observed by Abs600 via a continuous measurement (Sections 6.3.1 and 6.3.5).

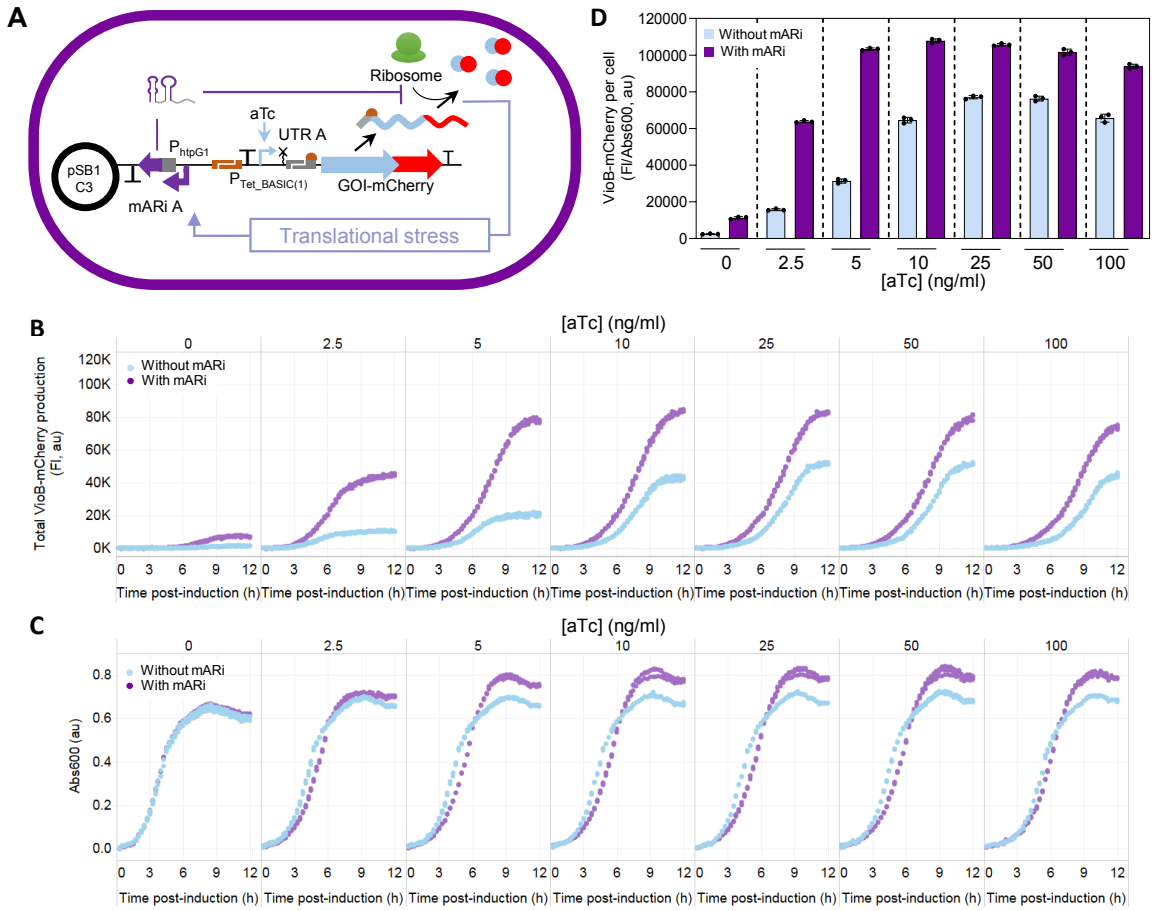


Figure 4.5 The mARi-based negative feedback improved VioB-mCherry production in *E. coli* BL21(DE3) strains. (A) A schematic of a dynamic negative feedback system implemented in a VioB-mCherry production. (B) The accumulation of VioB-mCherry and (C) growth profile over the time of induction with aTc concentration 0-100 ng/ml. The engineered strains were grown in LB medium at 37°C and 600 rpm in a 96 -well plate. To examine the performance of the feedback controller under different transcript numbers, aTc inducer with a concentration of 0-100 ng/ml was added after 1 h outgrowth. (D) The VioB-mCherry production per cell in response to varied aTc induction. Data were taken at 12 h post-induction. Bars represent the average of triplicates data points (denotes as a single dot) ± SD.

The inclusion of mARi expression cassette resulted in improvement production of total VioB-mCherry at all induction levels (Figure 4.5B). Remarkably, the feedback controller enhanced the total VioB-mCherry production up to 4-fold with 5 ng/ml aTc induction. The production improvement was observed from the early logarithmic phase. The use of stress-driven feedback controller was speculated to simultaneously modulate the host-stress level and improve the recombinant protein production despite the fluctuation in shared transcriptional resources. Importantly, the native production of Hfq chaperone required in the mARi-based negative feedback is also controlled via heat-shock stress promoters such as σ^{32}-dependent and σ^{70}-dependent promoters[128] (Section 1.6.3). This means that both the mARi actuator and Hfq chaperone in this biomolecular feedback control system are activated by host-stress.

Hence, the design provides effective and dynamic negative feedback control for shared cellular resource allocation between native growth and heterologous protein production.

This improvement of production was further confirmed through Sodium Dodecyl Sulphate-Poly Acrylamide Gel Electrophoresis (SDS-PAGE) analysis of whole-cell extracts (Section 6.2.12). The improvement of VioB-mCherry expression was clearly shown by a protein band with the size of 138.6 kDa (shown with an arrow at Figure 4.6) that can be compared between the regulated and unregulated circuits in the presence and absence of 5 ng/ml aTc inducer.

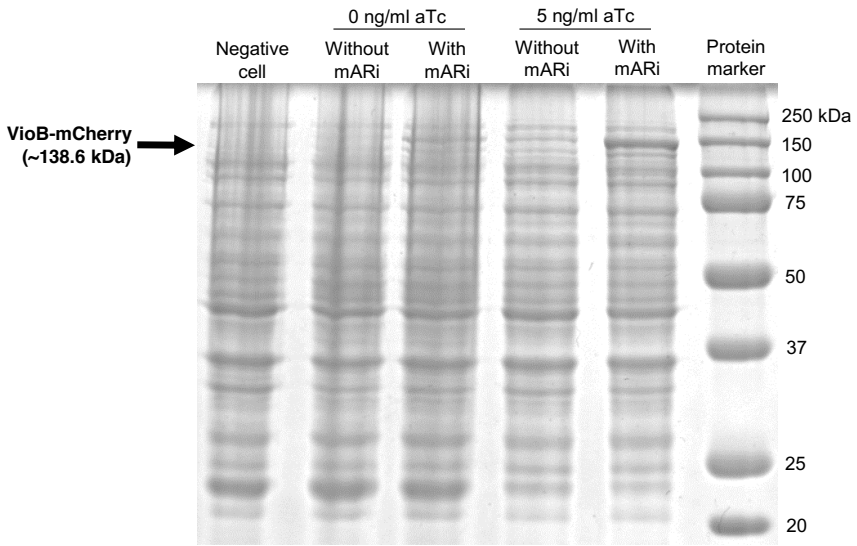


Figure 4.6 SDS-PAGE confirmation of VioB-mCherry production. The protein production of VioB-mCherry in the presence and absence of mARi are shown. The VioB-mCherry with the theoretical size of 138.6 kDa is indicated with an arrow. This theoretical molecular weight of VioB-mCherry was computed using Protein Molecular Weight Calculator (https://www.sciencegateway.org/tools/proteinmw.htm). The BL21(DE3) strains were grown in 5 ml of LB medium (30 ml tube) at 37°C and 220 rpm. 5 ng/ml of aTc was used to induce protein expression. The protein samples were taken from expressing strains at 12 h post-induction.

There was a significant difference in the observed growth profiles between strains carrying regulated and unregulated circuits (Figure 4.5C). Curiously, the *E. coli* BL21(DE3) strains with the P_{htpG1}-mARi controller showed slightly slower growth in the early logarithmic phase when compared to the strains lacking the controller. The use of mARi-based negative feedback control, however, extended the exponential growth phase leading to a higher cell density of the engineered strains when grown in LB medium at 37°C. This observation may correlate to the dynamic nature of the mARi-based feedback loop, including the molecular mechanism of mARi in repressing its targets and the varying concentration of the native helper chaperones. The detail of the mechanisms behind the mARi silencing are not fully understood yet and

require further investigation. It also should be noted that the molecular abundance of two key proteins (HtpG and Hfq) in the feedback system is affected by bacterial growth phase[83,130].

Notably, the engineered strains with the P_{htpG1}-mARi controller resulted in higher total production of VioB-mCherry and prolonged exponential growth phase. As such, this may arise from a growing population of non-producing mutant cells. To test this, single-cell analysis of the BL21(DE3) strains with regulated, unregulated, and negative (empty plasmid) were performed (Section 6.3.6). The homogeneity of producing strains across different growth phases was shown in Figure 4.7. This result further confirmed that the extended growth phase was not caused by non-producing escape mutants.

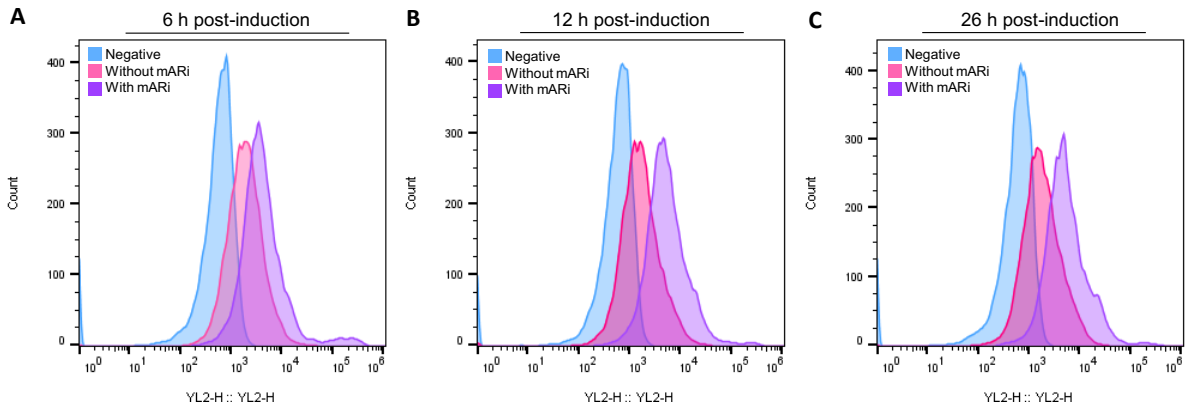


Figure 4.7 Uniformity of the representative VioB-mCherry production strains at (A) 6 h, (B) 12 h, and (C) 26 h post-induction. The BL21(DE3) strains were grown in LB medium at 37°C and 600 rpm. 5 ng/ml of aTc was used to induce protein expression.

Following this, the yield of VioB-mCherry was calculated by dividing the measured total VioB-mCherry to Abs600. The mARi-based negative feedback improved the yield production of VioB-mCherry across different induction levels (Figure 4.5D). At 12 h post-induction, there was a notable shift of maximal yield production in the presence of the feedback controller. Interestingly, the engineered strain with the P_{htpG1}-mARi controller showed a high (4-fold) improvement of VioB-mCherry production per cell at lower mRNA transcript concentration with aTc induction (5 ng/ml). At higher induction levels, the fold improvement was reduced. This effect may be due to the nature of post-transcriptional regulation of mARi, especially the transcript expression ratio of mARi and its target, as previously discussed in Sections 3.7 and 3.8.

Beyond that, the biomolecular feedback system tested in this section enabled efficient translational resource allocation by optimising yield production, even with lower mRNA transcript concentrations. This feature of mARi-based negative feedback is probably interesting and useful for biotechnological application, especially industrially-relevant enzymes or therapeutic recombinant protein production at industrial scale.

4.6 mARi-based negative feedback is portable for different tested production strains

The improvement of protein production and extended exponential growth shown in the *E. coli* BL21(DE3) strains further prompted the implementation of the designed feedback system in a different host strain, such as *E. coli* DH10b cells. *E. coli* DH10b cells are commonly used for cloning, hence one can directly use the positive clones to examine the protein improvement without the need to isolate and propagate the assembled plasmid into a production strain such as BL21(DE3). The intention here was to greatly simplify the routine workflow of molecular cloning for enzyme or protein overexpression purposes.

It should be noted that different strains have alternative genotype. As aforementioned in Sections 4.2 and 4.5, two core parameters required for the mARi-based feedback controller are a high fold activation of P_{htpG1} (see Sections 4.3 and 4.4) and the ability for mARi to effectively repress its cognate target at the post-transcriptional level (see Chapter 3). In this context, the different molecular abundances of HtpG and Hfq chaperones in both strains may affect the performance and robustness of the previously designed feedback system. Moreover, the abundance of those chaperones has been known to be affected by growth phases and genotype background of host strains[83,130,205]. The P_{htpG1} activity has been previously assessed in *E. coli* DH10b cells and been shown to display a rapid and strong response in the presence of stress-inducing VioB-mCherry overexpression[82]. Whereas, the high constitutively expressed mARi has resulted in effective repression activity in *E. coli* DH10b and relatively constant across different growth phases (see Section 3.11).

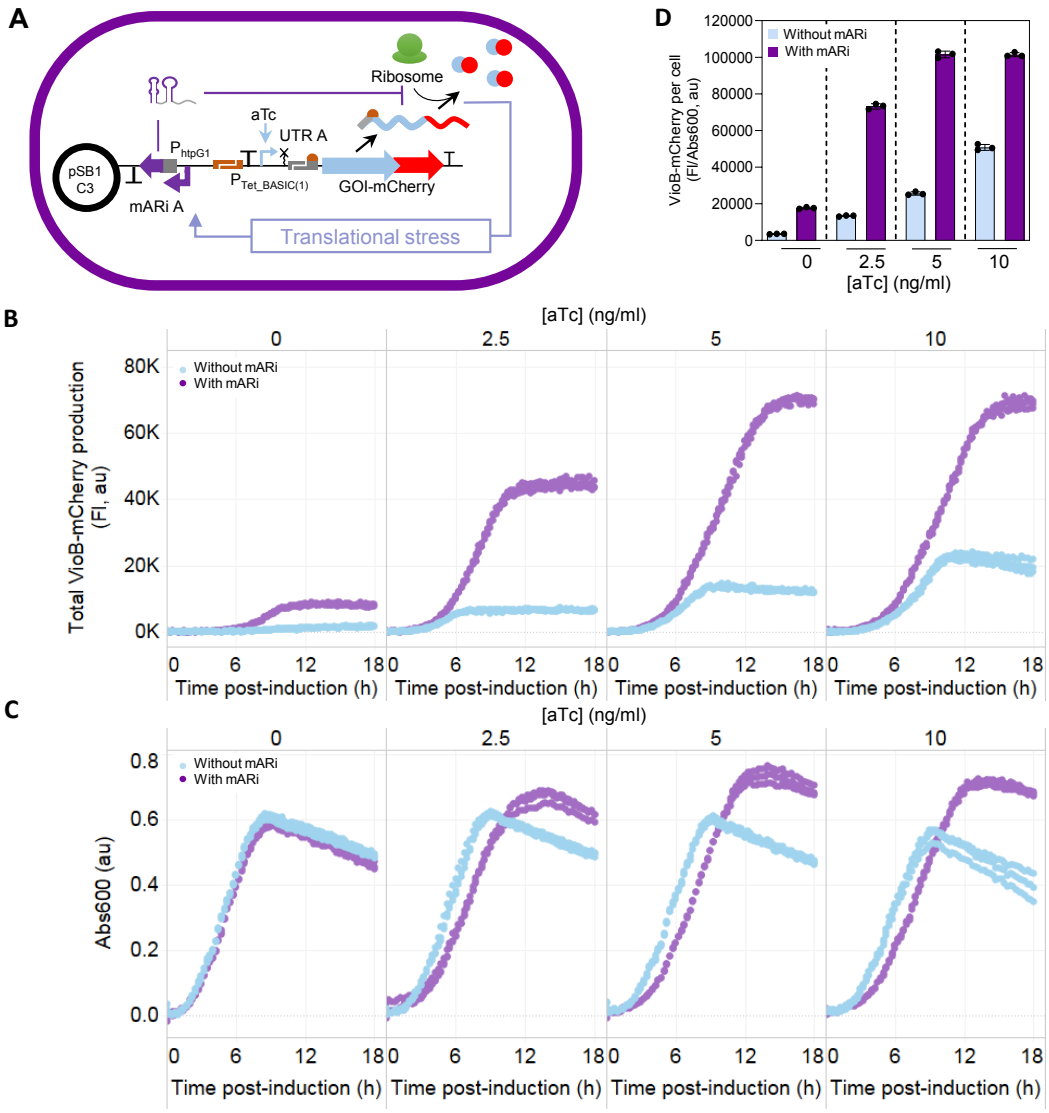


Figure 4.8 The implementation of mARi-based negative feedback to improve VioB-mCherry production in *E. coli* DH10b strains. (A) A schematic of a dynamic negative feedback system implemented in a VioB-mCherry production. (B) The accumulation of VioB-mCherry and (C) growth profile over the time of induction with aTc concentration 0-10 ng/ml. The engineered strains were then grown in LB media at 37°C and 600 rpm. The transcription of VioB-mCherry was triggered by aTc induction at concentration 0-10 ng/ml. The readout of mCherry and Abs600 were then monitored in a plate-based continuous measurement. (D) The VioB-mCherry production per cell in response to varied aTc induction. Data were taken at 18 h post-induction. Dots represent individual biological replicates. Bars represent the average of triplicates data points (denotes as a single dot) ± SD.

A similar improvement in productivity (Figure 4.8B) and exponential growth extension (Figure 4.8C) was observed in the DH10b cells over a period of 18 h post-induction. The improved growth was clearly pronounced in the DH10b cells with a feedback controller. Meanwhile, the strains without biomolecular controller showed a decline of cell density after entering stationary phase. In addition to the previously observed in BL21(DE3) strains (Figure 4.5C), the discrepancy of the growth in the early logarithmic phase was also shown in DH10b cells. Notably, up to a 5-fold yield improvement was obtained in the expression of mARi-based negative feedback controller with 5 ng/ml aTc induction (Figure 4.8D). The data obtained in

this section revealed the portability and robustness of the feedback system to enhance stress-inducing protein production across different tested host strains.

4.7 mARi-based negative feedback exhibits robust performance at different incubation temperatures

As a function of feedback control in genetic circuits, the utilisation of mARi-based negative feedback controller was intended to robustly improve protein production in diverse growth conditions. I therefore tested, the robustness of the designed feedback controller against external perturbations such as temperature. Temperature is known as one of the determinants that affects cellular component synthesis and bacterial growth[83,227–229]. To test this, the engineered DH10b strains with regulated and unregulated circuits were cultured at two different temperatures, 30 and 37°C.

The incubation of the engineered DH10b strains at 37°C resulted in a short time to attain maximal protein production (Figure 4.9A) and reach stationary phase (Figure 4.9B). Lowering incubation temperature resulted in lower overall VioB-mCherry production and growth. In particular, the DH10b cells with the P_{htpG1}-mARi controller showed a longer lag phase than the strains without feedback controller, although at 24 h post-induction both strains reached a similar cell density. Interestingly, a slightly higher total VioB-mCherry expression of strains with the biomolecular feedback controller was observed at 30°C incubation than 37°C after 24 h post-induction. Furthermore, yield improvement of VioB-mCherry expression across two different temperatures were unaffected by the presence of biomolecular controller (Figure 4.9C). These observations further confirmed that the mARi-based negative feedback controller is robust to stress-inducing protein production under different incubation temperatures.

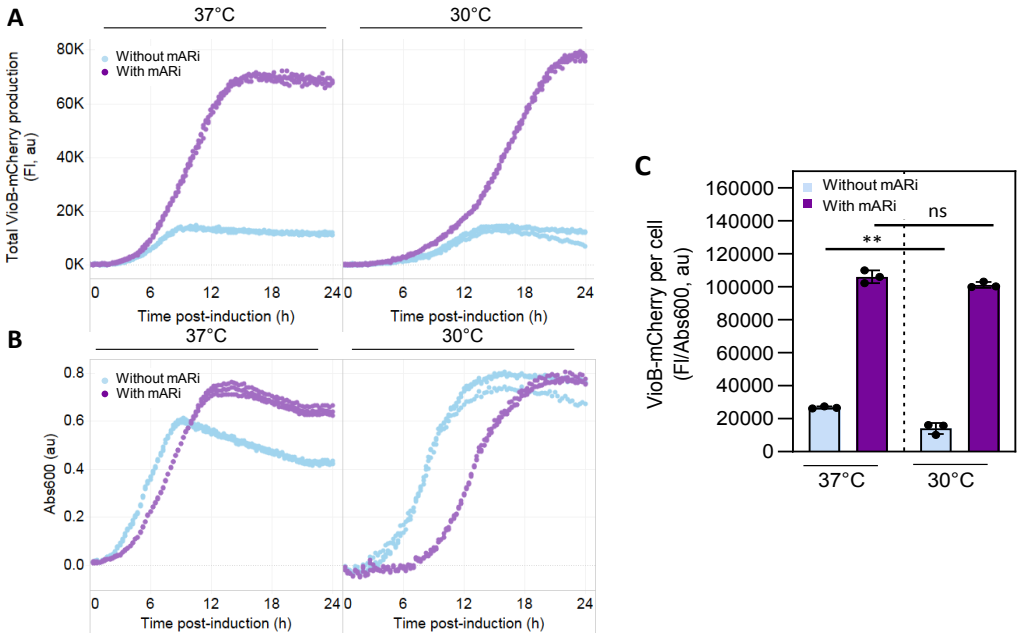


Figure 4.9 The robustness of mARi-based dynamic feedback in different incubation temperatures. (A) Total VioB-mCherry production and (B) growth profile over the time of induction with 5 ng/ml of aTc. The DH10b strains were grown in LB media at 37°C (left panel) and 30°C (right panel). (C) The VioB-mCherry production per cell at 24h post-induction. Bars represent the average of triplicates data points (denotes as a single dot) ± SD. Statistically significant differences determined using Student's t-test (** represents $p<0.01$, ns represents not significant).

4.8 mARi-based negative feedback is robust in different growth media

Another key aspect that influences bacterial growth and protein production is the nutrient composition in the growth medium. Essential nutrients such as phosphorus, nitrogen, and carbon are required for the synthesis of intracellular components, especially for protein and nucleic acid synthesis[230]. The limitation of these nutrients leads to three distinct strategies of ribosome allocation employed to achieve similar protein production rate under a slower growth rate. As such, different growth media may affect the performance and robustness of the mARi-based negative feedback systems in improving VioB-mCherry production. To investigate this hypothesis, *E. coli* DH10b cells with and without the P_{htpG1}-mARi controller were grown in different media to represent the different components of the essential nutrients. Two complex media (LB and 2YT+4% glucose) and a defined medium (EZ MOPS RDM+4% glucose) were used to culture the DH10b strains.

At a slower growth (using 30°C), both production strains showed a consistent protein production (Figure 5.10A) and growth pattern (Figure 4.10B) in all growth media. A relatively

constant improvement of VioB-mCherry overexpression was also observed across three different media (Figure 4.10C). These results prove the robustness of the stress-driven feedback controller to improve productivity despite varying growth medium compositions.

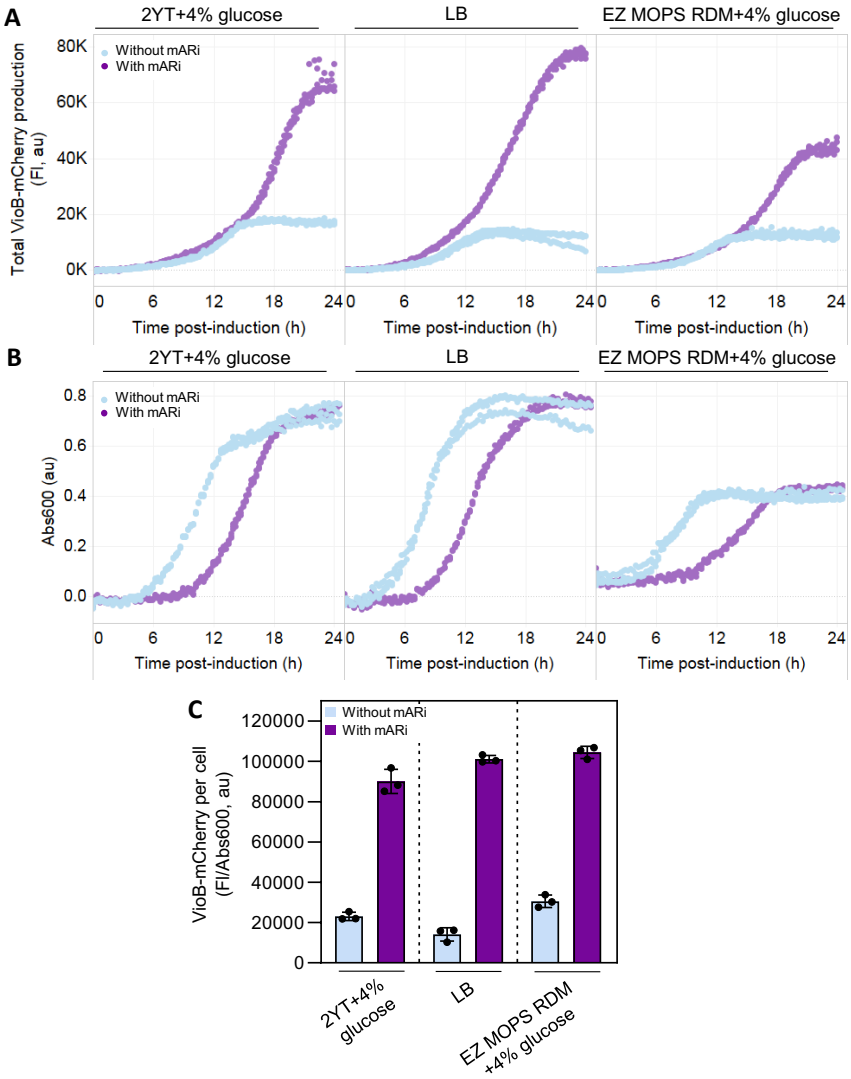


Figure 4.10 The robustness of mARi-based dynamic feedback in a variety of growth media. (A) Total VioB-mCherry production and (B) growth profile over the time of induction. The DH10b strains were grown in 2YT+4% glucose (left panel), LB (middle panel), and EZ MOPS RDM+4% glucose (right panel). The strains were then grown at 30°C and 600 rpm with 5 ng/ml aTc inducer. (C) The comparison of VioB-mCherry production per cell calculated at 24h post-induction. Bars represent the average of triplicates data points (denotes as a single dot) ± SD.

4.9 mARi-based negative feedback displays robust performance in scale-up production

With the robustness of the mARi-based negative feedback system shown across different host strains (Sections 4.5 and 4.6), incubation temperatures (Section 4.7) and growth media

(Section 4.8), the next step was to investigate its robustness in the scale-up of production by using different working volumes, geometric containers, and shaking speeds. This experimental setup was designed to represent the different physical and mechanical stresses inflicted on engineered strains and how they affect growth and protein production. In brief, the robustness of the feedback to improve the production of VioB-mCherry were tested at two different production scales in: (i) 100 µl cultures grown in 96 well plates with agitation at 600 rpm (Mikura shaker) and (ii) 50 ml cultures grown in 250 ml baffled flasks with agitation at 220 rpm (Innova shaker).

In both culture conditions, the strains with mARi-based feedback exhibited significant improvement of VioB-mCherry production over the strains lacking feedback controller (Figure 4.11A). Meanwhile, the total VioB-mCherry remained constant for the non-controlled systems at all conditions. With the feedback controller, the VioB-mCherry expression was greater in a 50 ml production volume than 100 µl production volume. The cell density of engineered strains grown in baffled-flasks were found to be higher than those in a microplate (Figure 4.11B). As a result, the yield improvement remained constant for two different growth conditions (Figure 4.11C). The improvement of VioB-mCherry production in a 50 ml further confirmed by a spun-down pellet from 1 ml of triplicate bacterial cultures (Figure 4.11D). In this section, the utilisation of mARi-based controller has been shown to maintain the robustness of the system in larger scale production condition.

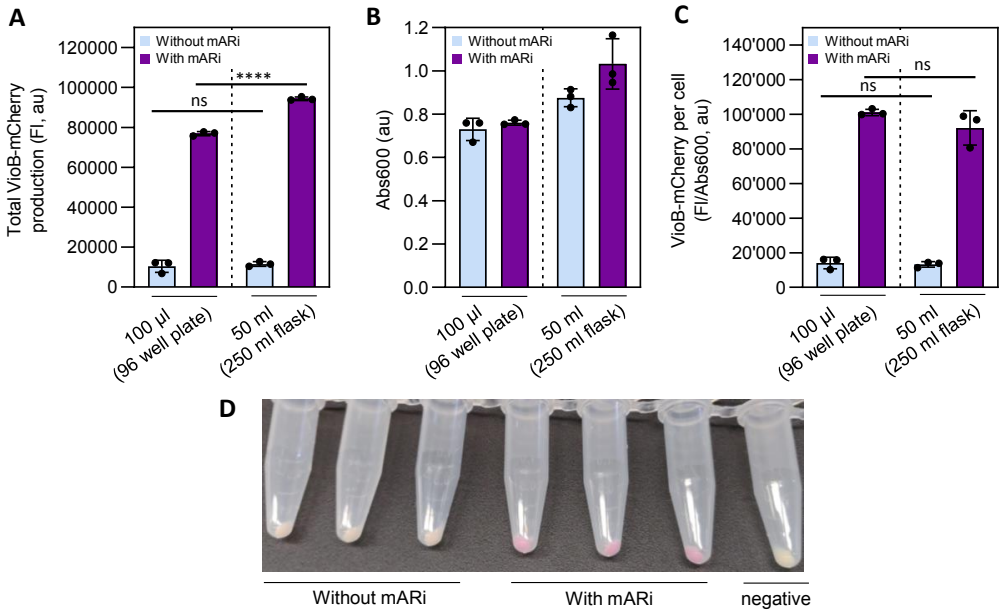


Figure 4.11 The robustness of mARi-based dynamic feedback in different working volumes. (A) Total VioB-mCherry production, (B) Abs600, and (C) The yield of VioB-mCherry production per cell at 24 h post-induction. The DH10b strains were grown in LB medium at 30°C and 600 rpm with 100 µl culture (96-well plate) and 220 rpm with 50 ml culture (250 ml flask). (D) Pellet from 1 ml of the culture grown in 50 ml culture after 24 h post-induction. 5 ng/ml aTc was added to induce the expression of VioB-mCherry after 1 h of outgrowth. Bars represent the average of triplicates data points (black dots) ± SD. Statistically significant differences determined using Student's t-test (**** represents p<0.0001, ns represents not significant).

4.10 Overexpression of an alternative stress-inducing protein results in a distinct profile of host-stress response

In addition to its robustness, the mARi-based negative feedback controller can be broadly applied due to its modular design. To demonstrate this modularity, the mARi-based negative feedback system was next employed to optimise the production of a different stress-inducing protein. An endogenous protein, β-galactosidase (LacZ) was chosen to exemplify an alternative stress-inducing protein. This protein has a native function to break-down disaccharide lactose into glucose and galactose. LacZ is encoded by the *lacZ* gene which from part of the Lac operon and has been widely used as a visible cloning marker in blue-white screening[231].

The *lacZ* gene has a 3 kb length and was extracted from the genome of *E. coli* BL21(DE3) using PCR amplification (Section 6.2.11). The extracted coding sequence without a stop codon was then fused onto the coding sequence of *mCherry* to enable online measurement of protein production. A DNA linker that functions to connect two proteins[62], namely the fusion linker

(FL2), was used to connect the 3' end of *lacZ* and 5' end of *mCherry*. Similar to the previously constructed negative feedback system made with the VioB-mCherry expression cassette, the expression cassette of LacZ-mCherry was created by using UTR A-RBSa to drive the LacZ-mCherry expression under the control of $P_{Tet\_BASIC(1)}$.

P_{htpG1} is supposed to be as a universal burden sensor for different types of protein expression[82]. Following the work in Sections 4.3 and 4.4, the activation response of P_{htpG1} was first investigated upon the induction of VioB-mCherry (Figure 4.12A) and LacZ-mCherry (Figure 4.12B) expression. To confirm the host-stress activation, the *E. coli* BL21(DE3) strains were co-transformed with expression plasmids and the P_{htpG1}-sfGFP plasmid[82]. Two different concentrations of aTc inducers were used to express these two different proteins: (i) 5 ng/ml of aTc for VioB-mCherry, and (ii) 50 ng/ml of aTc for LacZ-mCherry. The total protein production, P_{htpG1} activity, and growth profile were monitored over the time of induction in a plate-based continuous measurement (Section 6.3.4).

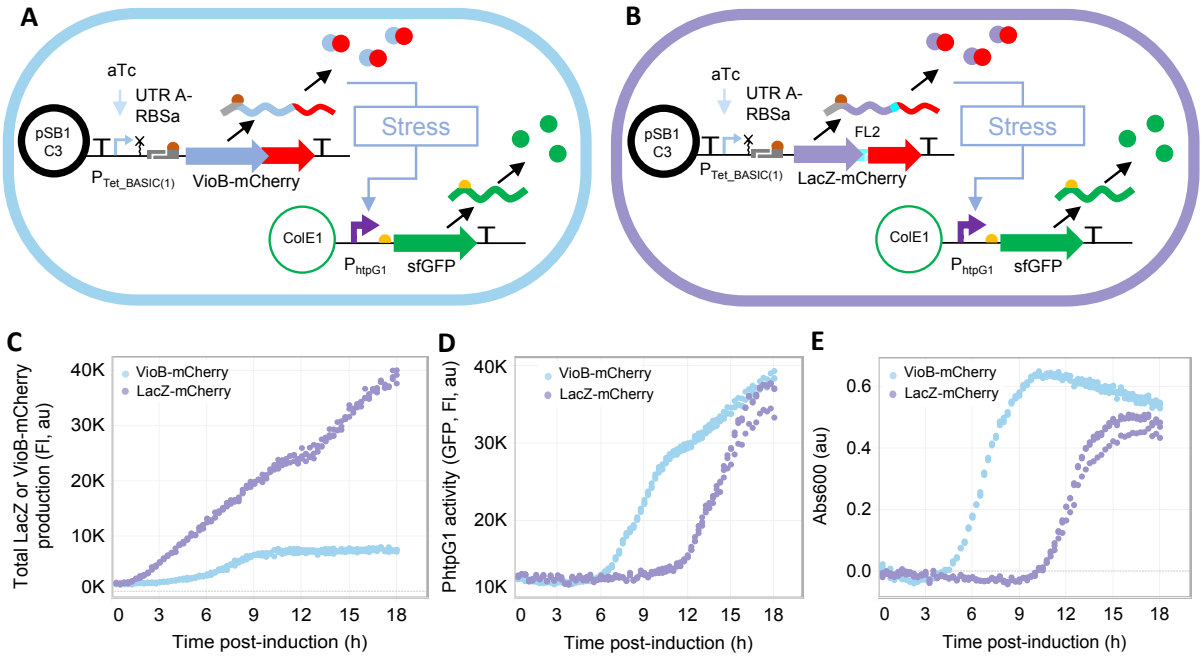


Figure 4.12 Characterisation of a stress-sensitive promoter, P_{htpG1} over the time of induction for VioB-mCherry and LacZ-mCherry producer strains. A schematic of a system used to evaluate the native cellular stress-response towards the expression of 2 different recombinant protein productions: (A) VioB-mCherry and (B) LacZ-mCherry. $P_{Tet\_BASIC(1)}$ and a strong RBS (RBSa) were used to drive the expression of VioB/LacZ-mCherry in the expression system. The system was constructed of the recombinant protein expression plasmids in a pSB1C3 backbone and a reporter plasmid containing P_{htpG1} to drive the expression of sfGFP in a ColE1 backbone[82]. (C) The total VioB/LacZ-mCherry production, D) P_{htpG1} activity readout, and E) growth profile over the time of induction. The engineered BL21(DE3) carrying these expression plasmids were then grown in LB media at 37°C and 600 rpm. After 1 h of outgrowth, 5 ng/ml of aTc and 50 ng/ml of aTc were added for VioB-mCherry and LacZ-mCherry, respectively. Dots represent independent data point of three independent replicates.

As expected, the LacZ-mCherry expression induced with 50 ng/ml aTc produced a higher yield than VioB-mCherry expression induced with 5 ng/ml aTc (Figure 4.12C). The difference of total protein production shown by red fluorescence signal was a consequence of the different types of protein and inducer concentrations used. Typically, endogenous sequences such as *lacZ* are preferable to be expressed than heterologous sequence i.e. *VioB*. This produced a distinct profile of host-stress response as can be seen from the P_{htpG1} activation (Figure 4.12D). Compared to VioB-mCherry, a slower activation of P_{htpG1} was observed during LacZ-mCherry production. Moreover, at the end of the incubation (18 h post-induction), both VioB-mCherry and LacZ-mCherry expression reached a similar output of P_{htpG1} activity. In spite of having a similar transcript length, the different codon compositions for both expression cassettes may cause different host-stress responses. In addition to the total protein production and host-stress response, each protein resulted in a distinct growth profile (Figure 4.12E). A slightly extended lag phase was observed for strains expressing LacZ-mCherry.

Through this initial investigation, it was shown that the P_{htpG1} can be used to sense the host-stress induced by LacZ-mCherry overexpression. Thus, the potential implementation of an adaptive feedback control could be explored to optimise the production and growth of engineered BL21(DE3) strains expressing LacZ-mCherry.

4.11 mARi-based negative feedback is modular for different stress-inducing proteins expression

Having shown that the P_{htpG1} was responsive to LacZ-mCherry overexpression, the P_{htpG1}-mARi was next used to optimise the production of this recombinant protein. In particular, the P_{htpG1}-mARi A was employed to target UTR A used in the LacZ-mCherry expression cassette (Figure 4.13A). Both the mARi-based controller and LacZ-mCherry transcriptional units were assembled in a single plasmid system (pSB1C3 backbone) to create a closed-loop feedback control system. The plasmid without the mARi-based controller was used as a control for an open-loop system. Both assembled plasmids were then propagated in *E. coli* BL21(DE3) cells.

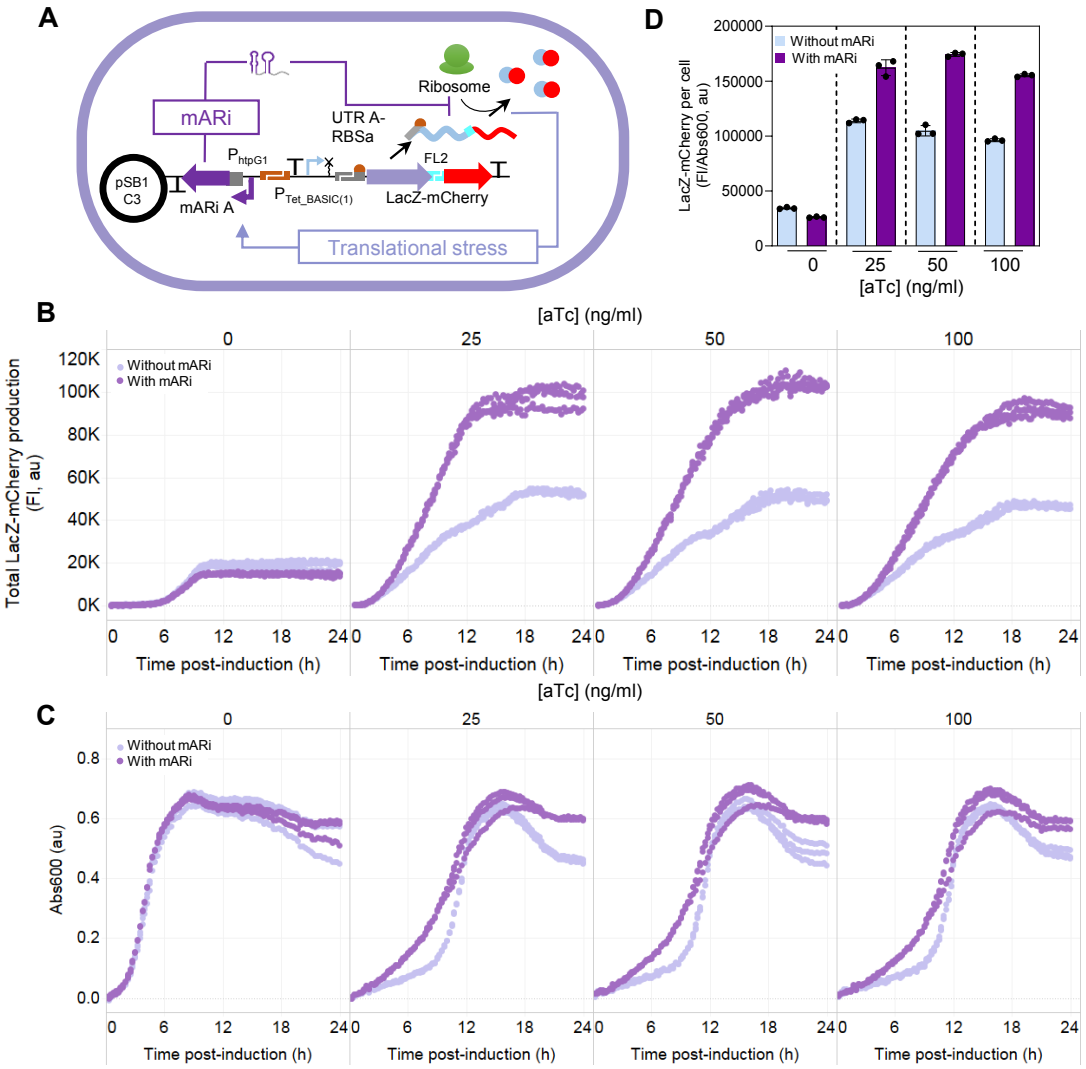


Figure 4.13 The implementation of mARi-based negative feedback to improve LacZ-mCherry production in *E. coli* BL21(DE3) strains. (A) A schematic diagram of a dynamic negative feedback system implemented in a LacZ-mCherry production. (B) The accumulation of LacZ-mCherry and (C) growth profile over the time of induction with aTc concentration 0-100 ng/ml. The cells were grown in LB media at 37°C and 600 rpm. D) The LacZ-mCherry production per cell in response to varied aTc induction. Data were taken at 24 h post-induction. Bars represent the average of triplicates data points (denotes as a single dot) ± SD.

Like the previous experiments (Section 4.5), the presence of mARi-mediated post-transcriptional control increased the total production of stress-inducing protein (i.e. LacZ-mCherry) up to two-fold (Figure 4.13B). As the feedback strength is influenced by the activity of P_{htpG1}, in this case, the different pattern of P_{htpG1} activation affected protein production. From the production curve, there is a small bump at 12 h post-induction, which is presumably the time when P_{htpG1} is highly activated (Figure 4.12D). Interestingly, the P_{htpG1}-mARi system also slightly enhanced the growth of engineered strains by shortening the lag phase and increased cell density after 24 h post-induction (Figure 4.13C). With the feedback system, at

all induction levels, the yield of LacZ-mCherry was higher than strains lacking the mARi-based negative feedback controller (Figure 4.13D).

The stress-driven negative feedback system presented here is also supported by the modular construction of genetic circuits using BASIC DNA assembly[62]. Specifically, this feature enables the easy changing of reusable bioparts required in the construction of expression cassettes such as a promoter, RBS, backbone, and protein of interests. In this section, the adaptive negative feedback was shown to improve the production of different stress-inducing proteins, without further tuning of the feedback regulator or even modify the host genotype. Therefore, our approach is of great use for the adoption of the negative feedback controllers for various genetic designs.

4.12 Summary

The overexpression of recombinant proteins (e.g. VioB-mCherry) driven by a strong RBS in a high plasmid copy number can significantly reduce the shared expression resources in a cell, limit recombinant protein production and trigger host-stress response. Shared translational resources (i.e. ribosomes) have been perceived to limit the production of recombinant proteins. In this chapter, a stress-responsive biomolecular feedback controller was proposed and built to effectively modulate translational resource allocation. This adaptive controller allows for dynamic and rapid switching of gene expression at a post-transcriptional level upon the triggering of cellular stress, e.g. by the overexpression of inducible constructs. The feedback strength is governed by a combination of an mARi-based gene repression system and a native promoter related to the heat-shock response, P_{htpG1}[82]. In the VioB-mCherry production strains (using *E. coli* BL21(DE3)), the total VioB-mCherry production resulted in up to 4-fold improvement in the presence of P_{htpG1}-mARi compared to the strains lacking feedback controller. Interestingly, the presence of a dynamic feedback system also extended the exponential growth phase of production strains during growth at 37°C. These improvements have also been shown in the cloning strain (DH10b) by up to 5-fold improvement, demonstrating the portability and robustness of the adaptive negative feedback controller for different tested host strains. Furthermore, the mARi-based negative

feedback features robust performance across different incubation temperatures, growth media, and working volumes. As a modular and portable controller, the P_{htpG1}-mARi system was then directly implemented to optimise the production of a LacZ-mCherry fusion protein. With the adaptive feedback controller, the productivity of LacZ-mCherry was increased by up to 200%. This was only required substituting the coding sequence of the GOI, without further tuning of the feedback controller system. Remarkably, this work demonstrates one implementation of mARi-based gene regulation system that is modular, robust and portable for biotechnological application, especially for improved protein production under stressful conditions.

Chapter 5

Overview, conclusion, and future work

5.1 Overview

Beyond their complexity and stochasticity, living cells can be exploited as a programmable chassis with varied engineered functionalities[232,233]. These programmable functions are encoded in the genetic circuits that can be built, modified, and inserted into host cells. As such, the encoded functions in the synthetic circuits are executed by employing endogenous machinery. Typical biological circuits are built using a bottom-up approach, where genetic devices and simple modules are combined to compose larger and more complex systems[232–234]. The capability to customise biological circuits has moved further from simple traditional genetic engineering to applying engineering concepts in the design and construction of a library of circuits[142]. Recently, remarkable progress has been made and underway to make engineering biology easier, faster, more robust, and reliable, including the standardised manufacturing process and modularisation of genetic toolboxes[175].

Further to this, the advancement of modern DNA assembly methods has provided standardised and modular construction of various synthetic circuits. This allows for the implementation of a modular design approach and increases the capabilities in the design construction of a variety of constructs. One of standardised and modular DNA assembly frameworks used thorough this project is BASIC DNA assembly[62]. In particular, a set of synthetic and orthogonal DNA linkers attached to standardised and modular bioparts and determine the order of assembly[34,62]. To this end, the domestication and diversification of existing bioparts are required to efficiently design and rapidly prototype complex synthetic genetic circuits. This project has focused on expanding biomolecular components, especially the standardised and modular a set of bacterial promoters at the transcriptional level, and modular RNAi system for gene regulation at the post-transcriptional level.

The first effort to standardise genetic controller was focused on bacterial promoter as an important regulatory element to initiate gene transcription. In particular, the physical structure of a set of bacterial promoters was systematically standardised while retaining their modular structure and compatibility to BASIC DNA assembly[62]. In this standardised architecture, a synthetic terminator and UP element were incorporated in the upstream core

promoter of interest to prevent unwanted interaction from/to the previous transcription. A self-cleaving ribozyme was placed immediately downstream core promoter sequence, aiming to standardise 5′ UTR sequences within translational initiation region. This strategy relies on automated cleavage activity of mRNA processing at a specific site to remove variability of extraneous base-pair in the 5′ UTR region[169,178]. This simple strategy in standardising 5′ UTR was achieved without the need to produce additional burdensome protein like Csy4 protein in a CRISPR-based RNA processing[174]. This strategy also provided a competitive advantage over the use of bicistronic design[235], where double RBSs are used in the design that may result in ribosome sequestration.

The standardised unit of the bacterial promoter was created with unique BASIC flanking sequences in its 5′ and 3′ boundaries as adhesive ends for attachment of BASIC DNA linkers of interest. This combination allows for modular design construction and gene expression, in particular with the use of UTR-RBS linker downstream the standardised promoter. Notably, the utilisation of ribozyme insulator and UTR-RBS linker within constitutive and inducible expression cassettes resulted in a greater output level than expression cassettes lacking ribozyme insulator and UTR-RBS linker. Presumably, this improvement was caused by the increase in mRNA stability[181,182].

To enable the use of standardised bacterial promoter for regulating multiple transcriptional units, three alternative promoter groups were devised by differing terminators and ribozyme sequences while preserving the UP element. Throughout this work, eight constitutive promoters and seven inducible promoters were standardised within these promoter groups and further functionally characterised. The performance of standardised inducible promoter was modulated by the balance ratio between the abundance of transcription factor, copy number of operator site, and dosage of external inducer. Functional characterisation and implementation of standardised inducible promoters in host-strains with genome-integrated regulatory cassette (*E. coli* marionette strains[80]) further demonstrated the compatibility and portability of these promoters. Additionally, this simplifies the implementation of standardised inducible promoters for multiplexed and independent control of multi-gene expression as shown in Sections 2.12 and 2.13 with avoidance to include a huge regulatory cassette in the construction of synthetic circuits.

In order to expand the standardised and modular biomolecular controllers, the work then continued with the development of a genetic controller at post-transcriptional level. In particular, the use of modular Artificial RNA interference (mARi)-based regulatory system to downregulate gene expression was investigated. The mARi-based regulation was developed using a native sRNA scaffold, namely MicC, which can be recognised by native Hfq chaperone. In this regard, the use of endogenous regulatory machinery is beneficial for enabling a portable, fast response, and lean repression system when compared to other controllers such as transcription factor, zinc finger, TALE, or CRISPRi, where the expression of accessory proteins can potentially be toxic, add cellular burden, and impact host morphology[236–238].

The modular design of mARi is inherited from the highly composable structure of *trans*-encoded sRNAs. A modification in the seed sequence whilst retaining the sRNA scaffold has enabled reprogramming natural sRNA with any target of interests that do not exist in nature[102,107,188]. To increase their modularity, target sites in mARis were rationally designed based on the sequence within UTR-RBS linkers for BASIC DNA assembly. Sequences specific to the UTR-RBS type while independent from RBS and GOI contexts were shown as an ideal target site for developing mARis. Thus, mARis can be deployed to control gene expression once corresponding reusable UTR-RBS linkers are used in the design of expression cassettes. This approach provides an advantage over earlier works, especially by avoiding the use of leader sequence[188,239], bespoke design[144], and target insertion[104]. Additionally, the standardised UTR-RBS linkers have been computationally generated and validated to ensure their orthogonality in the DNA assembly process[34]. By targeting these DNA linkers, it was expected that the mARi exerts an orthogonal post-transcriptional gene regulation with a reduced cross-interaction to non-cognate target, bioparts, plasmid backbones, and host genotypes[34].

By applying the same design principle, the new set of orthogonal mARis could easily be generated by substituting the target/seed sequence and employing the MicC scaffold. In total, five pairs of orthogonal mARis/UTRs were created and tested in this work. These mARis were operated for a multiplexed and simultaneous post-transcriptional regulation in multi-gene systems, including multiple transcriptional units and operon. Remarkably, this

implementation was carried out simply by assembling regulators and targets using BASIC DNA assembly without re-optimisation of the regulator.

Gene repression is achieved by directing mARi to the standardised sequence within 5' UTR region of the mRNA. As the mARi targeted the TIR region, the repression mechanism was presumably achieved through the steric hindrance of the RBS that prevents ribosome recruitment. This translation initiation inhibition was perceived to be mediated by Hfq chaperone recruitment to MicC scaffold. This is in agreement with previous works which describe the role of native Hfq in sRNA-based gene repression[104,240]. It is important to note that the designed mARi omits RNase E binding site, which is essential for RNase E-dependent cleavage. RNase E has a crucial role to accommodate rapid degradation of sRNA and mRNA target[241]. However, RNase E is also known to degrade unpaired sRNA prior to mRNA pairing[110]. The elimination of RNase E in mARi design was anticipated to reduce background degradation of free mARis and improve their half-lives. Another benefit of the RNase E site exclusion is that mARi-based regulation can be implemented in various tested *E. coli* strains independently from host genotype (Section 3.11).

Although translation initiation inhibition without active degradation of the mRNA has been reported to repress gene expression sufficiently[242], this silencing pathway is likely to result in intermediate and transient gene repression. Functional characterisation has further confirmed that the efficacy of mARi is modulated by relative expression ratio, molecular copy number, and spatial organisation of transcripts. Expressing mARi in a high expression ratio and a single plasmid system with medium plasmid copy number resulted in up to 80% repression activity.

In fact, such intermediate and transient gene regulations have an important role in the regulatory network, either in natural and engineered systems[104,127,143]. For instance, *trans*-encoded sRNA has been identified to actuate response system in the native stress-related adaptation[91,241]. This potential implementation is further supported by regulatory property, design modularity, and dynamic feature of RNA-RNA interaction. Additionally, the RNAi-based regulation system has been proposed as an actuator in a synthetic negative feedback control[144,221,243]. Given this prior knowledge and inspired by native stress response, mARi was

then functionalised as an ideal actuator in a dynamic feedback control to maximise recombinant protein production.

Recombinant protein production has been known as an important and interesting biotechnology application with a huge industrial interest[11,12,14]. However, the over-expression of recombinant protein in a certain design can impose low yield production, diminish cell fitness, and trigger host-stress response[82,155,211]. One of the possible reasons for this is the competition of shared cellular resources with endogenous gene expression when the resources are limited[82,153,210,211]. The overexpression of a long fusion protein, VioB-mCherry which is driven by a standardised inducible promoter and strong RBS in a high plasmid copy number turned on host-stress response. This was reflected by activity of previously identified stress-responsive promoter, P_{htpG1}[82] (Sections 4.3 and 4.4). Balancing gene expression through careful selection of regulators controlling GOI has been proposed as a strategy to resolve this issue[211,244]. Yet, this approach relies on bespoke design rules and further optimisation. On the contrary, the use of dynamic control offers more practical advantages than static control due to its automatic adjustment of gene expression[78].

In this project, a negative feedback loop was created by conditionally expressing mARi under the control of P_{htpG1}. Two key parameters determining the performance of this dynamic feedback were the magnitude of the promoter's activity driving mARi and dynamic profile of the actuator. The strength of P_{htpG1} dictates the amount of mARi and the feedback gain produced. A high expression ratio (mARi/mRNA) was substantially required to achieve effective repression (Section 3.7). A stronger activity of P_{htpG1} also stimulates greater feedback gain which is important for the robustness of the feedback system[221]. In terms of dynamic profile, stoichiometry titration through constitutively expressed transcripts has suggested that mARi exerts a graded-like response rather than ultra-sensitive manner in repressing its target (Section 3.7). This functional feature is suitable for developing a dynamic feedback control as the repression activity is proportional to the input of host-stress level[202].

Apart from the genetically encoded regulator, these unique biomolecular controllers take benefits from the existence of an endogenous regulatory system in host cells. In nature, this native regulatory machinery is a part of essential stress responses in adaptation

mechanisms[242]. In particular, the activation of host-stress response induces the expression of native Hfq chaperone[128]. Dynamic expression of mARi controller and Hfq in the presence of stress-inducing stimuli would consequently result in an efficient stress-driven biomolecular feedback system. Hypothetically, this rapid switching control at post-transcriptional level alleviates cellular stress by transiently reducing the amount of mRNA of interest. This is envisioned to create an effective proteome partitioning between inducible recombinant protein production and host metabolism maintenance. As a result, up to 5-fold improvement of total VioB-mCherry production was achieved with the conditional expression of mARi driven by P_{htpG1}. Interestingly, this adaptive feedback control also extended the exponential growth phase of production strains in a fast-growing condition (i.e. 37°C). The production improvement was robustly managed in a range of tested perturbations that may affect the fluctuation of cellular resources, including incubation temperature, growth medium, scale production, and host strain.

Another competitive advantage of this feedback system is its modularity inherited from the mARi and modular design assembly approach using BASIC DNA assembly[62]. The modularity of the design enabled the easy construction of synthetic circuits and the exchange of the protein of interest. A production improvement was observed with the presence of P_{htpG1}-mARi (Section 4.11), simply by substituting the GOI without further re-optimisation of the feedback regulator or even modify the host genotype. As an example, this adaptive feedback control has also been implemented to optimise production of a different stress-inducing protein (i.e. LacZ-mCherry).

As far as I am aware, this adaptive and robust mARi-based negative feedback system is the first example of stress-driven negative feedback control at the post-transcriptional level via reusable sRNA and is readily integrated into modular design framework. The mARi-based feedback provides a significant improvement over the existing stress-driven negative feedback control for dynamic optimisation of protein production, especially with its easy implementation, low cellular cost, portable, fast-responding, and modular features. Previous stress-driven feedback systems have been built based on transcriptional factor[155] and CRISPRi system[82]. As aforementioned, the utilisation of regulatory protein in such negative feedback adds extra metabolic load and creates delay[236–238]. Additionally, the strong binding of

gRNA/dCas9 to the mRNA target makes this approach less dynamic[82]. By contrast, temporal repression endowed by RNAi-mediated gene regulation enables reversible due to its nature RNA-RNA interaction, which is important for dynamic control of synthetic circuits.

As the engineering biology progresses from simple to more complex systems, the next engineered circuits would be constructed from well-characterised, standardised, and modular bioparts, manufactured with high-throughput DNA assembly platform, incorporated into dynamic interaction against host response, and confer robust performance towards different perturbations[177,245]. Overall, this project exemplifies a synergy between standardised, modular, and robust bioparts; modular design assembly for higher-level circuit design; and dynamic circuit-host interaction. Hence, it offers promising approach for future engineering framework of large and complex synthetic circuits with efficient and robust design.

5.2 Conclusion

This project demonstrated rational design and development, systematic evaluation, *in vivo* characterisation, and implementation of regulatory toolboxes in bacterial (i.e. *E. coli*) synthetic circuits. Both transcriptional and post-transcriptional controllers exhibited a relatively robust performance towards various genetic and cellular contexts. These regulatory properties are essential to guide their further utilisation in the design of synthetic genetic circuits. As an exemplar, the standardised inducible promoter was employed to upregulate recombinant protein expression whereas the intermediate and dynamic response of mARi was exploited to downregulate gene expression and create a stress-driven feedback control. This combination resulted in improved and robust production of recombinant proteins. These standardised, modular, and robust biomolecular regulators open up a new paradigm for efficient design of engineered biological circuits. The key to the design and development of these regulators is their integration into modular design and high-throughput DNA assembly framework, i.e. BASIC DNA assembly[62] thus greatly simplify their adoption and implementation in a rapid prototyping and modular construction of synthetic circuit of interest.

5.3 Recommendations for future work

While this project presents a valuable genetic toolbox that will be useful for the synthetic biology community, further follow-up experiments may be required to improve the performance and expand the diversity of standardised and modular regulatory platforms as well as their broadening their applications. Hence, there is a number of recommendations relevant for the continuation of this project in the future.

In this work, the library of standardised bacterial promoters has been developed and showed the fluctuated output level across different group promoters for the same set of core constitutive promoters (Section 2.6). In practical implementations, a near-identical gene expression across alternative group promoters is normally desirable. It is suspected that the varied residue in the sTSRV scaffold of the engineered RiboJ may significantly change the 5' UTR conformational and has a major contribution to the heterogeneous expression output (Section 2.4). The mutated non-essential sequences using Random DNA Sequence Generator (http://www.faculty.ucr.edu/~mmaduro/random.htm) resulted in this variable sequence[36]. For the next iteration, a new set of derived RiboJs may be generated with additional consideration of these sequence compositions function to maintain output expression level.

The *in vivo* characterisation of standardised inducible promoter sets would ideally be continued within a range of given inducer concentrations. This is aimed to complete characterisation data and confer kinetic parameter of standardised inducible promoters. The characterisation would be performed for engineered strains with the presence of plasmid-based and genome-integrated TFs. It may be interesting to compare the kinetic parameter of standardised inducible promoters in both cases.

This work has demonstrated that the standardised inducible promoters are compatible and portable in different marionette strains (Sections 2.10 and 2.11). Furthermore, this facilitates easy implementation, tight regulation, and independent control of multiple transcriptional units (Section 2.12 and 2.13). In the recent study, it has been shown that marionette strains can be implemented to modulate and balance multi-gene expression in a biosynthetic

pathway such as lycopene production[80]. The productivity is optimised by titrating concentration of corresponding chemical inducers. Therefore, it is possible to employ these standardised inducible promoters for various multi-gene pathways. However, careful promoter selection should be considered to maintain response function and minimise cross interaction between inducible systems of interest.

As an initial proof of concept, the mARi has been built based on MicC sRNA scaffold and shown significant repression of mRNA target. It is possible to further improve the efficacy of mARi through several strategies. Firstly, the mARi could be designed using different natural sRNA scaffolds (i.e. Spot42, SgrS, DsrA, MicF, etc)[101,104,106]. Secondly, the RNAi scaffold can be modified to change its binding affinity towards Hfq chaperone[102,137]. Thirdly, the sequence composition of seed/target sites can be refined to optimise their binding affinity. This approach may result in higher repression and orthogonality. Therefore, it should also be coupled with an improved algorithm to create a compatible, neutral, and orthogonal sequence for the assembly process using BASIC DNA assembly.

It is worth noted that the mARi system has been built using the Hfq-dependent scaffold that can be recognised by Hfq chaperone for an effective repression activity. Since this chaperone is highly conserved and present in a wide range of organism[120,123–125,133,134], it is possible to test the portability of the system in different organisms.

In Chapter 4, it has been shown that the use of stress-driven feedback control improves the production of recombinant proteins. The investigation of dynamic cellular mechanisms in the presence of this dynamic feedback control is one of interesting topics for further exploration using several advanced methods. Quantitative real-time PCR approach could be used to estimate and compare the abundance of transcript levels[104]. RNA-seq, Ribosomal profiling, and protein expression profiling strategies may also be employed to infer the dynamic resource allocation[246–248]. The dynamic behaviour of the engineered strains at a single cell level could also be investigated through *in vivo* super-resolution imaging system[249] to understand the dynamics of RNAPs and ribosomes over the time of induction. Furthermore, mathematical modelling could be developed to capture more insightful dynamic behaviour of the negative feedback system throughout this work.

Moreover, this negative feedback has successfully been applied to increase the production of recombinant proteins (VioB-mCherry and LacZ-mCherry) in a batch system (Chapter 4). The use of a batch production system has a limited amount of nutrients in the given volume over the production period. Scale-up of production can be conducted in a bigger capacity, for example in a 5-liter fermentor, and potentially through the use of a continuous production system. Additionally, the designed feedback system could be applied to various industrially relevant enzymes, such as laccase and cellulase.

The mARi-based feedback has been functionalised for creating an efficient design and dynamic control of engineered circuits. In particular, the expression of mARi was driven by a native stress-response (i.e. P_{htpG1}) that has been previously identified[82]. This promoter inherits its native activity. A modification in the core -35 and -10 elements could potentially increase the promoter activity for a higher feedback gain of the system. Further, this implementation is not only limited to stress-driven response. But also, it is possible to repurpose the negative feedback system for other inputs. With the underlying modular structure of mARi, the new negative feedback can be obtained by changing various responsive promoters (sensor) and retaining the actuator.

By using a modular design framework such as BASIC DNA assembly[62], multi-gene systems with alternative configurations can be easily created from interchangeable bioparts. Additionally, multiplexed and simultaneous gene regulation has been demonstrated to the multi-gene systems, but only limited to fluorescent reporters. In the future works, the designed feedback of multi-gene systems can be implemented to dynamically balance protein expression composing biosynthetic pathways during simultaneous growth and production phases. The dynamic system can be constructed from various possible combinatorial inputs and responses correspond to reduce toxicity of by-product accumulation, synchronise resource allocation, coordinate stress responses, and response to different physiological stages[156]. As a result, an improved growth, yield, and titer production of product of interest may be achieved.

Finally, the easy incorporation of the controller would be useful with the interconnection to different types of regulators for creating distinctive regulatory functions[146,250]. This further explores the potential implementation of mARi and feedback control beyond the application

in metabolic pathways. Moreover, the compact size, robust performance, compatibility, and portability of the controllers would facilitate further implementation in a multilayer regulatory network for robust performance of synthetic genetic circuits. This feature may be useful for developing future larger synthetic circuits with sophisticated functionalities and robust performance.

Chapter 6

Materials and methods

6.1 Materials

6.1.1 Bacterial strains

Bacterial cells were obtained from single colonies and stored at -80°C.

Table 6.1 Bacterial strains used in this study.

Strain	Genotype	Antibiotic resistance	Source
DH5α	(K-12 strain, F⁻ *endA1 glnV44 thi 1 recA1 relA1 gyrA96 deoR nupG purB20* ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^- m_K^+$), λ⁻)	None	NEB, #C2987H
DH10b	(K-12 strain, F-*mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ (*ara leu*) 7697 *gal*U *gal*K *rps*L *nup*G λ-)	None	Laboratory collection
DH10b_marionette (sAJM.1504)[80]		Chloramphenicol	Voigt Lab (Addgene plasmid #108251)
MG1655	(K-12 strain, F⁻ λ⁻ *ilvG⁻ rfb-50 rph-1*)	None	Laboratory collection
BL21(DE3)	(B strain, F⁻*omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm* (DE3))	None	Laboratory collection
BL21star(DE3)	(B strain, F⁻*omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm rne*131 (DE3))	None	Invitrogen, #C601003
BL21(DE3)_marionette(sAJM.1505)[80]		Chloramphenicol	Voigt Lab (Addgene plasmid #108253)

6.1.2 Media and carbon sources

The medium used for growing the bacteria strains are shown in Table 6.2. For sterilization purposes, all media were autoclaved (121°C for 15 minutes) or filter sterilized (0.22 µm) prior use.

Table 6.2 Bacterial growth medium used in this study.

Medium	Composition
LB (Luria-Bertani Broth)	25 g/l LB powder (Formedium)
LB agar	25 g/l LB powder (Formedium), 10 g/l agar (Formedium)
SOC	31.5 g/l SOC powder (Formedium)
2YT	16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl
EZ MOPS RDM	14 g/l EZ supplement+nucleotide (Formedium), 8.37 g/l MOPS (Formedium), 0.7167 g/l Tricine (SigmaAldrich), 3.56 g/l mineral mix (Formedium), 0.23 g/l K2HPO4 (SigmaAlfdrich)

Additional glucose solution was used as a carbon source. Glucose with stock concentration 40% was made by dissolving 400 g/l glucose (SigmaAldrich) in dH_2O. Glucose solution was then filter sterilised (0.22 µm) prior use and stored at 4°C.

6.1.3 Antibiotics

All antibiotics were dissolved in specific solution at the following stock concentration and filter sterilized (0.22 µm) prior use. The stock antibiotics were kept in the -20°C. To achieve working concentration, the antibiotics were diluted 1000x into fresh media.

Table 6.3 Antibiotics used in this study.

Antibiotic	Solvent	Stock concentration	Working concentration
Carbenicillin	dH_2O	100 mg/ml	100 µg/ml
Kanamycin	dH_2O	50 mg/ml	50 µg/ml
Chloramphenicol	Ethanol	25 mg/ml	25 µg/ml

6.1.4 Inducers

All inducers were dissolved in specific solution at the following stock concentration (Table 6.4) and filter sterilized (0.22 µm) prior use. The stock antibiotics were kept in the -20°C.

Table 6.4 Inducers used in this study.

Inducible promoter	Inducer	Solvent	Stock concentration
P_{Tet}	aTc	DMSO	100 µg/ml
P_{PhlF}	DAPG	Ethanol	25 mM
P_{BAD}	Arabinose	Arabinose	1 M
P_{Lac}	IPTG	Water	1 M
P_{Tac}	IPTG	Water	1 M
P_{Cym}	Cuminic acid	Ethanol	100 mM
P_{Cin}	3OHC14-AHL	DMF	10 mM
P_{Lux}	3OC6-HSL	DMF	10 mM
P_{Van}	Vanilic acid	Ethanol	100 mM

6.1.5 Kits

The commercially available kits in Table 6.5 were used in this study.

Table 6.5 Kits used in this study.

Kit	Supplier and catalog number	Used for
Miniprep I	Omega biotek, # D6943-02	Miniprep plasmid
Gel extraction kit	Omega biotek, # D2500-01	Gel extraction
Phusion polymerase	NEB, # M0530S	Site directed mutagenesis
CloneJet PCR Cloning Kit	Thermo Fisher Scientific, # K1231	Storage plasmid for BASIC bioparts

6.1.6 Buffers

All buffers were filter sterilized (0.22 µm) prior use apart from TBE (Tris-borate-EDTA) and running buffer.

Table 6.6 Buffers used in this study.

Buffer	Composition	Used for
Inoue Transformation Buffer (ITB)	10.88 g/l $MnCl_2.4H_2O$, 2.2 g/l $CaCl_2.2H_2O$, 18.65 g/l	Chemically competent cell, store at 4°C.
0.5 M PIPES pH 6.7	151 g/l PIPES	Chemically competent cell, adjust pH to 6.7 using KOH, store at -20°C.
100 mM $CaCl_2$	14.701 g/l $CaCl_2.2H_2O$	Chemically competent cell
Phosphate Buffer Saline (PBS)	1 tablet of PBS in 200 ml dH_2O	General use
TBE (Tris-borate-EDTA)	100 ml TBE 10x in 900 ml dH_2O	Agarose gel electrophoresis
10x SDS Running Buffer	30 g/l Tris base, 144 g/l glycine, and 10 g/l DS	SDS-PAGE

6.1.7 Dyes, enzymes, and reagents

The following dye, enzyme, and reagent were used in this study.

Table 6.7 Dyes and enzymes used in this study.

Dye and enzymes	Supplier and catalog number	Used for
SYBR Safe	Invitrogen, # S33102	Agarose gel electrophoresis
6x loading buffer	Home-made	Agarose gel electrophoresis
1 kb DNA ladder	NEB, # N3232L	Agarose gel electrophoresis
100 bp DNA ladder	NEB, # N3231S	Agarose gel electrophoresis
Protein ladder	Bio-rad, #1610374	SDS-PAGE
Commassie Brilliant Blue R-250	SigmaAldrich, # 1125530025	SDS-PAGE staining solution
*Bsa*I-HFv2	NEB, # R3733L	Restriction digest
*Dpn*I	NEB, # R0176S	Plasmid background digest
T4 Ligase	Promega, # M1804	Ligation
Magnetic beads (AMPureXP)	Beckman Coulter, # A63881	BASIC assembly
Pfu polymerase and buffer	Home-made	Colony PCR

6.1.8 DNA

All DNA sequence were synthesised as gBlocks from Integrated DNA Technology (IDT) and DNA fragments from TWIST Bioscience. DNA sequences of bioparts used in this work are listed in Appendix D. DNA linkers were synthesized from IDT and Biolegio. The sequence of DNA linkers used in this work are listed in Appendix G. Double stranded primers were synthesized from IDT with 5’ phosphorylation when used for site directed mutagenesis PCR. Primer was used for this work are listed in Appendix I.

6.2 Molecular biology methods

6.2.1 Bacterial culture preservation

The bacterial cultures used in this study were preserved in glycerol stock, both in Eppendorf tube and 96-well plate. 200 µl of 80% autoclaved glycerol was transferred into a 1.5 ml Eppendorf tube, then 200 µl of liquid bacterial culture was added and gently mixed. For 96-well plate format, 75 µl of 80% autoclaved glycerol was transferred into a 96-well plate Costar flat bottom (SLS, #3370), then 75 µl of liquid bacterial culture was added and gently mixed. The bacterial culture stock then was stored in -80°C.

6.2.2 Plasmid extraction

Plasmid DNA was extracted from *E. coli* DH5α using a miniprep kit from Omega Biotek kit in accordance with the manufacturer's instructions.

6.2.3 Quantitation of plasmid DNA concentration

DNA quantification from miniprepped plasmids was carried out using the NanoDrop OneC (Thermo Fisher Scientific) in accordance with the manufacturer's instruction.

6.2.4 Site directed mutagenesis and bioparts creation via PCR

Briefly, target DNA was amplified using 10 µl of 5X HF Buffer, 1 µl of 10mM of dNTPs 10 mM, 2.5 µl of forward primer 10 µM, 2.5 µl of reverse primer 10 µM, 50 pg of template DNA, 0.5 µl DNA Phusion Polymerase (NEB), and ddH_2O up to 50 µl. The following touchdown PCR protocol in the thermal cycler was used for site-directed mutagenesis; (1) 95°C for 3 min, start loop 25 cycles, (2) 95°C for 30 sec, (3) gradient from 65 to 70°C for 30 sec, (4) 72°C for 2 min, close loop, (5) 72°C for 10 min, and (6) 4°C hold forever. Primers used for mutagenesis purposes are listed in Appendix I. Subsequently, we continued ligation using 5 µl PCR product, 4 µl ddH_2O, 1 µl 10x T4 ligase buffer (Promega), and 0.5 µl T4 DNA ligase (Promega). Chemically competent cells of *E. coli* were transformed with 5 µl of ligation products. Sequence verification from mutagenesis of BASIC part was done using primer SP01 and SP02, while point mutation of assembled plasmid was verified using P07 and SP32 (Appendix I). The

list of bioparts and plasmids generated through site directed mutagenesis is listed in Appendix H.

6.2.5 Agarose gel electrophoresis

DNA fragments from restriction digest or PCR product were separated using 0.8-1% agarose gel prepared in 1x TBE buffer. 5 µl of sample was combined with 2 µl purple loading dye (6x, NEB) and 2 µl SYBR Green (10x stock, Invitrogen), and then loaded into the wells. The agarose gels were run at 100-125 V for 110-120 min. A DNA ladder (100bp or 1 kb, NEB) was used for size determination. DNA visualisation was performed under blue light and images were visualised using a Fuji LAS-3000 imaging system.

6.2.6 Gel extraction and DNA purification

DNA fragments were extracted from excised agarose gel pieces using Gel Extraction Kit (Omega Biotek) in accordance with the manufacturer's instructions.

6.2.7 BASIC assembly

Plasmid construction was done using the BASIC DNA assembly method as described in the original research[62,63] and updated version (manuscript under review). The details of BASIC DNA assembly could be found via BASIC assembly webpage (https://www.basic-assembly.org/blank-1).

6.2.7.1 BASIC bioparts preparation in the storage vectors

BASIC bioparts were generated by DNA-synthesis companies (IDT and Twist Bioscience), PCR mutagenesis to alter the sequences, and PCR to extract DNA fragments from *E. coli* genome. The BASIC bioparts should not contain internal *Bsa*I restriction sites. The DNA fragments of BASIC bioparts were designed with flanked BASIC prefix, *i*P (5'-TCTGGTGGGTCTCTGTCC-3') and suffix, *i*S (5'-GGCTCGGGAGACCTATCG-3') sequences. These DNA fragments can directly be ligated to specific DNA linkers for direct assembly process or stored into a storage vector. For the latter, the DNA fragments were ligated into methylated linkers (LMA and LMB). Thus, the ligation of bioparts into a storage vector enables reusing bioparts for different designs and assembly purposes. In the beginning of this study, a commercial storage vector, pJET1.2 (Thermo Fisher Scientific) was used. The pJET1.2 cloning was performed in accordance with

the manufacturer's instructions. Because pJET1.2 contains 3 recognition sites of *BsaI* that may reduce the effectiveness of BASIC assembly, we used custom pUC-Amp backbone for a storage vector in the later stage of this study. The custom pUC-Amp backbone integrated BASIC prefix and suffix, thus 2 parts BASIC assembly was carried out to assemble DNA fragments into a pUC-Amp backbone mediated by the use of methylated linkers (LMA and LMB). An ideal concentration of BASIC bioparts is about 50 ng/µl for each kb of DNA plasmid.

6.2.7.2 BASIC DNA linker preparation

As aforementioned in section 6.1.8, lyophilised BASIC DNA linkers were synthesised by IDT and Biolegio (*www.biolegio.com*). The BASIC DNA linkers synthesised through IDT consist of 2 separate fragments for each linker section: an adaptor (smaller fragment) and linker (longer fragment). The lyophilised BASIC DNA linkers were then eluted into TE buffer as a stock solution. The BASIC DNA linkers synthesised from Biolegio consist of a combined linker (small and long parts) for each linker section. The lyophilised BASIC DNA linkers from Biolegio were then eluted into 200 ul of linker annealing buffer provided as working solution. To facilitate the annealing process, both working solutions obtained from IDT and Biolegio then heated up at 95°C and after 5 min allow them cool down to room temperature. The stock and working solution were stored at -20°C until next use for up to 3 months.

6.2.7.3 BASIC linker-bioparts ligation

The BASIC linker-bioparts ligation were setup in a 200 µl tube with 30 µl total volume. The following reaction was setup for each BASIC linker-bioparts ligation: 3 µl 10x Promega T4 buffer, 1 µl of prefix linker, 1 µl of suffix linker, 200 nmol of BASIC biopart DNA (i.e. 1 µl if a 4kb plasmid is supplied at concentration of 200ng/µl), and dH_2O to 28.5 µl. 1µl of *Bsa*I-HFv2 (20 units, NEB) and 0.5 µl of T4 ligase (0.5-1.5 units, Promega) were added and mixed by pipetting up and down. The mixture was then incubated in a thermocycler with the following setup program: cycle 20x [37°C for 2 min, 20°C for 1 min], 55°C for 20 min, store at 4°C.

6.2.7.4 Linker-ligated bioparts purification

Linker-ligated part were done using magnetic beads purification. A fresh 70% ethanol should be prepared. 54 µl of magnetic beads (AMPure XP magnetic DNA purification kit, Beckman

Coulter) per reaction was transferred into 96 well Falcon plate (Falcon #351177). 30 µl of the BASIC linker-bioparts ligation mixture was added into the magnetic beads and resuspend by pipetting up and down several times. The mixture was incubated for 5 min at room temperature to allow DNA binding to magnetic beads. The whole Falcon plate was then placed onto the magnetic plate and left for 2 min to form immobilised rings. The solution was carefully aspirated from the centre of each well without disrupting the beads. 190 µl of 70% ethanol (Fischer Scientific) was added and carefully discarded after 30 sec incubation at room temperature. This step was repeated 2 times to remove the unbind DNA fragments. The ethanol left was evaporated by incubating the plate for about 5 min at room temperature. The plate was then removed from the magnetic plate and 30 µl of deionised water was used to elute the formed rings. Thorough mixing was required to ensure all beads were fully incorporated. After 1 min incubation at room temperature, the Falcon plate was then placed back onto the magnetic plate and left for 1 min to form immobilised rings. About 28 µl of the solution as purified product was carefully aspirated from the wells. If not immediately being used the purified product was stored at -20°C.

6.2.7.5 BASIC assembly from purified linker-ligated bioparts

The BASIC assembly from purified linker-ligated bioparts were setup in a 200 µl tube with the following reaction: 1 µl 10x NEB CutSmart buffer, 1 µl of each purified linker-ligated bioparts, and H_2O to total volume 10 µl. The mixture was resuspended by pipetting up and down and incubated in a thermocycler machine with the following setup program: 55°C for 45 min, store at 4°C. If not used for transformation immediately, the assembled mixture can be stored at 4°C over-night or at -20°C for up to a month.

6.2.8 Competent cell preparation

Chemically competent cells of DH5α, DH10b, and DH10b_marionette were prepared using Inoue protocol. A single colony of the bacterial strain was inoculated into 15 ml LB media in a 50 ml Falcon tube for overnight growth at 37 °C with shaking (220 rpm). 5-10 ml of overnight culture was inoculated into 300 ml fresh SOC media. The culture was then grown at 18°C and 220 rpm for 18-24 h until reach OD600 between 0.3 – 0.5. The culture was then transferred to 5 Falcon tube 50 ml and incubated on ice for 10 min before being centrifuged (4500 rpm, 5 min, 4°C). The harvested cells were resuspended gently with 5 ml cold ITB supplemented

with 2% PIPES pH 6.7. The mix then was combined in 2 Falcon tube 50 ml and added to 40 ml with chilled ITB+PIPES. After incubation for 10 min on ice, the mix then were centrifuged (4500 rpm, 5 min, 4°C). The pellet from 2 Falcon tube was resuspended again with 5 ml ITB+PIPES and combined together with a total volume 10 ml. 0.7 ml DMSO was added to this cell emulsion and incubate the cell for further 10 min on ice. The cells were then dispensed into 1.5 ml microtubes in 200 µl aliquots and stored at -80°C for later heat shock transformation. For quality control, a random aliquote was subjected to contamination test and quantify the transformation efficiency. The contamination test was done by spreading 50 µl of competent cells into agar plates with antibiotics (Kanamycin, Carbenicillin, Chloramphenicol, and Gentamycin). Plates then were incubated overnight at 37°C. Clean plates confirmed that there is no contamination issue on the batch of competent cells. Transformation efficiency test was done by transforming 25 µl of competent cell with 0.5 µl 50 pg/µl commercial pUC19 plasmid. 125 µl of SOC was added to the cells after heat-shock. 10 µl of the recovered cells were diluted into 990 µl SOC. 3, 30, and 300 µl of these diluted cells were then plated into agar plates supplemented with Carbenicillin. The average number of colonies per 1 µl of the diluted cells plated the multiplied by $0.6x10^9$ to obtain the CFU/µg pUC19 plasmid.

Chemically competent cells of BL21(DE3) and BL21(DE3)_marionette were prepared using $CaCl_2$ method. A single colony was inoculated into 5 ml LB and grown overnight. 1 ml of the overnight growth then inoculated into 100 ml LB in 250 ml flask. The culture was then grown at 37°C and 220 rpm until reach OD600 between 0.3 – 0.4. Keep the culture on ice for 15 min. Cells then centrifuged (5000 rpm, 4°C, 10 mins). The pellet was resuspended with 10 ml chilled $CaCl_2$ 100 mM. Then cells were incubated on ice for 20 min. Cells then were centrifuged (5000 rpm, 4˚C, 10 mins). Pellet was then resuspended in 5 ml chilled $CaCl_2$ 100 mM supplemented 15% glycerol. The cells were then dispensed into 1.5 ml microtubes in 200 µl aliquots and stored at -80°C for later heat shock transformation.

6.2.9 Heat shock transformation

5-10 µl assembly or 0.5-1 µl plasmid DNA (around 20 ng/µl) was added to 25-50 µl chemically competent cells which had been thawed on ice. The mix was gently mixed and incubated on ice for 5-10 min before being heat shocked at 42 °C for 45 sec. The cells were then incubated

on ice for 2 min and recovered in 0.5 ml of SOC for 1 h with shaking (220 rpm) at 37°C. The culture then was spun down and 400 ul supernatant was taken out. 100 µl of the remaining recovered cells was suspended and spread onto a LB agar plate containing the appropriate antibiotic to select transformants harbouring the plasmid of interest. The plate was incubated overnight at 37°C.

6.2.10 Assembly confirmation

The assembled plasmids were confirmed by colony Polymerase Chain Reaction (PCR), *BsaI* restriction digest, and Sanger DNA sequencing.

6.2.10.1 Colony PCR

The single colony was picked and mixed into 50 µl of deionised sterile water as template for PCR. The colony PCR was setup in a 200 µl tube with the following reaction: 2.5 µl 10x *Pfu* buffer, 0.5 µl of 10 mM dNTPs, 0.5 µl of 10 µM forward primer, 0.5 µl of 10 µM reverse primer, 1 µl of diluted colony and dH_2O to 24.5 µl. 0.5 µl of home-made *Pfu* polymerase were added and mixed by pipetting up and down. The mixture was then incubated in a thermocycler with the following setup program: 95°C for 5 min, cycle 30x [95°C for 30 sec, 60°C for 30 sec, 72°C for 2.5 min], 72°C for 10 min, store at 4°C. 10 µL of the PCR product was checked on a 0.8-1% agarose gel.

6.2.10.2 *Bsa*I restriction digest

Confirmation of assembled plasmid also could be done by performing *BsaI* restriction digestion. This strategy is facilitated by the use of methylated linkers (LMA and LMB) during the plasmid construction that allows the digestion of assembled plasmid to excise the fragment from the backbone. The *BsaI* restriction digestion was setup in a 200 µl tube with the following reaction: 1 µl 10x CutSmart buffer, 2 µl of miniprepped plasmid, and dH_2O to 9 µl. 1µl of *Bsa*I-HFv2 (20 units, NEB) was added and mixed by pipetting up and down. The mixture was then incubated in a thermocycler with the following setup program: 37°C for 30 min, 80°C for 10 min, store at 4°C. 10 µL of the digestion product was checked on a 0.8-1% agarose gel.

6.2.10.3 DNA sequencing

DNA sequencing was performed by Source Bioscience (Cambridge, UK) using Sanger DNA sequencing. 5 µl samples with concentration around 100 ng/µl was used as a template whereas 5 µl primers with concentration around 3.2 pg/µl were required for a sequencing reaction. Both samples and primers were prepared in a separate tube.

6.2.11 Extraction of *lacZ* gene and construction of *lacZ-mCherry* fusion

The *lacZ* gene was extracted from *E. coli* BL21(DE3) genome using a touchdown PCR method (Section 6.2.10.1). The template for fragment amplification was obtained from a single colony of BL21(DE3) diluted into 50 µl of dH_2O. A pair of primers, P205 and P206 (Appendix I) were used to amplify specific region flanking LacZ gene with BASIC prefix and suffix. The PCR product was then confirmed through gel agarose electrophoresis. Subsequently, a gel purification was performed to a band with a size of ~3 kb. The purified DNA fragment then was used directly for creating a *lacZ-mCherry* fusion by using a fusion linker, namely FL2. The *lacZ* fragment was flanked by LMA-P and FL2-S (Appendix G) whereas *mCherry* part was ligated into FL2-P and LMB-S (Appendix G). These linker-ligated parts were then assembled into a pUC-Amp backbone for creating a *lacZ-mCherry* fusion.

6.2.12 Sodium Dodecyl Sulphate-Poly Acrylamide Gel Electrophoresis (SDS-PAGE)

Confirmation of total VioB-mCherry production was performed using SDS-PAGE as described[251]. Overnight culture was diluted 200x in 10 ml LB media supplemented with chloramphenicol. After 1 h outgrowth, aTc inducer at final concentration 5 ng/ml was added. The culture then grown at 37°C and 600 rpm for 12 h followed by Abs600 measurement. About 1 ml culture with normalised Abs600 was then centrifuged (13,000 g for 1 min). 100 ul of 5x SDS loading dye (0.6 M TRIS-HCl pH 6.8, 25% glycerol, 2% SDS, 0.05% b-mercaptoethanol, 0.1% Bromophenol Blue) was added to the pellet followed by 2 min vortexing and 5 min boiling at 99°C. 5 ul of sample mix was then loaded into SDS-PAGE with concentration 12% of separating gel and 5% of stacking gel. 100 mL of 10 x SDS running buffer (250 mM TRIS, 1.92 M Glycine, 35 mM SDS) was diluted in 1 L of dH_2O and placed into the tank. After running gel at 100 V for 2 h in 1x SDS running buffer, the gel was stained with Commassie Brilliant Blue solution (1.2 mM Coomassie Blue R-250 dye, 45% Methanol, 10 % glacial acetic acid) for 10 min. The destaining process was performed with destaining solution

(20% methanol, 10% glacial acetic acid) in 2 h sections until most of the unbound stain was removed. The image was captured using a Fuji LAS-3000 imaging system.

6.3 Experimental assay methods

6.3.1 Continuous plate reader measurement

Time-series measurements were performed in 96 well plates with biological triplicates for each construct. *E. coli* carrying an empty backbone (containing only an origin of replication and an antibiotic resistance gene) was used as a negative control. *E. coli* colonies were inoculated from glycerol stock into LB media supplemented with appropriate antibiotics. The colonies were grown for 13 h at 30°C and 600 rpm in a benchtop shaker (Mikura shaker). The culture was then diluted 200x (10 µl into 190 µl then 10 µl into 90 µl) in LB media using an automated liquid handling robot (CyBio Felix). For the constitutive gene expression, the plate was then incubated in a microplate reader (Clariostar, BMG Labtech) with continuous shaking at 37°C or 30°C and 600 rpm for 12-24 h. For the inducible expression of the recombinant protein assay, plates were grown for outgrowth at 37°C or 30°C with agitation at 600 rpm for 1 h before adding aTc inducer followed by continuous incubation and reading. Absorbance at 600 nm, GFP fluorescence (F:482-16/F:530-40), RFP fluorescence (F:540-20/F:590-20), and BFP measurement (398-8/455-10) were measured continuously every 15 min.

6.3.2 Flow cytometer measurement

During this project, two different machines (i.e. BD Fortessa and Attune NxT) were used for flow cytometer. The flow cytometer settings for the BD Fortessa were: 650 volts for forward scattered (FSC), 225 volts for side scattered (SSC), 520 volts for Alexa Fluor 488, 500 volts for mCherry and threshold 1200 for FSC and 200 SSC. The flow cytometer settings for the Attune NxT were: 500 volts for forward scattered (FSC), 420 volts for side scattered (SSC), 550 volts for yellow light (YL), and a threshold of 0.3x1000. In total 10,000 events were collected for each sample. The data was stored as an FCS file 3.0 and analysis was done using FlowJo V10.

6.3.3 mARi assay

Time-series measurements were performed in 96 well plates with biological triplicates for each construct. *E. coli* carrying an empty backbone (containing only an origin of replication and an antibiotic resistance gene) was used as a negative control. *E. coli* colonies were inoculated from glycerol stock into LB media supplemented with appropriate antibiotics. The colonies were grown for 13 h at 30°C and 600 rpm in a benchtop shaker (Mikura shaker). The culture was then diluted 200x (10 µl into 190 µl then 10 µl into 90 µl) in LB media using an automated liquid handling robot (CyBio Felix). For the constitutive gene expression, the plate was grown at 37°C and 600 rpm in a benchtop shaker (Mikura shaker). After 6 h of incubation, the Abs600 was measured and checked. Abs600 typically reached values in the range 0.6–0.7. Two microliters of liquid culture were taken and diluted into 200 µl Phosphate Buffer Saline (PBS) supplemented with 2 mg/ml Kanamycin. The fluorescence data of single cells was collected using the BD Fortessa flow cytometer.

6.3.4 Measurement of P_{htpG1} assay

For the inducible expression of the recombinant protein assay, plates were grown for outgrowth at 37 or 30°C with agitation at 600 rpm for 1 h before adding aTc inducer followed by continuous incubation and reading (see Section 6.3.1). Absorbance at 600 nm and GFP fluorescence (F:482-16/F:530-40) were measured continuously every 15 min.

6.3.5 Measurement of VioB-mCherry production

For the inducible expression of the recombinant protein assay, plates were grown for outgrowth at 37°C or 30°C with agitation at 600 rpm for 1 h before adding aTc inducer followed by continuous incubation and reading (see Section 6.3.1). Absorbance at 600 nm, and mCherry fluorescence (F:540-20/F:590-20) were measured continuously every 15 min.

6.3.6 Validation of VioB-mCherry in a single cell

Single colonies of BL21(DE3) cell containing VioB-mCherry expression plasmids with and without mARi were inoculated into 100 µl of LB media supplemented with Chloramphenicol. The BL21(DE3) strains were grown for 13 h at 30°C and 600 rpm in a benchtop shaker (Mikura shaker). The culture was then diluted 200x (10 µl into 190 µl then 10 µl into 90 µl) in LB media using an automated liquid handling robot (CyBio Felix). Plates were grown for outgrowth at

37°C or 30°C with agitation at 600 rpm for 1 h before adding 5 ng/ml aTc inducer. The Abs600 was measured and two microliters of liquid culture were taken at 6 h (exponential phase), 12 h (stationary phase), and 26 h (late stationary phase) post-induction. The sampled liquid culture were diluted into 200 µl Phosphate Buffer Saline (PBS) supplemented with 2 mg/ml Kanamycin. The fluorescence data of single cells was collected using the Attune NxT flow cytometer.

6.3.7 Shake-flask growth for VioB-mCherry production

Single colonies of DH10b cell containing VioB-mCherry expression plasmids with and without mARi were inoculated into 5 ml of LB media supplemented with Chloramphenicol. The cultures were then grown in the 30°C shaking incubator. The overnight culture was diluted 200x in 50 ml LB media supplemented with Chloramphenicol, in 250 ml shake flasks. After 1 h outgrowth, aTc inducer at final concentration 5 ng/ml was added. The cultures were then grown in the 30°C shaking incubator (220 rpm) for 24 h. At the end of the experiment, absorbance at 600 nm and RFP fluorescence (F:540-20/F:590-20) were then measured. 1 ml of culture was then spun down to obtain the pellet.

6.4 Computational method, data processing, and analysis

6.4.1 *In silico* assembly

The BASIC bioparts used in this work are in Benchling's inventory and access is available upon request. The sequence file (.xdna) were generated from Serial Cloner. To generate a full sequence of the assembled plasmids, *in silico* plasmid assembly were done from individual .xdna file using a custom Python script (Python 2.7) written by Dr Marko Storch.

6.4.2 Sequencing result analysis

The sequencing results were visualised and analysed using alignment feature towards their particular sequences in Benchling.

6.4.3 Computational method for designing mARi

The sequences of UTR-RBS linkers used in BASIC DNA assembly[62] were designed and validated using R2oDNA Designer[34]. The secondary structures of the modular mARi was predicted using RNAFold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) and visualized using VARNA GUI[252] (http://varna.lri.fr/index.php?lang=en&page=tutorial&css=varna). The predicted interactions of designed modular ARi and mRNA target were simulated through IntaRNA[198,199] (http://rna.informatik.unifreiburg.de/IntaRNA/Input.jsp). Free binding energy of mARi and mRNA target site was estimated using a web-based service of the RNA folding software DINAMelt[253] (http://mfold.rna.albany.edu/?q=DINAMelt/Two-state-melting) using parameters for RNA molecule at 37°C, 1 M of natrium concentration, 0 M of magnesium concentration, and strand concentration is 0.00001 M.

6.4.4 RBS strength prediction

The predicted expression strength of UTR-RBS linker was estimated using EMOPEC[33](http://emopec.biosustain.dtu.dk/) and RBS calculator v2.0[31,32] (https://salislab.net/software/forward) with the input sequence starts from Transcription Start Site (+1) into 100 bp of CDS downstream. Six base-pair space between RBS to ATG was used as a parameter in the calculation as default length of base-pair in BASIC assembly between RBS and ATG.

6.4.5 Identity scores calculation of the UTR variants

The percent identity of mARi-UTR pairs has been calculated using EMBOSS needle method (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) with default setting.

6.4.6 Off-target prediction of the designed mARi variants

The off-target effect was computed using CopraRNA[199,208,209] (http://rna.informatik.uni-freiburg.de/CopraRNA/Input.jsp) with default setting. NC_000913 (*E. coli* MG1655), NC_010473 (*E. coli* DH10b), NC_012892 and NC_012971 (*E. coli* BL21(DE3)) were used as host references.

6.4.7 Estimation of VioB-mCherry molecular weight

The molecular weight of VioB-mCherry was calculated using Protein Molecular Weight Calculator (https://www.sciencegateway.org/tools/proteinmw.htm). A full amino acid sequence of VioB-mCherry with 1244 residue sequence resulted in protein weight of 138.64 kDa.

6.4.8 Data processing for continuous microplate reader measurement

The data of continues plate reader measurement were extracted using MARS (BMG Labtech) and saved in .xlsx file. These output files (i.e. Abs600 and GFP) and plate layout (in .xlsx file) were used as inputs for a simple data processing using custom R script (Appendix J) in RStudio environment. The script is deposited to a Github repository (https://github.com/MsDwijayanti/Data-processing-Feliostar). This script has been tested using R version 3.6.1. Briefly, the normalised Abs600 was calculated by subtracting the value of Abs600 in each well by the mean of media only wells. The normalised Abs600 were fitted into logistic equation using Growthcurver package[254]. Typical growth fitting is provided in Appendix K. The fluorescence intensity was corrected by the mean of auto-fluorescence value of negative control (cells carrying backbone-only plasmid). The value Fl/Abs600 was calculated by dividing the value of corrected fluorescence by corrected Abs600. All output of the data processing e.g. Abs600_corrected, growth_parameter, GFP_corrected, and FlperAbs600 were saved in .csv file. The growth fitting curves were saved in .pdf file.

6.4.9 Data analysis for flow cytometry measurement

The data was stored as an FCS file 3.0 and analysis was done using FlowJo V10. Single cell population gating was performed after plotting FSC-H against SSC-H and histograms of channel of fluorescences. The outputs from the BD Fortessa were using Alexa Fluor and mCherry while the outputs from the Attune NxT were using BL1 and YL2. The chosen gating covered 90-98% of the total population. The fluorescence intensity of the sample was calculated by subtracting the geometric mean of the sample with auto-fluorescence from the negative control (cells carrying backbone-only plasmid).

6.4.10 Calculation of normalised fluorescence in mARi assay

The normalised fluorescence was calculated by dividing the fluorescence intensity of the sample with mARi by the fluorescence intensity of the sample without mARi expression.

6.4.11 Statistical analysis

Statistical analysis was calculated in Prism v.8.0 (GraphPad). The values were compared using a Student's t-test for unpaired comparisons and one-way ANOVA.

6.4.12 Data visualisation

Data were visualised using Prism v.8.0 (GraphPad), Tableau10.1, and Inkscape.

References

1. Li, G. W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014).
2. Darlington, A. P. S., Kim, J., Jiménez, J. I. & Bates, D. G. Design of a translation resource allocation controller to manage cellular resource limitations. *IFAC-PapersOnLine* **50**, 12653–12660 (2017).
3. Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R. & Hwa, T. Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).
4. Patange, O., Schwall, C., Jones, M., Villava, C., Griffith, D. A., Phillips, A. & Locke, J. C. W. Escherichia coli can survive stress by noisy growth modulation. *Nat. Commun.* **9**, (2018).
5. Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K. & Lewis, K. Persisters: A distinct physiological state of E. coli. *BMC Microbiol.* (2006). doi:10.1186/1471-2180-6-53
6. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. The complete genome sequence of Escherichia coli K-12. *Science* (1997). doi:10.1126/science.277.5331.1453
7. Perna, N. T., Plunkett, G., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., Pósfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E. T., Potamousis, K. D., Apodaca, J., Anantharaman, T. S., Lin, J., Yen, G., Schwartz, D. C., Welch, R. A. & Blattner, F. R. Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. *Nature* (2001). doi:10.1038/35054089
8. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. Escherichia coli physiology in Luria-Bertani broth. *J. Bacteriol.* (2007). doi:10.1128/JB.01368-07
9. Paliy, O. & Gunasekera, T. S. Growth of E. coli BL21 in minimal media with different gluconeogenic carbon sources and salt contents. *Appl. Microbiol. Biotechnol.* (2007). doi:10.1007/s00253-006-0554-8

10. Nyström, T. Stationary-Phase Physiology. *Annu. Rev. Microbiol.* (2004). doi:10.1146/annurev.micro.58.030603.123818

11. Baeshen, M. N., Al-Hejin, A. M., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., Baeshen, N. A. & Redwan, E. M. Production of biopharmaceuticals in E. Coli: Current scenario and future perspectives. *Journal of Microbiology and Biotechnology* **25**, 953–962 (2015).

12. Ferrer-Miralles, N. & Villaverde, A. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microbial Cell Factories* **12**, (2013).

13. Chen, X., Zhou, L., Tian, K., Kumar, A., Singh, S., Prior, B. A. & Wang, Z. Metabolic engineering of Escherichia coli: A sustainable industrial platform for bio-based chemical production. *Biotechnology Advances* (2013). doi:10.1016/j.biotechadv.2013.02.009

14. Huang, C. J., Lin, H. & Yang, X. Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. *Journal of Industrial Microbiology and Biotechnology* (2012). doi:10.1007/s10295-011-1082-9

15. Riglar, D. T., Giessen, T. W., Baym, M., Kerns, S. J., Niederhuber, M. J., Bronson, R. T., Kotula, J. W., Gerber, G. K., Way, J. C. & Silver, P. A. Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* (2017). doi:10.1038/nbt.3879

16. Riglar, D. T. & Silver, P. A. Engineering bacteria for diagnostic and therapeutic applications. *Nature Reviews Microbiology* (2018). doi:10.1038/nrmicro.2017.172

17. Danino, T., Prindle, A., Kwong, G. A., Skalak, M., Li, H., Allen, K., Hasty, J. & Bhatia, S. N. Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med.* **7**, 289ra84-289ra84 (2015).

18. Charrier, M., Li, D., Mann, V. R., Yun, L., Jani, S., Rad, B., Cohen, B. E., Ashby, P. D., Ryan, K. R. & Ajo-Franklin, C. M. Engineering the S-Layer of Caulobacter crescentus as a Foundation for Stable, High-Density, 2D Living Materials. *ACS Synth. Biol.* (2019). doi:10.1021/acssynbio.8b00448

19. Green, A. A., Kim, J., Ma, D., Silver, P. A., Collins, J. J. & Yin, P. Complex cellular logic computation using ribocomputing devices. *Nature* (2017). doi:10.1038/nature23271

20. Goñi-Moreno, A. & Nikel, P. I. High-performance biocomputing in synthetic biology-integrated transcriptional and metabolic circuits. *Front. Bioeng. Biotechnol.* (2019). doi:10.3389/fbioe.2019.00040

21. Peters, J. M., Vangeloff, A. D. & Landick, R. Bacterial transcription terminators: The RNA 3′-end chronicles. *Journal of Molecular Biology* (2011). doi:10.1016/j.jmb.2011.03.036
22. Rosenberg, M. & Court, D. Regulatory Sequences Involved in the Promotion and Termination of RNA Transcription. *Annu. Rev. Genet.* (1979). doi:10.1146/annurev.ge.13.120179.001535
23. Richardson, J. P. Loading Rho to terminate transcription. *Cell* (2003). doi:10.1016/S0092-8674(03)00554-3
24. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology* (2008). doi:10.1038/nbt1413
25. Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: New engineering rules for an emerging discipline. *Molecular Systems Biology* (2006). doi:10.1038/msb4100073
26. Ellis, T., Adie, T. & Baldwin, G. S. DNA assembly for synthetic biology: From parts to pathways and beyond. *Integr. Biol.* **3**, 109–118 (2011).
27. Chao, R., Mishra, S., Si, T. & Zhao, H. Engineering biological systems using automated biofoundries. *Metab. Eng.* **42**, 98–108 (2017).
28. Kelwick, R., MacDonald, J. T., Webb, A. J. & Freemont, P. Developments in the tools and methodologies of synthetic biology. *Frontiers in Bioengineering and Biotechnology* **2**, (2014).
29. Jessop-Fabre, M. M. & Sonnenschein, N. Improving reproducibility in synthetic biology. *Front. Bioeng. Biotechnol.* **7**, 1–6 (2019).
30. Chao, R., Yuan, Y. & Zhao, H. Recent advances in DNA assembly technologies. 1–9 (2018). doi:10.1111/1567-1364.12171
31. Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **42**, 2646–2659 (2014).
32. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
33. Bonde, M. T., Pedersen, M., Klausen, M. S., Jensen, S. I., Wulff, T., Harrison, S., Nielsen, A. T., Herrgård, M. J. & Sommer, M. O. A. Predictable tuning of protein expression in bacteria. *Nat. Methods* **13**, 233–236 (2016).
34. Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T. & MacDonald, J. T.

R2oDNA Designer: Computational design of biologically-neutral synthetic DNA sequences. *ACS Synth. Biol.* 140131140734008 (2014). doi:10.1021/sb4001323

35. Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M. & Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* (2011). doi:10.1002/jcc.21596

36. Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross, D., Densmore, D. & Voigt, C. A. Genetic circuit design automation. *Science (80-.).* **352**, (2016).

37. Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K. A., Currin, A., Rattray, N. J. W., Taylor, S., Spiess, R., Sung, R., Williams, A. R., Fellows, D., Stanford, N. J., Mulherin, P., Le Feuvre, R., Barran, P., Goodacre, R., Turner, N. J., Goble, C., Chen, G. G., Kell, D. B., Micklefield, J., Breitling, R., Takano, E., Faulon, J. & Scrutton, N. S. An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.* **1**, 1–10 (2018).

38. Exley, K., Reynolds, C. R., Suckling, L., Chee, S. M., Tsipa, A., Freemont, P. S., McClymont, D. & Kitney, R. I. Utilising datasheets for the informed automated design and build of a synthetic metabolic pathway. *J. Biol. Eng.* **13**, (2019).

39. King, R. D., Whelan, K. E., Jones, F. M., Reiser, P. G. K., Bryant, C. H., Muggleton, S. H., Kell, D. B. & Oliver, S. G. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252 (2004).

40. Chambers, S., Kitney, R. & Freemont, P. The Foundry: The DNA synthesis and construction Foundry at Imperial College. *Biochem. Soc. Trans.* (2016). doi:10.1042/BST20160007

41. Hillson, N., Caddick, M., Cai, Y., Carrasco, J. A., Chang, M. W., Curach, N. C., Bell, D. J., Le Feuvre, R., Friedman, D. C., Fu, X., Gold, N. D., Herrgård, M. J., Holowko, M. B., Johnson, J. R., Johnson, R. A., Keasling, J. D., Kitney, R. I., Kondo, A., Liu, C., Martin, V. J. J., Menolascina, F., Ogino, C., Patron, N. J., Pavan, M., Poh, C. L., Pretorius, I. S., Rosser, S. J., Scrutton, N. S., Storch, M., Tekotte, H., Travnik, E., Vickers, C. E., Yew, W. S., Yuan, Y., Zhao, H. & Freemont, P. S. Building a global alliance of biofoundries. *Nature Communications* (2019). doi:10.1038/s41467-019-10079-2

42. Fletcher, L., Rosser, S. & Elfick, A. Exploring synthetic and systems biology at the

University of Edinburgh. *Biochemical Society Transactions* (2016). doi:10.1042/BST20160006

43. McClymont, D. W. & Freemont, P. S. With all due respect to Maholo, lab automation isn't anthropomorphic. *Nature Biotechnology* **35**, 312–314 (2017).
44. Yachie, N., Consortium, R. B. & Natsume, T. Robotic crowd biology with Maholo LabDroids. *Nature Biotechnology* (2017). doi:10.1038/nbt.3758
45. Casini, A., Storch, M., Baldwin, G. S. & Ellis, T. Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.* **16**, 568–576 (2015).
46. Rebatchouk, D., Daraselia, N. & Narita, J. O. NOMAD: A versatile strategy for in vitro DNA manipulation applied to promoter analysis and vector design. *Proc. Natl. Acad. Sci. U. S. A.* (1996). doi:10.1073/pnas.93.20.10891
47. Shetty, R. P., Endy, D. & Knight, T. F. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 1–12 (2008).
48. Yamazaki, K., de Mora, K. & Saitoh, K. BioBrick-based 'Quick Gene Assembly' in vitro. *Synth. Biol.* **2**, 1–8 (2017).
49. Xu, P., Vansiri, A., Bhan, N. & Koffas, M. A. G. EPathBrick: A synthetic biology platform for engineering metabolic pathways in E. coli. *ACS Synth. Biol.* **1**, 256–266 (2012).
50. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* (2008). doi:10.1371/journal.pone.0003647
51. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: A one-pot DNA shuffling method based on type ils restriction enzymes. *PLoS One* **4**, (2009).
52. Woodruff, L. B. A., Gorochowski, T. E., Roehner, N., Mikkelsen, S., Densmore, D., Gordon, D. B., Nicol, R. & Voigt, A. Registry in a tube : multiplexed pools of retrievable parts for genetic design space exploration. **45**, 1553–1565 (2018).
53. Casini, A., Chang, F. Y., Eluere, R., King, A. M., Young, E. M., Dudley, Q. M., Karim, A., Pratt, K., Bristol, C., Forget, A., Ghodasara, A., Warden-Rothman, R., Gan, R., Cristofaro, A., Borujeni, A. E., Ryu, M. H., Li, J., Kwon, Y. C., Wang, H., Tatsis, E., Rodriguez-Lopez, C., O'Connor, S., Medema, M. H., Fischbach, M. A., Jewett, M. C., Voigt, C. & Gordon, D. B. A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology. *J. Am. Chem. Soc.* **140**, 4302–4316 (2018).
54. Smanski, M. J., Bhatia, S., Zhao, D., Park, Y., Woodruff, L. B. A., Giannoukos, G., Ciulla, D., Busby, M., Calderon, J., Nicol, R., Gordon, D. B., Densmore, D. & Voigt, C. A.

Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* **32**, (2014).

55. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS One* (2011). doi:10.1371/journal.pone.0016765

56. Weber, E., Birkenfeld, J., Franz, J., Gritzan, U., Linden, L. & Trautwein, M. Modular Protein Expression Toolbox (MoPET), a standardized assembly system for defined expression constructs and expression optimization libraries. *PLoS One* **12**, 1–14 (2017).

57. Iverson, S. V, Haddock, T. L., Beal, J. & Densmore, D. M. CIDAR MoClo: Improved MoClo Assembly Standard and New E. coli Part Library Enable Rapid Combinatorial Design for Synthetic and Traditional Biology. (2016). doi:10.1021/acssynbio.5b00124

58. Moore, S. J., Lai, H., Kelwick, R. J. R., Chee, S. M., Bell, D. J., Polizzi, K. M. & Freemont, P. S. EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. (2016). doi:10.1021/acssynbio.6b00031

59. Lai, H. E., Moore, S., Polizzi, K. & Freemont, P. EcoFlex: A multifunctional moclo kit for E. coli synthetic biology. in *Methods in Molecular Biology* (2018). doi:10.1007/978-1-4939-7795-6_25

60. Moore, S. J., Macdonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A. & Aw, R. Rapid acquisition and model-based analysis of cell-free transcription – translation reactions from nonmodel bacteria. **115**, (2018).

61. Lin, D. & O'Callaghan, C. A. MetClo: Methylase-assisted hierarchical DNA assembly using a single type IIS restriction enzyme. *Nucleic Acids Res.* **46**, (2018).

62. Storch, M., Casini, A., Mackrow, B., Fleming, T., Trewhitt, H., Ellis, T. & Baldwin, G. S. BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology. *ACS Synth. Biol.* **4**, 781–787 (2015).

63. Storch, M., Casini, A., Mackrow, B., Ellis, T. & Baldwin, G. S. BASIC: A simple and accurate modular DNA assembly method. in *Methods in Molecular Biology* (2017). doi:10.1007/978-1-4939-6343-0_6

64. Zhang, S., Zhao, X., Tao, Y. & Lou, C. A novel approach for metabolic pathway optimization : Oligo-linker mediated assembly (OLMA) method. *J. Biol. Eng.* 1–10 (2015). doi:10.1186/s13036-015-0021-0

65. Van Hove, B., Guidi, C., De Wannemaeker, L., Maertens, J. & De Mey, M. Recursive DNA

Assembly Using Protected Oligonucleotide Duplex Assisted Cloning (PODAC). *ACS Synth. Biol.* **6**, 943–949 (2017).

66. Taylor, G. M., Mordaka, P. M. & Heap, J. T. Start-Stop Assembly : a functionally scarless DNA assembly system optimised for metabolic engineering . 1–31 (2018).
67. Sleight, S. C. & Sauro, H. M. Visualization of evolutionary stability dynamics and competitive fitness of Escherichia coli engineered with randomized multigene circuits. *ACS Synth. Biol.* (2013). doi:10.1021/sb400055h
68. Storch, M., Haines, M. C. & Baldwin, G. S. DNA-BOT: A low-cost, automated DNA assembly platform for synthetic biology. *bioRxiv* 832139 (2019). doi:10.1101/832139
69. Anindyajati, Anita Artarini, A., Riani, C. & Retnoningrum, D. S. Plasmid copy number determination by quantitative polymerase chain reaction. *Sci. Pharm.* (2016). doi:10.3797/scipharm.ISP.2015.02
70. Jahn, M., Vorpahl, C., Hübschmann, T., Harms, H. & Müller, S. Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microb. Cell Fact.* 1–12 (2016). doi:10.1186/s12934-016-0610-8
71. Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De Las Heras, A., Páez-Espino, A. D., Durante-Rodríguez, G., Kim, J., Nikel, P. I., Platero, R. & De Lorenzo, V. The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, 666–675 (2013).
72. Martínez-Garćía, E., Aparicio, T., Goñi-Moreno, A., Fraile, S. & De Lorenzo, V. SEVA 2.0: An update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res.* **43**, D1183–D1189 (2015).
73. Knight, T. Idempotent Vector Design for Standard Assembly of Biobricks. *MIT Libr.* 1–11 (2003). doi:http://hdl.handle.net/1721.1/21168
74. Kim, S. H., Cavaleiro, A. M., Rennig, M. & Nørholm, M. H. H. SEVA Linkers: A Versatile and Automatable DNA Backbone Exchange Standard for Synthetic Biology. *ACS Synth. Biol.* **5**, 1177–1181 (2016).
75. Harley, C. B. & Reynolds, R. P. Analysis of E.coli pormoter sequences. *Nucleic Acids Res.* **15**, 2343–2361 (1987).
76. Liang, S. T., Bipatnath, M., Xu, Y. C., Chen, S. L., Dennis, P., Ehrenberg, M. & Bremer, H. Activities of constitutive promoters in Escherichia coli. *J. Mol. Biol.* **292**, 19–37 (1999).

77. Geisel, N. Constitutive versus responsive gene expression strategies for growth in changing environments. *PLoS One* **6**, 23–25 (2011).

78. Holtz, W. J. & Keasling, J. D. Engineering Static and Dynamic Control of Synthetic Pathways. *Cell* **140**, 19–23 (2010).

79. Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de Mora, K., Glieberman, A. L., Monie, D. D. & Endy, D. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* (2009). doi:10.1186/1754-1611-3-4

80. Meyer, A. J., Segall-shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* **15**, (2019).

81. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR / O , the TetR / O and AraC / I 1 -I 2 regulatory elements. **25**, 1203–1210 (1997).

82. Ceroni, F., Boo, A., Furini, S., Gorochowski, T. E., Borkowski, O., Ladak, Y. N., Awan, A. R., Gilbert, C., Stan, G. B. & Ellis, T. Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393 (2018).

83. Anthony Mason, C., Dünner, J., Indra, P. & Colangelo, T. Heat-induced expression and chemically induced expression of the Escherichia coli stress protein HtpG are affected by the growth environment. *Appl. Environ. Microbiol.* **65**, 3433–3440 (1999).

84. Qoronfleh, M. W., Debouck, C. & Keller, J. Identification and characterization of novel low-temperature-inducible promoters of Escherichia coli. *J. Bacteriol.* (1992). doi:10.1128/jb.174.24.7902-7909.1992

85. Kelly, C. L., Liu, Z., Yoshihara, A., Jenkinson, S. F., Wormald, M. R., Otero, J. M., Estévez, A. M., Kato, A., Marqvorsen, M. H. S., Fleet, G. W. J., Estevez, R. J., Izumori, K. & Heap, J. T. Synthetic chemical inducers and genetic decoupling enable orthogonal control of the rhaBAD promoter. *ACS Synth. Biol.* acssynbio.6b00030 (2016). doi:10.1021/acssynbio.6b00030

86. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. *J. Bacteriol.* (1995). doi:10.1128/jb.177.14.4121-4130.1995

87. de Boer, H. A., Comstock, L. J. & Vasser, M. The tac promoter: a functional hybrid

derived from the trp and lac promoters. *Proc. Natl. Acad. Sci. U. S. A.* (1983). doi:10.1073/pnas.80.1.21

88. Skerra, A. Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli. *Gene* **151**, 131–135 (1994).
89. Stanton, B. C., Nielsen, A. A. K., Tamsir, A., Clancy, K., Peterson, T. & Voigt, C. A. orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2013).
90. Waters, L. S. & Storz, G. Regulatory RNAs in Bacteria. *Cell* **136**, 615–628 (2009).
91. Gottesman, S. & Storz, G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* **3**, 1–16 (2011).
92. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Molecular Cell* (2011). doi:10.1016/j.molcel.2011.08.022
93. Man, S., Cheng, R., Miao, C., Gong, Q., Gu, Y., Lu, X., Han, F. & Yu, W. Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res.* **39**, e50–e50 (2011).
94. Michaux, C., Verneuil, N., Hartke, A. & Giard, J. C. Physiological roles of small RNA molecules. *Microbiology (United Kingdom)* (2014). doi:10.1099/mic.0.076208-0
95. Thompson, K. M., Rhodius, V. A. & Gottesman, S. σE regulates and is regulated by a small RNA in Escherichia coli. *J. Bacteriol.* (2007). doi:10.1128/JB.00020-07
96. Semsey, S., Andersson, A. M. C., Krishna, S., Jensen, M. H., Massé, E. & Sneppen, K. Genetic regulation of fluxes: Iron homeostasis of Escherichia coli. *Nucleic Acids Res.* (2006). doi:10.1093/nar/gkl627
97. Svenningsen, S. L., Waters, C. M. & Bassler, B. L. A negative feedback loop involving small RNAs accelerates Vibrio cholerae's transition out of quorum-sensing mode. *Genes Dev.* (2008). doi:10.1101/gad.1629908
98. Kalamorz, F., Reichenbach, B., März, W., Rak, B. & Görke, B. Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in Escherichia coli. *Mol. Microbiol.* (2007). doi:10.1111/j.1365-2958.2007.05888.x
99. Chappell, J., Watters, K. E., Takahashi, M. K. & Lucks, J. B. A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. *Curr. Opin. Chem. Biol.* **28**, 47–56 (2015).
100. Mehta, P., Goyal, S. & Wingreen, N. S. A quantitative comparison of sRNA-based and

protein-based gene regulation. *Mol. Syst. Biol.* **4**, (2008).

101. Hussein, R. & Lim, H. N. Direct comparison of small RNA and transcription factor signaling. *Nucleic Acids Res.* **40**, 7269–7279 (2012).
102. Sharma, V., Yamamura, A. & Yokobayashi, Y. Engineering artificial small RNAs for Conditional gene silencing in escherichia coli. *ACS Synth. Biol.* **1**, 6–13 (2012).
103. Park, H., Yoon, Y., Suk, S., Lee, J. Y. & Lee, Y. Effects of different target sites on antisense RNA-mediated regulation of gene expression. *BMB Rep.* **47**, 619–624 (2014).
104. Ghodasara, A. & Voigt, C. A. Balancing gene expression without library construction via a reusable sRNA pool. *Nucleic Acids Res.* **45**, 8116–8127 (2017).
105. Vogel, J. & Luisi, B. F. Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9**, 578–589 (2011).
106. Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H. & Lee, S. Y. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nat. Biotechnol.* **31**, 170–174 (2013).
107. Yoo, S. M., Na, D. & Lee, S. Y. Design and use of synthetic regulatory small RNAs to control gene expression in Escherichia coli. *Nat. Protoc.* **8**, 1694–707 (2013).
108. Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C. D. & Vogel, J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.* **16**, 840–846 (2009).
109. Rice, J. B., Balasubramanian, D. & Vanderpool, C. K. Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proc. Natl. Acad. Sci.* **109**, E2691–E2698 (2012).
110. Viegas, S. C., Silva, I. J., Saramago, M., Domingues, S. & Arraiano, C. M. Regulation of the small regulatory RNA MicA by ribonuclease III: A target-dependent pathway. *Nucleic Acids Res.* (2011). doi:10.1093/nar/gkq1239
111. Lay, N. De, Schu, D. J. & Gottesman, S. Bacterial Small RNA-based Negative Regulation : Hfq and Its. **288**, 7996–8003 (2013).
112. Ishikawa, H., Otaka, H., Maki, K., Morita, T. & Aiba, H. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3 9 poly (U) tail. 1062–1074 (2012). doi:10.1261/rna.031575.111.understand
113. Morita, T. & Aiba, H. RNase E action at a distance: Degradation of target mRNAs

mediated by an Hfq-binding small RNA in bacteria. *Genes Dev.* **25**, 294–298 (2011).

114. Morita, T., Maki, K. & Aiba, H. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* **19**, 2176–86 (2005).
115. Aiba, H. Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.* **10**, 134–139 (2007).
116. Fei, J., Singh, D., Zhang, Q., Park, S., Balasubramanian, D., Golding, I., Vanderpool, C. K. & Ha, T. RNA biochemistry. Determination of in vivo target search kinetics of regulatory noncoding RNA. *Science (80-.).* **347**, 1371–4 (2015).
117. Geissmann, T. A. & Touati, D. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J.* **23**, 396–405 (2004).
118. Van Assche, E., Van Puyvelde, S., Vanderleyden, J. & Steenackers, H. P. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front. Microbiol.* **6**, 141 (2015).
119. Wroblewska, Z. & Olejniczak, M. Hfq assists small RNAs in binding to the coding sequence of ompD mRNA and in rearranging its structure. *RNA* 979–994 (2016). doi:10.1261/rna.055251.115
120. Franze De Fernandez, M. T., Eoyang, L. & August, J. T. Factor fraction required for the synthesis of bacteriophage Qβ-RNA. *Nature* **219**, 588–290 (1968).
121. Chen, S., Zhang, A., Blyn, L. B. & Storz, G. MicC , a Second Small-RNA Regulator of Omp Protein Expression in Escherichia coli MicC , a Second Small-RNA Regulator of Omp Protein Expression in Escherichia coli. *J. Bacteriol.* **186**, 6689–6697 (2004).
122. Sittka, A., Pfeiffer, V., Tedin, K. & Vogel, J. The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol. Microbiol.* **63**, 193–217 (2007).
123. Huis in ‘t Veld, R. A. G., Kramer, G., van der Ende, A., Speijer, D. & Pannekoek, Y. The Hfq regulon of Neisseria meningitidis. *FEBS Open Bio* **7**, 777–788 (2017).
124. Puerta-Fernández, E. & Vioque, A. Hfq is required for optimal nitrate assimilation in the cyanobacterium Anabaena sp. strain PCC 7120. *J. Bacteriol.* **193**, 3546–3555 (2011).
125. Sun, T., Li, S., Song, X., Pei, G., Diao, J., Cui, J., Shi, M., Chen, L. & Zhang, W. Re-direction of carbon flux to key precursor malonyl-CoA via artificial small RNAs in photosynthetic Synechocystis sp. PCC 6803. *Biotechnol. Biofuels* **11**, 1–16 (2018).
126. Kakoschke, T. K., Kakoschke, S. C., Zeuzem, C., Bouabe, H., Adler, K., Heesemann, J. &

Rossier, O. The RNA Chaperone Hfq Is Essential for Virulence and Modulates the Expression of Four Adhesins in Yersinia enterocolitica. *Nat. Publ. Gr.* 1–17 (2016). doi:10.1038/srep29275

127. Levine, E., Zhang, Z., Kuhlman, T. & Hwa, T. Quantitative Characteristics of Gene Regulation by Small RNA. *PLoS Biol.* **5**, (2007).
128. Valentin-hansen, P. & Eriksen, M. MicroReview The bacterial Sm-like protein Hfq : a key player in RNA transactions. *MicroReview* **51**, 1525–1533 (2004).
129. Tsui, H. C. T., Feng, G. & Winkler, M. E. Negative regulation of mutS and mutH repair gene expression by the Hfq and RpoS global regulators of Escherichia coli K-12. *J. Bacteriol.* (1997). doi:10.1128/jb.179.23.7476-7487.1997
130. Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. & Ishihama, A. Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q(β). *J. Bacteriol.* **176**, 531–534 (1994).
131. Azam, T. A., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J. Bacteriol.* **181**, 6361–6370 (1999).
132. Carpousis, A. J. The RNA Degradosome of Escherichia coli : An mRNA-Degrading Machine Assembled on RNase E. (2007). doi:10.1146/annurev.micro.61.080706.093440
133. Vytvytska, O., Moll, I., Kaberdin, V. R., Gabain, A. Von & Bla, U. Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. 1109–1118 (2000).
134. Moll, I. & Bla, U. D. O. Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. 976–984 (2005). doi:10.1261/rna.2360205.crystallography
135. Wagner, E. G. H. Kill the messenger : bacterial antisense RNA promotes mRNA decay. *Nat. Publ. Gr.* **16**, 804–806 (2009).
136. Pfeiffer, V., Papenfort, K., Lucchini, S. & Hinton, J. C. D. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. **16**, (2009).
137. Noh, M., Yoo, S. M., Kim, W. J. & Lee, S. Y. Gene Expression Knockdown by Modulating Synthetic Small RNA Expression in Escherichia coli. *Cell Syst.* (2017). doi:10.1016/j.cels.2017.08.016
138. Rodrigo, G., Landrain, T. E. & Jaramillo, A. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc. Natl. Acad. Sci. U.*

S. A. **109**, 15271–6 (2012).

139. Peters, G., Coussement, P., Maertens, J., Lammertyn, J. & De Mey, M. Putting RNA to work: Translating RNA fundamentals into biotechnological engineering practice. *Biotechnol. Adv.* (2015). doi:10.1016/j.biotechadv.2015.10.011
140. Bradley, R. W., Buck, M. & Wang, B. Tools and Principles for Microbial Gene Circuit Engineering. *Journal of Molecular Biology* (2016). doi:10.1016/j.jmb.2015.10.004
141. Xiang, Y., Dalchau, N. & Wang, B. Scaling up genetic circuit design for cellular computing: advances and prospects. *Nat. Comput.* **17**, 833–853 (2018).
142. Nielsen, A. A. K., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression. *Current Opinion in Chemical Biology* **17**, 878–892 (2013).
143. Nitzan, M., Rehani, R. & Margalit, H. Integration of Bacterial Small RNAs in Regulatory Networks. *Annu. Rev. Biophys.* **46**, 131–148 (2017).
144. Kelly, C. L., Harris, A. W. K., Steel, H., Hancock, E. J., Heap, J. T. & Papachristodoulou, A. Synthetic negative feedback circuits using engineered small RNAs. *Nucleic Acids Res.* **46**, 9875–9889 (2018).
145. Huang, H. H., Qian, Y. & Del Vecchio, D. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nat. Commun.* (2018). doi:10.1038/s41467-018-07899-z
146. Lee, Y. J., Kim, S. J. & Moon, T. S. Multilevel Regulation of Bacterial Gene Expression with the Combined STAR and Antisense RNA System. *ACS Synth. Biol.* **7**, 853–865 (2018).
147. Meyer, S., Chappell, J., Sankar, S., Chew, R. & Lucks, J. B. Improving fold activation of small transcription activating RNAs (STARs) with rational RNA engineering strategies. *Biotechnol. Bioeng.* **113**, 216–225 (2016).
148. Chappell, J., Takahashi, M. K. & Lucks, J. B. Creating small transcription activating RNAs. *Nat. Chem. Biol.* (2015). doi:10.1038/nchembio.1737
149. Chappell, J., Westbrook, A., Verosloff, M. & Lucks, J. B. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nat. Commun.* (2017). doi:10.1038/s41467-017-01082-6
150. Lee, Y. J., Kim, S. J., Moon, T. S., Saberi, F., Kamali, M., Najafi, A., Yazdanparast, A., Moghaddam, M. M., Sedlyarova, N., Shamovsky, I., Bharati, B. K., Epshtein, V., Chen, J.

J., Gottesman, S., Schroeder, R., Nudler, E., Govindarajan, S., Nevo-Dinur, K., Amster-Choder, O., Chappell, J., Takahashi, M. K., Meyer, S., Loughrey, D., Watters, K. E., Lucks, J., Chen, H., Shiroguchi, K., Ge, H., Xie, X. S., Chen, S., Zhang, A., Blyn, L. B., Storz, G., Saramago, M., Bárria, C., dos Santos, R. F., Silva, I. J., Pobre, V., Domingues, S., Andrade, J. M., Viegas, S. C., Arraiano, C. M., Sun, T., Li, S., Song, X., Pei, G., Diao, J., Cui, J., Shi, M., Chen, L., Zhang, W., Search, H., Journals, C., Contact, A., Iopscience, M., Address, I. P., Teimouri, A. M., Manuscript, T. A., Publishing, I. O. P., By, C. C., Manuscript, A., By, C. C., Manuscript, A., By, C. C., Osterman, I. A., Evfratov, S. A., Sergiev, P. V., Dontsova, O. A., Kery, M. B., Feldman, M., Livny, J., Tjaden, B., Dendooven, T., Luisi, B. F., Vecchio, D. Del, Dy, A. J., Qian, Y., Aj, D., Qian, Y., Design, L., Sprinzak, D., Elowitz, M. B., Noman, N., Inniss, M., Iba, H., Way, J. C., Torella, J. P., Lienert, F., Boehm, C. R., Chen, J. J., Way, J. C., Silver, P. A., Ceroni, F., Algar, R., Stan, G., Ellis, T., Purcell, O., Lu, T. K., Jang, S. S., Jang, S. S., Yang, J., Seo, S. W. & Jung, G. Y. Comparison of mRNA features affecting translation initiation and reinitiation. *Nucleic Acids Res.* **11**, 1379–1395 (2016).

151. Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F. & Marraffini, L. A. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* (2013). doi:10.1093/nar/gkt520

152. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. & Lim, W. A. Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. *Cell* (2013). doi:10.1016/j.cell.2013.02.022

153. Borkowski, O., Ceroni, F., Stan, G. B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* **33**, 123–130 (2016).

154. Ceroni, F., Boo, A., Furini, S., Gorochowski, T. E., Borkowski, O., Ladak, Y. N., Awan, A. R., Gilbert, C., Stan, G. B. & Ellis, T. Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393 (2018).

155. Dragosits, M., Nicklas, D. & Tagkopoulos, I. A synthetic biology approach to self-regulatory recombinant protein production in Escherichia coli. *J. Biol. Eng.* **6**, 1–10 (2012).

156. Moser, F., Espah Borujeni, A., Ghodasara, A. N., Cameron, E., Park, Y. & Voigt, C. A. Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* **14**, e8605 (2018).

157. Ceroni, F., Algar, R., Stan, G. B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* (2015). doi:10.1038/nmeth.3339

158. Gill, R. T., Valdes, J. J. & Bentley, W. E. A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in Escherichia coli. *Metab. Eng.* (2000). doi:10.1006/mben.2000.0148

159. Dürrschmid, K., Reischer, H., Schmidt-Heck, W., Hrebicek, T., Guthke, R., Rizzi, A. & Bayer, K. Monitoring of transcriptome and proteome profiles to investigate the cellular response of E. coli towards recombinant protein expression under defined chemostat conditions. *J. Biotechnol.* (2008). doi:10.1016/j.jbiotec.2008.02.013

160. Stanton, B. C., Nielsen, A. A. K., Tamsir, A., Clancy, K., Peterson, T. & Voigt, C. A. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).

161. Segall-Shapiro, T. H., Sontag, E. D. & Voigt, C. A. Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat. Biotechnol.* **36**, 352–358 (2018).

162. Wang, B., Kitney, R. I., Joly, N. & Buck, M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat. Commun.* **2**, 508–509 (2011).

163. Sleight, S. C., Bartley, B. A., Lieviant, J. A. & Sauro, H. M. Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng.* 12 (2010). doi:10.1186/1754-1611-4-12

164. Shen, P. & Huang, H. V. Homologous recombination in Escherichia coli: dependence on substrate length and homology. *Genetics* **112**, 441–457 (1986).

165. Fujitani, Y., Yamamoto, K. & Kobayashi, I. Dependence of frequency of homologous recombination on the homology length. *Genetics* **140**, 797–809 (1995).

166. Nielsen, A. A. K., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression. *Current Opinion in Chemical Biology* **17**, 878–892 (2013).

167. Lucks, J. B., Qi, L., Whitaker, W. R. & Arkin, A. P. Toward scalable parts families for predictable design of biological circuits. *Current Opinion in Microbiology* (2008). doi:10.1016/j.mib.2008.10.002

168. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131–1141 (2011).

169. Lou, C., Stanton, B., Chen, Y. J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).

170. Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K., Gao, Y., Arkin, A. P., Endy, D. & Church, G. M. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 14024–14029 (2013).

171. Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P. & Endy, D. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).

172. Carr, S. B., Beal, J. & Densmore, D. M. Reducing DNA context dependence in bacterial promoters. *PLoS One* (2017). doi:10.1371/journal.pone.0176013

173. McCall, M. J., Hendry, P. & Jennings, P. A. Minimal sequence requirements for ribozyme activity. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5710–5714 (1992).

174. Qi, L., Haurwitz, R. E., Shao, W., Doudna, J. A. & Arkin, A. P. RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.* **30**, 1002–1006 (2012).

175. Decoene, T., De Paepe, B., Maertens, J., Coussement, P., Peters, G., De Maeseneire, S. L. & De Mey, M. Standardization in synthetic biology: an engineering discipline coming of age. *Critical Reviews in Biotechnology* **38**, 647–656 (2018).

176. Chen, Y. J., Liu, P., Nielsen, A. A. K., Brophy, J. A. N., Clancy, K., Peterson, T. & Voigt, C. A. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659–664 (2013).

177. Del Vecchio, D. Modularity, context-dependence, and insulation in engineered biological circuits. *Trends in Biotechnology* **33**, 111–119 (2015).

178. Buzayan, J. M., Gerlach, W. L. & Bruening, G. Satellite tobacco ringspot virus RNA: A subset of the RNA sequence is sufficient for autolytic processing. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8859–8862 (1986).

179. Carrier, T. A. & Keasling, J. D. Engineering mRNA stability in E. coli by the addition of synthetic hairpins using a 5' cassette system. *Biotechnol. Bioeng.* **55**, 577–580 (1997).

180. Clifton, K. P., Jones, E. M., Paudel, S., Marken, J. P., Monette, C. E., Halleran, A. D., Epp, L. & Saha, M. S. The genetic insulator RiboJ increases expression of insulated genes. *J. Biol. Eng.* (2018). doi:10.1186/s13036-018-0115-6

181. Paudel, S., Marken, J. P., Monette, C. E., Jones, E. M., Epp, L., Clifton, K. P., Halleran, A. D. & Saha, M. S. The genetic insulator RiboJ increases expression of insulated genes. *J. Biol. Eng.* **12**, 1–6 (2018).

182. Smit, M. & Duin, J. Van. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7668 (1990).

183. Gorochowski, T. E., Espah Borujeni, A., Park, Y., Nielsen, A. A., Zhang, J., Der, B. S., Gordon, D. B. & Voigt, C. A. Genetic circuit characterization and debugging using RNA -seq . *Mol. Syst. Biol.* (2017). doi:10.15252/msb.20167461

184. Lin, M. T., Wang, C. Y., Xie, H. J., Cheung, C. H. Y., Hsieh, C. H., Juan, H. F., Chen, B. Sen & Lin, C. Novel Utilization of Terminators in the Design of Biologically Adjustable Synthetic Filters. *ACS Synth. Biol.* (2016). doi:10.1021/acssynbio.5b00174

185. Fang, H., Li, D., Kang, J., Jiang, P., Sun, J. & Zhang, D. Metabolic engineering of Escherichia coli for de novo biosynthesis of vitamin B12. *Nat. Commun.* (2018). doi:10.1038/s41467-018-07412-6

186. Georgi, C., Buerger, J., Hillen, W. & Berens, C. Promoter strength driving TetR determines the regulatory properties of tet-controlled expression systems. *PLoS One* **7**, (2012).

187. Subach, O. M., Cranfill, P. J., Davidson, M. W. & Verkhusha, V. V. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS One* (2011). doi:10.1371/journal.pone.0028674

188. Lahiry, A., Stimple, S. D., Wood, D. W. & Lease, R. A. Retargeting a Dual-Acting sRNA for Multiple mRNA Transcript Regulation. *ACS Synth. Biol.* **6**, 648–658 (2017).

189. Sagawa, S., Shin, J. E., Hussein, R. & Lim, H. N. Paradoxical suppression of small RNA activity at high Hfq concentrations due to random-order binding. *Nucleic Acids Res.* **43**, 8502–8515 (2015).

190. Hoynes-O'Connor, A. & Moon, T. S. Development of design rules for reliable antisense RNA behavior in E. coli. *ACS Synth. Biol.* acssynbio.6b00036 (2016). doi:10.1021/acssynbio.6b00036

191. Lee, Y. J. & Moon, T. S. Design rules of synthetic non-coding RNAs in bacteria. *Methods* **143**, 58–69 (2018).

192. Kelly, C. L., Harris, A. W. K., Steel, H., Hancock, E. J., Heap, J. T. & Papachristodoulou, A.

Synthetic negative feedback circuits using engineered small RNAs. *Nucleic Acids Res.* **46**, 9875–9889 (2018).

193. Massaiu, I., Pasotti, L., Casanova, M., Politi, N., Zucca, S., Cusella De Angelis, M. G. & Magni, P. Quantification of the gene silencing performances of rationally-designed synthetic small RNAs. *Syst. Synth. Biol.* **9**, 107–123 (2015).
194. Panja, S. & Woodson, S. A. Hfq proximity and orientation controls RNA annealing. *Nucleic Acids Res.* **40**, 8690–8697 (2012).
195. Vytvytska, O., Moll, I., Kaberdin, V. R., Von Gabain, A. & Bläsi, U. Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. *Genes Dev.* **14**, 1109–1118 (2000).
196. Večerek, B., Moll, I. & Bläsi, U. Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. *RNA* **11**, 976–984 (2005).
197. Hoynes-O'Connor, A. & Moon, T. S. Development of Design Rules for Reliable Antisense RNA Behavior in E. coli. *ACS Synth. Biol.* **5**, 1441–1454 (2016).
198. Busch, A., Richter, A. S. & Backofen, R. IntaRNA: Efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**, 2849–2856 (2008).
199. Wright, P. R., Georg, J., Mann, M., Sorescu, D. A., Richter, A. S., Lott, S., Kleinkauf, R., Hess, W. R. & Backofen, R. CopraRNA and IntaRNA: Predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res.* **42**, 119–123 (2014).
200. Kohler, R., Mooney, R. A., Mills, D. J., Landick, R. & Cramer, P. Architecture of a transcribing-translating expressome. *Science (80-.).* **356**, 194–197 (2017).
201. Ceroni, F., Furini, S., Stefan, A., Hochkoeppler, A. & Giordano, E. A synthetic post-transcriptional controller to explore the modular design of gene circuits. *ACS Synth. Biol.* **1**, 163–71 (2012).
202. Mitarai, N., Benjamin, J.-A. M., Krishna, S., Semsey, S., Csiszovszki, Z., Massé, E. & Sneppen, K. Dynamic features of gene expression control by small regulatory RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10655–10659 (2009).
203. Massaiu, I., Pasotti, L., Casanova, M., Politi, N., Zucca, S., Angelis, M. G. C. De & Magni, P. Quantification of the gene silencing performances of rationally- designed synthetic small RNAs. *Syst. Synth. Biol.* 107–123 (2015). doi:10.1007/s11693-015-9177-7
204. Teimouri, H., Korkmazhan, E., Stavans, J. & Levine, E. Sub-cellular mRNA localization

modulates the regulation of gene expression by small RNAs in bacteria. *Phys. Biol.* **14**, (2017).

205. Yoon, S. H., Han, M., Jeong, H., Lee, C. H., Xia, X., Lee, D., Shim, J. H., Lee, S. Y., Oh, T. K. & Kim, J. F. Comparative multi-omics systems analysis of Escherichia coli strains B and K-12. *Genome Biol.* **13**, R37 (2012).
206. Massé, E., Escorcia, F. E. & Gottesman, S. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev.* **17**, 2374–2383 (2003).
207. Lopez, P. J., Marchand, I., Joyce, S. A. & Dreyfus, M. The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol. Microbiol.* (1999). doi:10.1046/j.1365-2958.1999.01465.x
208. Raden, M., Ali, S. M., Alkhnbashi, O. S., Busch, A., Costa, F., Davis, J. A., Eggenhofer, F., Gelhausen, R., Georg, J., Heyne, S., Hiller, M., Kundu, K., Kleinkauf, R., Lott, S. C., Mohamed, M. M., Mattheis, A., Miladi, M., Richter, A. S., Will, S., Wolff, J., Wright, P. R. & Backofen, R. Freiburg RNA tools: A central online resource for RNA-focused research and teaching. *Nucleic Acids Res.* **46**, W25–W29 (2018).
209. Wright, P. R., Richter, A. S., Papenfort, K., Mann, M., Vogel, J., Hess, W. R., Backofen, R. & Georg, J. Comparative genomics boosts target prediction for bacterial small RNAs. *Proc. Natl. Acad. Sci.* **110**, E3487–E3496 (2013).
210. Boo, A., Ellis, T. & Stan, G. B. Host-aware synthetic biology. *Curr. Opin. Syst. Biol.* **14**, 66–72 (2019).
211. Ceroni, F., Algar, R., Stan, G. B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418 (2015).
212. Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747–747 (2014).
213. Cardinale, S., Joachimiak, M. P. & Arkin, A. P. Report Effects of Genetic Variation on the E . coli Host-Circuit Interface. *CellReports* **4**, 231–237 (2013).
214. Seizer, G., Som, T., Wei, X., Zhang, J., Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., Mukhopadhyay, A., Anthony Mason, C., Dünner, J., Indra, P., Colangelo, T., Juers, D. H., Matthews, B. W., Huber, R. E., Goelzer, A., Fromion, V., Cardinale, S., Joachimiak, M. P., Arkin, A. P., Scott, M., Mateescu, E. M., Zhang, Z. & Hwa, T. Effects of genetic variation on the e. coli host-circuit interface. *Genome Biol. Evol.* **330**, 524–537 (2011).

215. Gorochowski, T. E., Avcilar-Kucukgoze, I., Bovenberg, R. A. L., Roubos, J. A. & Ignatova, Z. A Minimal Model of Ribosome Allocation Dynamics Captures Trade-offs in Expression between Endogenous and Synthetic Genes. *ACS Synth. Biol.* **5**, 710–720 (2016).

216. Qian, Y. & Del Vecchio, D. Mitigation of ribosome competition through distributed sRNA feedback. in *2016 IEEE 55th Conference on Decision and Control, CDC 2016* (2016). doi:10.1109/CDC.2016.7798359

217. Borkowski, O., Bricio, C., Murgiano, M., Stan, G. & Ellis, T. Cell-free prediction of protein expression costs for growing cells. (2017).

218. Schlesinger, O., Chemla, Y., Heltberg, M., Ozer, E., Marshall, R., Noireaux, V., Jensen, M. H. & Alfonta, L. Tuning of Recombinant Protein Expression in Escherichia coli by Manipulating Transcription, Translation Initiation Rates, and Incorporation of Noncanonical Amino Acids. *ACS Synth. Biol.* (2017). doi:10.1021/acssynbio.7b00019

219. Cooling, M. T., Rouilly, V., Misirli, G., Lawson, J., Yu, T., Hallinan, J. & Wipat, A. Standard virtual biological parts: A repository of modular modeling components for synthetic biology. *Bioinformatics* (2010). doi:10.1093/bioinformatics/btq063

220. Jeschek, M., Gerngross, D. & Panke, S. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat. Commun.* (2016). doi:10.1038/ncomms11163

221. Huang, H. H., Qian, Y. & Del Vecchio, D. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nat. Commun.* **9**, 1–12 (2018).

222. Becskel, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* (2000). doi:10.1038/35014651

223. Kothari, V., Sharma, S. & Padia, D. Recent research advances on Chromobacterium violaceum. *Asian Pac. J. Trop. Med.* **10**, 744–752 (2017).

224. Fernandes, L. D., Moura, A. P. S. D. & Ciandrini, L. Gene length as a regulator for ribosome recruitment and protein synthesis: Theoretical insights. *Sci. Rep.* **7**, 1–11 (2017).

225. Seizer, G. & Som, T. The Origin of Replication of Plasmid p15A and Comparative Studies on the Nucleotide Sequences around the Origin of Related Plasmids. **32**, 119–129 (1983).

226. Nikolados, E. M., Weiße, A. Y., Ceroni, F. & Oyarzún, D. A. Growth Defects and Loss-of-Function in Synthetic Gene Circuits. *ACS Synth. Biol.* **8**, 1231–1240 (2019).

227. Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y. & Barkai, N. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *Elife* **6**, 1–21 (2017).

228. Nedwell, D. B. Effect of low temperature on microbial growth: Lowered affinity for substrates limits growth at low temperature. *FEMS Microbiol. Ecol.* **30**, 101–111 (1999).

229. Al Refaii, A. & Alix, J. H. Ribosome biogenesis is temperature-dependent and delayed in Escherichia coli lacking the chaperones DnaK or DnaJ. *Mol. Microbiol.* (2009). doi:10.1111/j.1365-2958.2008.06561.x

230. Li, S. H. J., Li, Z., Park, J. O., King, C. G., Rabinowitz, J. D., Wingreen, N. S. & Gitai, Z. Escherichia coli translation strategies differ across carbon, nitrogen and phosphorus limitation conditions. *Nat. Microbiol.* (2018). doi:10.1038/s41564-018-0199-2

231. Juers, D. H., Matthews, B. W. & Huber, R. E. LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Sci.* **21**, 1792–1807 (2012).

232. Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: From modules to systems. *Nature Reviews Molecular Cell Biology* **10**, 410–422 (2009).

233. Slusarczyk, A. L., Lin, A. & Weiss, R. Foundations for the design and implementation of synthetic genetic circuits. *Nature Reviews Genetics* (2012). doi:10.1038/nrg3227

234. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nature Reviews Microbiology* (2014). doi:10.1038/nrmicro3239

235. Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q.-A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P. & Endy, D. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–60 (2013).

236. Zhang, S. & Voigt, C. A. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* **46**, 11115–11125 (2018).

237. Liu, Q., Schumacher, J., Wan, X., Lou, C. & Wang, B. Orthogonality and Burdens of Heterologous and Gate Gene Circuits in E. coli. *ACS Synth. Biol.* **7**, 553–564 (2018).

238. Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B. & Cho, B. K. High-Level dCas9 Expression Induces Abnormal Cell Morphology in Escherichia coli. *ACS Synth. Biol.* **7**, 1085–1094 (2018).

239. Sagawa, S., Shin, J., Hussein, R. & Lim, H. N. Paradoxical suppression of small RNA

activity at high Hfq concentrations due to random-order binding. 1–14 (2015). doi:10.1093/nar/gkv777

240. Guisbert, E., Rhodius, V. A., Ahuja, N., Witkin, E. & Gross, C. A. Hfq modulates the σE-mediated envelope stress response and the σ32-mediated cytoplasmic stress response in Escherichia coli. *J. Bacteriol.* **189**, 1963–1973 (2007).

241. Massé, E., Escorcia, F. E. & Gottesman, S. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. 2374–2383 (2003). doi:10.1101/gad.1127103.pressed.

242. Morita, T., Mochizuki, Y. & Aiba, H. Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc. Natl. Acad. Sci. U. S. A.* (2006). doi:10.1073/pnas.0509638103

243. Stevens, J. T. & Carothers, J. M. Designing RNA-based genetic control systems for efficient production from engineered metabolic pathways. *ACS Synth. Biol.* **4**, 107–115 (2015).

244. Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K. & Gao, Y. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. **110**, (2013).

245. Nielsen, A. A. K., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression. *Curr. Opin. Chem. Biol.* **17**, 878–892 (2013).

246. Gorochowski, T. E., Borujeni, A. E., Park, Y., Nielsen, A. A. K., Zhang, J., Der, B. S., Gordon, D. B. & Voigt, C. A. Genetic circuit characterization and debugging using RNA-seq. 1–16 (2017). doi:10.15252/msb.20167461

247. Gorochowski, T. E., Chelysheva, I., Eriksen, M., Nair, P., Pedersen, S. & Ignatova, Z. Absolute quantification of translational regulation and burden using combined sequencing approaches. *Mol. Syst. Biol.* (2019). doi:10.15252/msb.20188719

248. Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E. M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* (2007). doi:10.1038/nbt1270

249. Bakshi, S., Siryaporn, A., Goulian, M. & Weisshaar, J. C. Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. *Mol. Microbiol.* **85**, 21–38 (2012).

250. Lee, Y. J., Hoynes-O'Connor, A., Leong, M. C. & Moon, T. S. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. *Nucleic Acids Res.* **44**, 2462–2473 (2016).

251. Sambrook, J. *Molecular cloning : a laboratory manual*. (Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001).

252. Blin, G., Denise, A., Dulucq, S., Herrbach, C. & Touzet, H. Alignments of RNA structures. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **7**, 309–322 (2010).

253. Markham, N. R. & Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **33**, 577–581 (2005).

254. Sprouffske, K. & Wagner, A. Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 17–20 (2016).

Appendices

Appendix A: Permission to reuse figures

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Oct 27, 2019

This Agreement between Ms. Ari Dwijayanti ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4695360502020
License date	Oct 24, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Structural & Molecular Biology
Licensed Content Title	Kill the messenger: bacterial antisense RNA promotes mRNA decay
Licensed Content Author	E Gerhart H Wagner
Licensed Content Date	Dec 31, 1969
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Engineering modular and robust biological controllers for efficient design and easy implementation in synthetic genetic circuits
Institution name	Imperial College London
Expected presentation date	Nov 2019
Portions	Figure 1: MicC: two targets, two mechanisms.
Requestor Location	Ms. Ari Dwijayanti 2, Twickenham gardens, Greenford Twickenham gardens, Greenford London, Select UB6 0LU United Kingdom Attn: Ms. Ari Dwijayanti

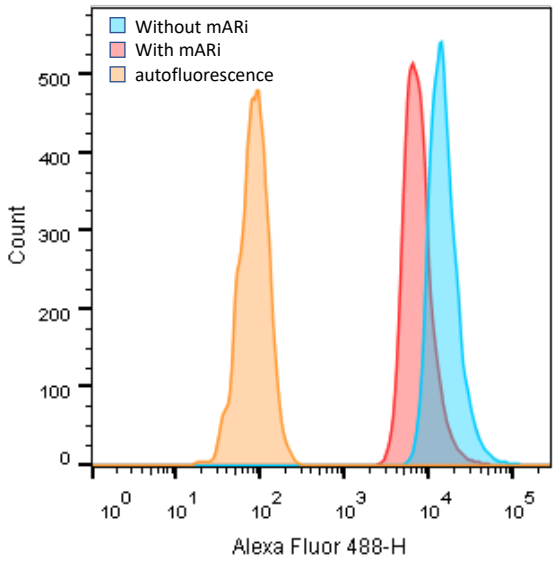
SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Oct 27, 2019

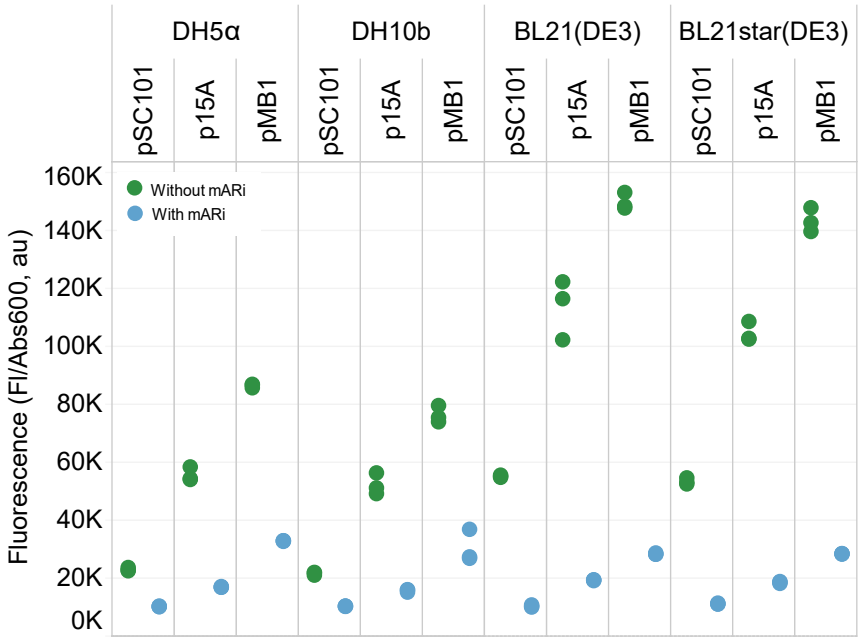
This Agreement between Ms. Ari Dwijayanti ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4695361079325
License date	Oct 24, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Structural & Molecular Biology
Licensed Content Title	Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation
Licensed Content Author	Verena Pfeiffer et al
Licensed Content Date	Jul 20, 2009
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Engineering modular and robust biological controllers for efficient design and easy implementation in synthetic genetic circuits
Institution name	Imperial College London
Expected presentation date	Nov 2019
Portions	Figure 1.b
Requestor Location	Ms. Ari Dwijayanti 2, Twickenham gardens, Greenford Twickenham gardens, Greenford London, Select UB6 0LU United Kingdom Attn: Ms. Ari Dwijayanti

Appendix B: Uniformity of producing strains by flow cytometry measurement



Appendix C: The sfGFP expression in the tested *E. coli* strains by plate reader measurement



Appendix D: List of bioparts used in this study

Part name	Type	Sequence[a] (5'→3')
$P_{J23119\_BASIC(1)}$	Standardised constitutive promoter[b]	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23111\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23104\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23101\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23108\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTCTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23106\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTTTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGG ACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23105\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGG ACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23116\_BASIC(1)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAG GACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23119\_BASIC(2)}$	Standardised constitutive promoter[b]	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGA AAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg

$P_{J23111\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACG AAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23104\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGA AAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23101\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGA AAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23108\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTCTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACG AAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23106\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGA AAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23105\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGA AAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23116\_BASIC(2)}$	Standardised constitutive promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACG AAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23119\_BASIC(3)}$	Standardised constitutive promoter[b]	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT GACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23111\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT GACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23104\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT GACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23101\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT TACAGCTAGCTCAGTCCTAGGTATTATGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg

$P_{J23108\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTCT GACAGCTAGCTCAGTCCTAGGTATAATGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23106\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT TACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23105\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT TACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23116\_BASIC(3)}$	Standardised constitutive promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT GACAGCTAGCTCAGTCCTAGGGACTATGCTAGCAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTA CCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{J23119\_noRiboJ}$	Constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCggctcgggagacctatcg
$P_{J23111\_noRiboJ}$	Constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCggctcgggagacctatcg
$P_{J23104\_noRiboJ}$	Constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCggctcgggagacctatcg
$P_{J23101\_noRiboJ}$	Constitutive promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCggctcgggagacctatcg
$P_{Tet\_BASIC(1)}$	Standardised inducible promoter[c]	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATAATG AGCACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgg gagacctatcg
$P_{Tet\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTTTTTCAGCAGGACGCACTGACCTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACAGCGCT CAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{Tet\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTTT TTCAGCAGGACGCACTGACCTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACAGTAGTCACCG GCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{Tet\_noRiboJ}$	Inducible promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATAATG AGCACggctcgggagacctatcg

$P_{Tac\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGCTCACA ATTAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggag acctatcg
$P_{Tac\_noRiboJ}$	Inducible promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGCTCACA ATTggctcgggagacctatcg
$P_{BAD\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTA ACCCCGCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCA GAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGAC GCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAA CAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{BAD\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCC CGCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAA AGTCCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTT TTATCGCAACTCTCTACTGTTTCTCCATACCCGAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGAAAGCGC CTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{BAD\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTA GAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTT ATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCC ACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATC GCAACTCTCTACTGTTTCTCCATACCCGAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTACCTCTA CAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{BAD\_noRiboJ}$	Inducible promoter	tctggtgggtctctgtccCCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGT GCCTACTCTGGAAAATCTACTTTTCATACTCCCGCCATTCAGAGAAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTG CGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGA CAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATA GCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGCTA GCggctcgggagacctatcg

$P_{Cin\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTCCCTTTGTGCGTCCAAACGGACGCACGGCGCTCTAAAGCGGGTCGCGATCTTTCAGATTCGCTCCTCGCGCTTTCAGT CTTTGTTTTGGCGCATGTCGTTATCGCAAAACCGCTGCACACTTTTGCGCGACATGCTCTGATCCCCCTCATCTGGGGGGGCCTATC TGAGGGAATTTCCGATCCGGCTCGCCTGAACCATTCTGCTTTCCACGAACTTGAAAACGCTAGCTGTCACCGGATGTGCTTTCCGG TCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{Cin\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTCCCTTTGTGCGTCCAAACGGACGCACGGCGCTCTAAAGCGGGTCGCGATCTTTCAGATTCGCTCCTCGCGCTTTCAGTCTTTG TTTTGGCGCATGTCGTTATCGCAAAACCGCTGCACACTTTTGCGCGACATGCTCTGATCCCCCTCATCTGGGGGGGCCTATCTGAG GGAATTTCCGATCCGGCTCGCCTGAACCATTCTGCTTTCCACGAACTTGAAAACGCTAGCGCTCAACGGGTGTGCTTCCCGTTCTG ATGAGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{Cin\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTCC CTTTGTGCGTCCAAACGGACGCACGGCGCTCTAAAGCGGGTCGCGATCTTTCAGATTCGCTCCTCGCGCTTTCAGTCTTTGTTTTGG CGCATGTCGTTATCGCAAAACCGCTGCACACTTTTGCGCGACATGCTCTGATCCCCCTCATCTGGGGGGGCCTATCTGAGGGAATT TCCGATCCGGCTCGCCTGAACCATTCTGCTTTCCACGAACTTGAAAACGCTAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGC CTGTGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{CymRC\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTAACAAACAGACAATCTGGTCTGTTTGTATTATGGAAAATTTTTCTGTATAATAGATTCAACAAACAGACAATCTGGTCT GTTTGTATTATAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAg gctcgggagacctatcg
$P_{CymRC\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTAACAAACAGACAATCTGGTCTGTTTGTATTATGGAAAATTTTTCTGTATAATAGATTCAACAAACAGACAATCTGGTCTGTTT GTATTATAGCGCTCAACGGGTGTGCTTCCCGTTCTGATGAGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctc gggagacctatcg
$P_{CymRC\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTAA CAAACAGACAATCTGGTCTGTTTGTATTATGGAAAATTTTTCTGTATAATAGATTCAACAAACAGACAATCTGGTCTGTTTGTATTA TAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacc tatcg
$P_{LuxB\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTACCTGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTACAGTCGAATAAAAGCTGTCACCGGATGTGCTTTCC GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{LuxB\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTACCTGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTACAGTCGAATAAAAGCGCTCAACGGGTGTGCTTCCCGTTC TGATGAGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg

$P_{LuxB\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTAC CTGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTACAGTCGAATAAAAGTAGTCACCGGCTGTGCTTGCCGGTCTGATG AGCCTGTGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{PhlF\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTCGACGTACGGTGGAATCTGATTCGTTACCAATTGACATGATACGAAACGTACCGTATCGTTAAGGTAGCTGTCACCG GATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{PhlF\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTCGACGTACGGTGGAATCTGATTCGTTACCAATTGACATGATACGAAACGTACCGTATCGTTAAGGTAGCGCTCAACGGGTG TGCTTCCCGTTCTGATGAGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{PhlF\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTCG ACGTACGGTGGAATCTGATTCGTTACCAATTGACATGATACGAAACGTACCGTATCGTTAAGGTAGTAGTCACCGGCTGTGCTTGC CGGTCTGATGAGCCTGTGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{VanCC\_BASIC(1)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCCGTGCCTACTCTG GAAAATCTATTGGATCCAATTGACAGCTAGCTCAGTCCTAGGTACCATTGGATCCAATAGCTGTCACCGGATGTGCTTTCCGGTCT GATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{VanCC\_BASIC(2)}$	Standardised inducible promoter	tctggtgggtctctgtccCTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCCGTGCCTACTCTGGAAA ATCTATTGGATCCAATTGACAGCTAGCTCAGTCCTAGGTACCATTGGATCCAATAGCGCTCAACGGGTGTGCTTCCCGTTCTGATG AGTCCGTGAGGACGAAAGCGCCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
$P_{VanCC\_BASIC(3)}$	Standardised inducible promoter	tctggtgggtctctgtccGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTGCCTACTCTGGAAAATCTAT TGGATCCAATTGACAGCTAGCTCAGTCCTAGGTACCATTGGATCCAATAGTAGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTG TGAAGGCGAAACTACCTCTACAAATAATTTTGTTTAAggctcgggagacctatcg
P_{J23105}_TetR	Repressor cassette	tctggtgggtctctgtccTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCTGAGCGCTCACAATTCTATGGACTATGTTTTCACACAG GAAAGGCCTCGATGTCCAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAATCGAAGGTTT AACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGAC GCCTTAGCCATTGAGATGTTAGATAGGCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAA CGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAGTAT GAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCA TTTTACTTTAGGTTGCGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCG CCATTATTACGACAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGA TTAGAAAAACAACTTAAATGTGAAAGTGGGTCCTAATAACTAGGGCCCATACCCTCCGGCAATTAAAAAAGCGGCTAACCACGCC GCTTTTTTTACGTCTGCAggctcgggagacctatcg
P_{J23116}_TetR	Repressor cassette	tctggtgggtctctgtccTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCTGAGCGCTCACAATTCTATGGACTATGTTTTCACACAG GAAAGGCCTCGATGTCCAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAATCGAAGGTTT

		AACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGAC GCCTTAGCCATTGAGATGTTAGATAGGCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAA CGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAGTAT GAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCA TTTTACTTTAGGTTGCGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCG CCATTATTACGACAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGA TTAGAAAAACAACTTAAATGTGAAAGTGGGTCCTAATAACTAGGGCCCATACCCTCCGGCAATTAAAAAAGCGGCTAACCACGCC GCTTTTTTTACGTCTGCAggctcgggagacctatcg
P_{LacI}-LacI	Repressor cassette	tctggtgggtctctgtccGCGGCGCGCCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCA GGGTGGTGAATATGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGC ACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCG GCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCG GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAG CTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAGGACGGTA CGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCG TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATG TCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGG GCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGATAGCT CATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCT CAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACC GCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGACAAggctcggga gacctatcg
P_{J23105}-AraC	Repressor cassette	tctggtgggtctctgtccTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCTACTAGAGAAAGAGGAGAAATACTAGATGGCTGAAGC GCAAAATGATCCCCTGCTGCCGGGATACTCGTTTAATGCCCATCTGGTGGCGGGTTTAACGCCGATTGAGGCCAACGGTTATCTCG ATTTTTTTATCGACCGACCGCTGGGAATGAAAGGTTATATTCTCAATCTCACCATTCGCGGTCAGGGGGTGGTGAAAAATCAGGG ACGAGAATTTGTTTGCCGACCGGGTGATATTTTGCTGTTCCCGCCAGGAGAGATTCATCACTACGGTCGTCATCCGGAGGCTCGCG AATGGTATCACCAGTGGGTTTACTTTCGTCCGCGCGCCTACTGGCATGAATGGCTTAACTGGCCGTCAATATTTGCCAATACGGGG TTCTTTCGCCCGGATGAAGCGCACCAGCCGCATTTCAGCGACCTGTTTGGGCAAATCATTAACGCCGGGCAAGGGGAAGGGCGCT ATTCGGAGCTGCTGGCGATAAATCTGCTTGAGCAATTGTTACTGCGGCGCATGGAAGCGATTAACGAGTCGCTCCATCCACCGAT GGATAATCGGGTACGCGAGGCTTGTCAGTACATCAGCGATCACCTGGCAGACAGCAATTTTGATATCGCCAGCGTCGCACAGCAT GTTTGCTTGTCGCCGTCGCGTCTGTCACATCTTTTCCGCCAGCAGTTAGGGATTAGCGTCTTAAGCTGGCGCGAGGACCAACGTAT CAGCCAGGCGAAGCTGCTTTTGAGCACCACCCGGATGCCTATCGCCACCGTCGGTCGCAATGTTGGTTTTGACGATCAACTCTATT

		TCTCGCGGGTATTTAAAAAATGCACCGGGGCCAGCCCGAGCGAGTTCCGTGCCGGTTAATAACCAATTATTGAAGGCCGCTAACG CAGCCTTTTTTTGTTTCTGGggctcgggagacctatcg
P_{J23101}_mARi A Position 1 (J23101_mARi A)	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTGAACACCGTCTCAGGTAAGT ATCAGTTGTAAAGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAggctcgggagacctatcg
P_{J23101}_mARi A Position 2	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTATCAGTTGTAAAAAGAGGGG AAATGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAggctcgggagacctatcg
P_{J23101}_mARi A Position 3	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTAAAAAGAGGGGAAATAGTCC ATGGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAggctcgggagacctatcg
P_{J23101}_mARi A Position 4	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAATGCGTAAAGGCGAAGAACTGT TGCTAGCATAATACCTAGGACTGAGCTAGCTGTAAAggctcgggagacctatcg
P_{J23119}_mARi A	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTGAACACCGTCTCAGGTAAGT ATCAGTTGTAAAGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAggctcgggagacctatcg
P_{J23111}_mARi A	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTGAACACCGTCTCAGGTAAGT ATCAGTTGTAAAGCTAGCACTATACCTAGGACTGAGCTAGCCGTCAAggctcgggagacctatcg
P_{J23104}_mARi A	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTGAACACCGTCTCAGGTAAGT ATCAGTTGTAAAGCTAGCACAATACCTAGGACTGAGCTAGCTGTCAAggctcgggagacctatcg
P_{J23119}_mARi B	mARi expression cassette	tctggtgggtctctgtccGGACCAAAACGAAAAAACACCCTTTCGGGTGTCTTTTCTGGAATTTGGTACCGAGAAAAAAAGCCCGGACGA CTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGGTATCTCGTGGTCTGACGGTAAA ATCTATTGTAAGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAggctcgggagacctatcg

P_{J23119}_mARi C	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTGTTACTATTGGCTGAGATAAG GGTAGCAGAAAAGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAggctcgggagacctatcg
P_{J23119}_mARi D	mARi expression cassette	tctggtgggtctctgtccGGACCAAAACGAAAAAACACCCTTTCGGGTGTCTTTTCTGGAATTTGGTACCGAGAAAAAAAGCCCGGACGA CTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTATACCCTGTCCCGGCTGTGTA TTCTATGTAAAGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAAGATTTTCCAGAGTAGGCACTTCAGGTCTATTATCACAAC GAGATGGTAGACCAGAAACAAAAAAACACCCCGAAGGGTGTTCAATAATTGGggctcgggagacctatcg
P_{J23119}_mARi E	mARi expression cassette	tctggtgggtctctgtccGGACCAAAACGAAAAAACACCCTTTCGGGTGTCTTTTCTGGAATTTGGTACCGAGAAAAAAAGCCCGGACGA CTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTATACAGCGTCTTGACGATAAA ACATATGTAAAGCTAGCATTATACCTAGGACTGAGCTAGCTGTCAAAGATTTTCCAGAGTAGGCACTTCAGGTCTATTATCACAAC GAGATGGTAGACCAGAAACAAAAAAACACCCCGAAGGGTGTTCAATAATTGGggctcgggagacctatcg
P_{htpG1}_mARi A	mARi expression cassette	tctggtgggtctctgtccTATAAACGCAGAAAGGCCCACCCGAAGGTGAGCCAGTGTGACTCTAGTAGAGAGCGTTCACCGACAAAAAC AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCTCGAGCTCGAGAAAAAAAGCCCGGACG ACTGTTCGGGCTTGTCTTTTTATATGTTGGAAAATCAGTGGCAATGCAATGGCCCAACAGAAAGTTGAACACCGTCTCAGGTAAGT ATCAGTTGTAAATGGGAGAGATGGGGACAAGGGAGAATAATTTCAAGCGAGAGCAGGTGATCCTGCTCTCGTggctcgggagacctat cg
sfGFP_no terminator	Reporter	tctggtgggtctctgtccATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGG TCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAAGTTCATCTGCACCACCGGTAAA CTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGACCACATGAAACAAC ACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGC GCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGG GCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAA GATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTG CTGCCGGATAATCACTATCTGAGCACCCAAAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAAT TTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTGTATAAGCGTCCGTAAggctcgggagacctatcg
sfGFP_B15	Reporter	tctggtgggtctctgtcc**ATG**CGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGG TCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAAGTTCATCTGCACCACCGGTAAA CTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGACCACATGAAACAAC ACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGC GCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGG GCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAA GATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTG

		CTGCCGGATAATCACTATCTGAGCACCCAAAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAAT TTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAAC GAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTT CGGGTGGGCCTTTCTGCGTTTATAggctcgggagacctatcg
sfGFP-ECK120033737_Term	Reporter	tctggtgggtctctgtcc**ATG**CGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTGGACGGCGATGTGAACGG TCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAAGTTCATCTGCACCACCGGTAAA CTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGCGCGTTATCCGGACCACATGAAACAAC ACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGC GCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGG GCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAA GATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTG CTGCCGGATAATCACTATCTGAGCACCCAAAGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAAT TTGTGACCGCTGCGGGCATCACCCACGGTATGGACGAGCTGTATAAGCGTCCGTAATAACGCTGATAGTGCTAGTGTAGATCGCT ACTAGAGGGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTTCGACCAAAGGTACTggctcgggagacctatcg
mCherry-B15_Term	Reporter	tctggtgggtctctgtcc**ATG**GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGA GGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGA AGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCA CCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGC GTGGTGACCGTGACCCAGGACTCCTCCTTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCG ACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGC GAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGT GCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAA CGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAA GGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGG TGGGCCTTTCTGCGTTTATAggctcgggagacctatcg
mCherry-B14_Term	Reporter	tctggtgggtctctgtcc**ATG**GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGA GGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGA AGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCA CCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGC GTGGTGACCGTGACCCAGGACTCCTCCTTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCG ACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGC GAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGT GCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAA

		CGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAAGGCTCGATCACGGCACTACACTCGTTGCTTTATC GGTATTGTTATTACAGAGTCCTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATATACTAGAGAGAGAATATAAAAAGC CAGATTATTAATCCGGCTTTTTTATTATTTggctcgggagacctatcg
BFP-B15_term	Reporter	tctggtgggtctctgtcc**ATG**TCCGAGTTGATCAAAGAGAACATGCATATGAAATTATATATGGAAGGCACTGTAGATAATCATCATTTTA AATGTACGTCGGAAGGCGAAGGTAAACCATATGAAGGTACGCAGACGATGCGCATCAAGGTGGTGGAGGGCGGTCCGCTGCCA TTCGCTTTCGATATTTTAGCCACGAGCTTCCTCTACGGTTCTAAAACTTTCATCAATCACACGCAGGGTATTCCGGACTTCTTTAAAC AGTCGTTCCCGGAGGGTTTCACCTGGGAACGCGTTACCACGTATGAAGATGGTGGTGTGCTTACGGCAACGCAGGACACGAGCC TTCAGGATGGGTGTTTGATTTACAACGTGAAAATTCGTGGTGTGAACTTCACGTCTAACGGCCCGGTGATGCAGAAAAAAACACT GGGTTGGGAAGCCTTTACCGAAACCCTGTATCCGGCGGACGGTGGCCTGGAAGGCCGTAATGATATGGCCTTGAAATTAGTCGG CGGTTCACACCTGATCGCGAACGCGAAAACAACCTATCGTAGTAAAAAACCAGCCAAAAACCTGAAAATGCCGggcGTCTACTACG TAGACTACCGTCTGGAGCGCattAAAGAGGCGAATAATGAAACCTATGTCGAGCAGCACGAAGTTGCGGTTGCACGCTATTGCGA TCTGCCCAGCAAACTGGGCCACAAGCTTAATGGTAGCTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAA AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAggctcgggagacctatcg
VioB-mCherry	Expression burden and reporter[82,211]	tctggtgggtctctgtcc**ATG**AGCATTCTGGATTTCCCGCGTATCCACTTCCGTGGCTGGGCCCGTGTCAATGCGCCGACCGCGAACCGC GATCCGCACGGCCACATCGATATGGCCAGCAATACCGTGGCGATGGCGGGTGAGCCGTTCGACCTGGCACGCCATCCTACGGAG TTCCACCGTCACCTGCGCTCCCTGGGTCCGCGCTTCGGCTTGGATGGTCGTGCTGACCCGGAAGGCCCGTTCAGCCTGGCCGAGG GCTACAACGCTGCCGGTAACAACCACTTTTCGTGGGAGAGCGCAACCGTTAGCCACGTGCAATGGGATGGCGGTGAGGCGGATC GTGGTGACGGTCTGGTCGGTGCTCGTTTGGCACTGTGGGGTCACTACAATGATTATCTGCGTACCACCTTCAATCGTGCTCGTTGG GTCGACAGCGACCCGACGCGCCGTGACGCTGCACAAATCTATGCGGGCCAATTCACCATTAGCCCGGCTGGTGCCGGTCCGGGTA CGCCGTGGCTGTTTACGGCAGACATTGATGATAGCCATGGTGCACGTTGGACGCGTGGCGGCCACATTGCAGAGCGTGGCGGCC ACTTCTTGGATGAAGAGTTTGGTCTGGCACGCCTGTTTCAGTTCTCTGTGCCGAAAGATCACCCACATTTTCTGTTTCACCCGGGTC CGTTTGATTCCGAGGCCTGGCGTCGTCTGCAATTGGCTCTGGAGGATGACGACGTTCTGGGTCTGACCGTGCAATATGCGTTGTTC AATATGAGCACCCCGCCTCAGCCGAACAGCCCGGTTTTTCACGATATGGTCGGTGTTGTCGGTCTGTGGCGTCGTGGTGAACTGG CGAGCTACCCGGCTGGTCGTCTGCTGCGTCCGCGTCAACCGGGTCTGGGTGACCTGACCCTGCGCGTCAACGGTGGTCGCGTTGC GCTGAATTTGGCGTGTGCCATTCCGTTCAGCACTCGTGCCGCGCAGCCAAGCGCACCGGACCGCCTGACCCCGGACCTGGGTGCC AAACTGCCGCTGGGCGATCTGCTGCTGCGTGATGAGGACGGCGCACTGTTGGCACGTGTGCCGCAGGCTCTGTACCAAGACTATT GGACGAATCACGGTATTGTGGACCTGCCGCTGCTGCGCGAACCGCGTGGTAGCTTGACCCTGAGCAGCGAACTGGCGGAGTGGC GTGAGCAAGACTGGGTCACCCAAAGCGACGCGTCTAACCTGTACCTGGAGGCACCGGATCGCCGTCACGGTCGCTTTTTCCCTGA GAGCATCGCGCTGCGCAGCTACTTTCGCGGTGAAGCGCGTGCGCGTCCGGATATCCCGCATCGTATCGAGGGCATGGGCCTGGT CGGCGTCGAATCTCGTCAGGATGGCGACGCTGCGGAATGGCGTCTGACGGGTCTGCGTCCGGGTCCGGCACGCATTGTTCTGGA CGATGGTGCCGAGGCGATCCCTCTGCGTGTTCTGCCTGACGATTGGGCGCTGGATGACGCGACCGTCGAAGAAGTGGATTACGC CTTTTTGTACCGCCACGTTATGGCGTATTACGAGCTGGTGTATCCATTCATGAGCGACAAGGTGTTTTCCCTGGCTGATCGTTGCAA

		ATGTGAAACGTACGCACGTCTGATGTGGCAGATGTGTGATCCGCAGAACCGCAACAAGTCCTATTACATGCCGAGCACCCGCGAA CTGTCGGCACCGAAAGCTCGTTTGTTCTTGAAGTATCTGGCCCACGTGGAAGGCCAGGCACGCCTGCAAGCACCTCCGCCAGCGG GTCCGGCACGCATTGAATCTAAAGCCCAGTTGGCGGCAGAGCTGCGTAAAGCCGTCGACCTGGAGCTGTCTGTGATGCTGCAATA CCTGTACGCGGCGTATAGCATTCCGAACTATGCACAGGGCCAACAACGTGTTCGTGACGGTGCGTGGACCGCCGAGCAGCTGCA ACTGGCGTGCGGTAGCGGTGACCGTCGCCGTGATGGCGGTATTCGTGCAGCACTGCTGGAAATTGCTCATGAAGAAATGATTCAT TACCTGGTCGTTAACAACCTGCTGATGGCCCTGGGCGAGCCGTTCTACGCGGGTGTCCCGCTGATGGGCGAAGCGGCACGTCAG GCGTTTGGCCTGGACACCGAGTTCGCTCTGGAACCGTTTAGCGAAAGCACGCTGGCACGTTTTGTTCGTCTGGAATGGCCGCACTT TATCCCAGCACCGGGCAAATCCATCGCGGACTGCTATGCCGCCATTCGTCAGGCGTTTTTGGATCTGCCGGACTTGTTTGGTGGCG AGGCAGGTAAGCGTGGCGGTGAACACCACCTGTTCCTGAATGAGCTGACCAACCGTGCGCATCCGGGTTATCAACTGGAAGTTTT CGATCGCGACTCGGCGCTGTTTGGTATTGCATTTGTGACCGATCAGGGCGAAGGTGGCGCTCTGGACAGCCCGCACTACGAACAT AGCCATTTTCAACGTCTGCGTGAAATGAGCGCGCGTATCATGGCTCAAAGCGCACCGTTCGAACCGGCGCTGCCGGCGTTGCGTA ATCCGGTTCTGGATGAGAGCCCGGGTTGCCAACGTGTCGCAGACGGTCGTGCGCGTGCGCTGATGGCATTGTACCAAGGCGTTTA TGAGCTGATGTTTGCGATGATGGCGCAGCACTTCGCCGTGAAACCGCTGGGTAGCTTGCGTCGCAGCCGCCTGATGAACGCAGCA ATCGATCTGATGACCGGTCTGTTGCGTCCGCTGAGCTGCGCGCTGATGAACCTGCCAAGCGGCATCGCCGGTCGCACGGCCGGTC CGCCGCTGCCGGGTCCGGTTGACACCCGTAGCTATGACGACTACGCGCTGGGCTGTCGCATGCTGGCACGCCGTTGCGAGCGTCT GCTGGAGCAGGCGAGCATGCTGGAACCGGGTTGGCTGCCGGATGCGCAGATGGAGCTGCTGGATTTCTATCGTCGCCAAATGCT GGACTTGGCGTGCGGCAAACTGAGCCGCGAGGCCGCCGAAGCTGCAGCAAAGGAGGCCGCAGCTAAAGAGGCTGCAGCGAAA GCGATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGT GAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCA AGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGAC ATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACC GTGACCCAGGACTCCTCCTTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCG TAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAG CAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCC CGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAG GGCCGCCACTCCACCTAAggctcgggagacctatcg
LacZ_no stop	Expression burden	tctggtgggtctctgtcc**ATG**ACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC AGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAG GCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTAACCTATCCCATTACGGT CAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCC AGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCG TTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGC

		AGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCA GCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGT GACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCG ATGAGCGTGGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCT CTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCG GATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGT CAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACTTTAACGCCGTGCGCTGTTCGCATTATC CGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGT GCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAA TCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCT GTCGATCCTTCCCGCCCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGCGCG TGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGAT CCTTTGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTAC AGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTG ATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGA AGCAAAACACCAGCAGCAGTTTTTCCAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCG ATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAG GTAAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGA ACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCG CCGCGTCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTTAACCGC CAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCT GGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCA GGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGCATCA GGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCG AGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGG GCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACG TCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCGAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAA CATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATC GACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACC ATTACCAGTTGGTCTGGTGTCAAAAAggctcgggagacctatcg

[a] DNA sequence colours in lowercase correspond to prefix, *i*P (blue) and suffix, *i*S (orange). DNA sequence in bold indicates starting codon. [b] Standardised constitutive promoter consists of a synthetic terminator, up element, core constitutive promoter and ribozyme insulator. [c] Standardised constitutive promoter consists of a synthetic terminator, up element, core constitutive promoter and ribozyme insulator.

Appendix E: List of plasmid backbones used in this study

Backbone	Sequence[a] (5'→3')
pSC101_KanR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAAC AATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTGTAGCCGTCCGCGTCTGAACAGCAAC ATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCG GAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCG ACCATTAAACATTTTATTCGCACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGG ATAGCGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGATCGTGTGTT TCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTAACGGCTGGCCGGTGGAACAGGTGT GGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACT GATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCT GCAAAAACGTCTGTTTCAGAAATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTG GACCGCGGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCCGGCCGGCCTCAGATCCTTCCGTATTTAGCCAGT ATGTTCTCTAGTGTGGTTCGTTGTTTTTGCGTGAGCCATGAGAACGAACCATTGAGATCATACTTACTTTGCATGTCACTCAAAAATTTTGCCTCAAAA CTGGTGAGCTGAATTTTTGCAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTCCGTTATGTAGGTAGGAATCTGATGTAATGGTTGTTGGTATTTTGTCA CCATTCATTTTTATCTGGTTGTTCTCAAGTTCGGTTACGAGATCCATTTGTCTATCTAGTTCAACTTGGAAAATCAACGTATCAGTCGGGCGGCCTCGC TTATCAACCACCAATTTCATATTGCTGTAAGTGTTTAAATCTTTACTTATTGGTTTCAAAACCCATTGGTTAAGCCTTTTAAACTCATGGTAGTTATTTTC AAGCATTAACATGAACTTAAATTCATCAAGGCTAATCTCTATATTTGCCTTGTGAGTTTTCTTTTGTGTTAGTTCTTTTAATAACCACTCATAAATCCTCA TAGAGTATTTGTTTTCAAAAGACTTAACATGTTCCAGATTATATTTTATGAATTTTTTTAACTGGAAAAGATAAGGCAATATCTCTTCACTAAAAACTA ATTCTAATTTTTCGCTTGAGAACTTGGCATAGTTTGTCCACTGGAAAATCTCAAAGCCTTTAACCAAAGGATTCCTGATTTCCACAGTTCTCGTCATCA GCTCTCTGGTTGCTTTAGCTAATACACCATAAGCATTTTCCCTACTGATGTTCATCATCTGAGCGTATTGGTTATAAGTGAACGATACCGTCCGTTCTTT CCTTGTAGGGTTTTCAATCGTGGGGTTGAGTAGTGCCACACAGCATAAAATTAGCTTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATCGCTAGTT CATTTGCTTTGAAAACAACTAATTCAGACATACATCTCAATTGGTCTAGGTGATTTTAATCACTATACCAATTGAGATGGGCTAGTCAATGATAATTAC ATGTCCTTTTCCTTTGAGTTGTGGGTATCTGTAAATTCTGCTAGACCTTTGCTGGAAAACTTGTAAATTCTGCTAGACCCTCTGTAAATTCCGCTAGAC CTTTGTGTGTTTTTTTTGTTTATATTCAAGTGGTTATAATTTATAGAATAAAGAAAGAATAAAAAAAGATAAAAAGAATAGATCCCAGCCCTGTGTAT AACTCACTACTTTAGTCAGTTCCGCAGTATTACAAAAGGATGTCGCAAACGCTGTTTGCTCCTCTACAAAACAGACCTTAAAACCCTAAAGGCTTAAG TAGCACCCTCGCAAGCTCGGGCAAATCGCTGAATATTCCTTTTGTCTCCGACCATCAGGCACCTGAGTCGCTGTCTTTTTCGTGACATTCAGTTCGCTG CGCTCACGGCTCTGGCAGTGAATGGGGGTAAATGGCACTACAGGCGCGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGC GTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG

p15A_KanR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAAC AATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTGTAGCCGTCCGCGTCTGAACAGCAAC ATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCG GAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCG ACCATTAAACATTTTATTCGCACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGG ATAGCGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGATCGTGTGTT TCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTAACGGCTGGCCGGTGGAACAGGTGT GGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACT GATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCT GCAAAAACGTCTGTTTCAGAAATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTG GACCGCGGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCCGGCCGGCCCTAGAAATATTTTATCTGATTAATA AGATGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGC TACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCCTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGACTA ACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC GGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTACCCGGAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGG CCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGT CCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTTGCCGCGGCCCTC TCACTTCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTAAGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGC GAGTCAGTGAGCGAGGAAGCGGAATATATCCGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAA CAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG
p15A_CamR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAACTTTCACCATAATGAA ATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTG ATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTT TAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATG AAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACC ACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATG TTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAA TATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAAT TACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTGACTTTTGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTAC GATTAGTCCGGCCGGCCCTAGAAATATTTTATCTGATTAATAAGATGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAA AAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTG

	TCCTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCC GGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTA CCCGGAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCAC GAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCTTGTCAGGG GGGCGGAGCCTATGGAAAAACGGCTTTTGCCGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTA AGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCGGCGCGCCCAGCTGTCTAGGGCGG CGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTT AATTAAGGCTCG
pUC_AmpR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTAGTAGCCCGCCTAATGAGCGGGCTTTTTTTTAATTCCCCTATTTGTTTATTTTTCTAA ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGCATTCAGCATTTTCGTGT GGCGCTGATTCCGTTTTTTGCGGCGTTTTGCCTGCCGGTGTTTGCGCATCCGGAAACCCTGGTGAAAGTGAAAGATGCGGAAGATCAACTGGGTGC GCGCGTGGGCTATATTGAACTGGATCTGAACAGCGGCAAAATTCTGGAATCTTTTCGTCCGGAAGAACGTTTTCCGATGATGAGCACCTTTAAAGTG CTGCTGTGCGGTGCGGTTCTGAGCCGTGTGGATGCGGGCCAGGAACAACTGGGCCGTCGTATTCATTATAGCCAGAACGATCTGGTGGAATATAGC CCGGTGACCGAAAAACATCTGACCGATGGCATGACCGTGCGTGAACTGTGCAGCGCGGCGATTACCATGAGCGATAACACCGCGGCGAACCTGCTG CTGACGACCATTGGCGGTCCGAAAGAACTGACCGCGTTTCTGCATAACATGGGCGATCATGTGACCCGTCTGGATCGTTGGGAACCGGAACTGAAC GAAGCGATTCCGAACGATGAACGTGATACCACCATGCCGGCAGCAATGGCGACCACCCTGCGTAAACTGCTGACGGGTGAGCTGCTGACCCTGGCA AGCCGCCAGCAACTGATTGATTGGATGGAAGCGGATAAAGTGGCGGGTCCGCTGCTGCGTAGCGCGCTGCCGGCTGGCTGGTTTATTGCGGATAA AAGCGGTGCGGGCGAACGTGGCAGCCGTGGCATTATTGCGGCGCTGGGCCCGGATGGTAAACCGAGCCGTATTGTGGTGATTTATACCACCGGCA GCCAGGCGACGATGGATGAACGTAACCGTCAGATTGCGGAAATTGGCGCGAGCCTGATTAAACATTGGTAAACCGATACAATTAAAGGCTCCTTTT GGAGCCTTTTTTTTTGGACGACCCTTGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCCGGCCGGCCCCGTAG AAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCG GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCA CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGG ACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA CTGAGATACCTACAGCGTGAGCTTTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTG CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG CGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCAGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGA GAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG

pMB1_KanR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAAC AATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTGTAGCCGTCCGCGTCTGAACAGCAAC ATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCG GAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCG ACCATTAAACATTTTATTCGCACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGG ATAGCGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGATCGTGTGTT TCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTAACGGCTGGCCGGTGGAACAGGTGT GGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACT GATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCT GCAAAAACGTCTGTTTCAGAAATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTG GACCGCGGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCCGGCCGGCCCCGTAGAAAAGATCAAAGGATCTT CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACT CTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC CGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGT GAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAG CTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGT GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGC GCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCC AGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCG GCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAA AGCTCATCAGCGTGGTCGTGCAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCT GATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATG AAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCG GCGGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTC GACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG
pMB1_AmpR	GTCCACTAGTCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT TGGTGAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTAGTAGCCCGCCTAATGAGCGGGCTTTTTTTTAATTCCCCTATTTGTTTATTTTTCTAA ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGCATTCAGCATTTTCGTGT GGCGCTGATTCCGTTTTTTGCGGCGTTTTGCCTGCCGGTGTTTGCGCATCCGGAAACCCTGGTGAAAGTGAAAGATGCGGAAGATCAACTGGGTGC GCGCGTGGGCTATATTGAACTGGATCTGAACAGCGGCAAAATTCTGGAATCTTTTCGTCCGGAAGAACGTTTTCCGATGATGAGCACCTTTAAAGTG

	CTGCTGTGCGGTGCGGTTCTGAGCCGTGTGGATGCGGGCCAGGAACAACTGGGCCGTCGTATTCATTATAGCCAGAACGATCTGGTGGAATATAGC CCGGTGACCGAAAAACATCTGACCGATGGCATGACCGTGCGTGAACTGTGCAGCGCGGCGATTACCATGAGCGATAACACCGCGGCGAACCTGCTG CTGACGACCATTGGCGGTCCGAAAGAACTGACCGCGTTTCTGCATAACATGGGCGATCATGTGACCCGTCTGGATCGTTGGGAACCGGAACTGAAC GAAGCGATTCCGAACGATGAACGTGATACCACCATGCCGGCAGCAATGGCGACCACCCTGCGTAAACTGCTGACGGGTGAGCTGCTGACCCTGGCA AGCCGCCAGCAACTGATTGATTGGATGGAAGCGGATAAAGTGGCGGGTCCGCTGCTGCGTAGCGCGCTGCCGGCTGGCTGGTTTATTGCGGATAA AAGCGGTGCGGGCGAACGTGGCAGCCGTGGCATTATTGCGGCGCTGGGCCCGGATGGTAAACCGAGCCGTATTGTGGTGATTTATACCACCGGCA GCCAGGCGACGATGGATGAACGTAACCGTCAGATTGCGGAAATTGGCGCGAGCCTGATTAAACATTGGTAAACCGATACAATTAAAGGCTCCTTTT GGAGCCTTTTTTTTTGGACGACCCTTGTCGGCTCGACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCCGGCCGGCCCCGTAG AAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCG GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCA CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGG ACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA CTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTG CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG CGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTG ATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA CGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG CGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGCAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAG CGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGG GGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAA CTGGCGGTATGGATGCGGCGGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAA CGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG
pSB1C3	CCAATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGATCTATCAACAGGAGTCCAAGCGAG CTCGATATCAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCATCACAAACGGCATGAT GAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCAC GTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAAAAACATATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCG TAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGA AAACGGTGTAACAAGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGAAATTCCGGATGAGCATTCATCAGGCGGGCAAG AATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATT GAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCT TAGCTCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCCGATCAACTCG

	AGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCAGAATTTCAGATAAAAAAAATCCTTA GCTTTCGCTAAGGATGATTTCTGGAATTCGCGGCCGCTTCTAGAGGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAAT CGTCCTTGACAGCTAGCTCAGTCCTAGGTATTGTGCTAGCTACTAGTGAAAGAGGAGAAATACTAGTATGGTTAGCAAAGGCGAGGCGGTTATCAA GGAGTTTATGCGTTTTAAGGTTCACATGGAGGGTAGCATGAATGGTCACGAGTTCGAGATCGAGGGTGAAGGCGAGGGTCGTCCGTACGAAGGCA CCCAGACCGCGAAGCTGAAAGTGACCAAGGGTGGCCCGCTGCCGTTCAGCTGGGACATCCTGAGCCCGCAGTTCATGTATGGCAGCCGTGCGTTTA CCAAACACCCGGCGGACATTCCGGATTACTATAAGCAAAGCTTCCCGGAAGGTTTTAAATGGGAGCGTGTTATGAACTTCGAAGATGGTGGCGCGG TGACCGTTACCCAGGACACCAGCCTGGAGGATGGCACCCTGATTTACAAGGTGAAACTGCGTGGCACCAACTTTCCGCCGGATGGTCCGGTTATGC AGAAGAAAACGATGGGTTGGGAAGCGAGCACCGAGCGTCTGTATCCGGAAGATGGCGTGCTGAAGGGTGATATCAAAATGGCGCTGCGTCTGAA GGACGGTGGCCGTTACCTGGCGGATTTTAAGACCACCTATAAAGCGAAGAAACCGGTGCAAATGCCGGGTGCGTACAACGTTGACCGTAAACTGGA TATTACCAGCCACAACGAGGATTATACCGTGGTTGAGCAATATGAGCGTAGCGAGGGTCGCCACAGCACCGGCGGCATGGACGAACTGTATAAGG GATCCTAATAACGCTGATAGTGCTAGTGTAGATCGCTACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTT ATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATATACTGGCTCGGGTAAGAACTCG CACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCTAATACTAGTAGCGGCCGCTGCAGTCCGGCAAAAAAGGGCAAGGTGTCACCACCCTGC CCTTTTTCTTTAAAACCGAAAAGATTACTTCGCGTTATGCAGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC CTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT CTGACAGCTCGAGGCTTGGATTCTCA

Appendix F: List of additional assembled plasmids used in this study

Plasmid name	Type[a]	Sequence (5'→3')	Notes
P_{J23119_RiboJ}-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTA GGTATAATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATT TTGTTTAAAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTG GACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGA AGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTG CGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTT CCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTG AAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATT ACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGC CGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGCGT GCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTA TGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTT ATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCTGTTG ATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAG AATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTG TAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCG GCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAAC GATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGCACC CCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAGCGG CGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGA TCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTA ACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGG CGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGA TCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAAATA	This study

		TGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCGCGG TCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGACCA AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT TCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACA GCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG GGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG AGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATCCA GCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGAAA TCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGTGT GACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCAAG CACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTTCG CGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAGCC AAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGCCC TCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGGGT TAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCCAT CGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGGCC AGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGTGA GGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACAAC AAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATTAC GTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACGTG GCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTATCG CCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCAGC GGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACA GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
P_{J23111_RiboJ}-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACGGCTAGCTCAGTCCTA GGTATAGTGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATT TTGTTTAAAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTG GACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGA	This study

		AGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTG CGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTT CCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTG AAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATT ACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGC CGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGCGT GCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTA TGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTT ATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCTGTTG ATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAG AATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTG TAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCG GCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAAC GATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGCACC CCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAGCGG CGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGA TCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTA ACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGG CGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGA TCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAAATA TGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCGCGG TCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGACCA AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT TCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACA GCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG GGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG AGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATCCA	

		GCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGAAA TCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGTGT GACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCAAG CACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTTCG CGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAGCC AAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGCCC TCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGGGT TAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCCAT CGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGGCC AGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGTGA GGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACAAC AAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATTAC GTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACGTG GCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTATCG CCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCAGC GGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACA GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
P_{J23104_RiboJ}-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTA GGTATTGTGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATT TTGTTTAAAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTG GACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGA AGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTG CGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTT CCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTG AAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATT ACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGC CGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGCGT GCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTA TGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTT ATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCTGTTG ATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAG	This study

		AATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTG TAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCG GCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAAC GATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGCACC CCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAGCGG CGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGA TCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTA ACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGG CGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGA TCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAAATA TGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCGCGG TCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGACCA AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT TCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACA GCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG GGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG AGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATCCA GCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGAAA TCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGTGT GACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCAAG CACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTTCG CGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAGCC AAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGCCC TCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGGGT TAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCCAT CGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGGCC AGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGTGA GGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACAAC	

		AAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATTAC GTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACGTG GCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTATCG CCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCAGC GGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACA GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
P_{J23101_RiboJ}-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTTACAGCTAGCTCAGTCCTAG GTATTATGCTAGCAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTT TGTTTAAAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCGAGCTGG ACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGACCCTGAA GTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTGTTTTGC GCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTACCATTTC CTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGAATTGA AGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTACATTA CGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAACTGGCC GACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAAGCGTG CTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCACGGTAT GGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC CTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTT ATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCTGTTG ATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAG AATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAGCTG TAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGAGCG GCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCGAAC GATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGCACC CCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAGCGG CGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAGCGA TCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAACGTA ACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCACGG CGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGCGGA TCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAAATA	This study

		TGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCGCGG TCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGACCA AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT TCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACA GCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG GGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG AGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATCCA GCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGAAA TCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGTGT GACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCAAG CACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTTCG CGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAGCC AAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGCCC TCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGGGT TAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCCAT CGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGGCC AGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGTGA GGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACAAC AAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATTAC GTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACGTG GCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTATCG CCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCAGC GGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACA GATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
$P_{J23119\_noRiboJ}$ -sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTA GGTATAATGCTAGCAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTC GAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGA CCCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGT	This study

		GTTTTGCGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTA CCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCG AATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTT ACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAA CTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAA AGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCA CGGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGAC TGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCT GTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGT GAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATAT ATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAG CTGTAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGA GCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCG AACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGC ACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAG CGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAG CGATCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAAC GTAACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCA CGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGC GGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAA ATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCG CGGTCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGA CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTT TTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACC AACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA AACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC GGAGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATC CAGCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGA	

		AATCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGT GTGACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCA AGCACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTT CGCGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAG CCAAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGC CCTCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGG GTTAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCC ATCGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGG CCAGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGT GAGGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACA ACAAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATT ACGTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACG TGGCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTAT CGCCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCA GCGGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAA CAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
$P_{J23111\_noRiboJ}$-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACGGCTAGCTCAGTCCTA GGTATAGTGCTAGCAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTC GAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGA CCCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGT GTTTTGCGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTA CCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCG AATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTT ACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAA CTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAA AGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCA CGGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGAC TGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCT GTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGT GAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATAT ATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAG	This study

		CTGTAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGA GCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCG AACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGC ACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAG CGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAG CGATCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAAC GTAACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCA CGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGC GGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAA ATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCG CGGTCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGA CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTT TTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACC AACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA AACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC GGAGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATC CAGCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGA AATCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGT GTGACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCA AGCACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTT CGCGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAG CCAAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGC CCTCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGG GTTAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCC ATCGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGG CCAGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGT GAGGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACA ACAAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATT ACGTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACG	

		TGGCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTAT CGCCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCA GCGGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAA CAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
$P_{J23104\_noRiboJ}$-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTGACAGCTAGCTCAGTCCTA GGTATTGTGCTAGCAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTC GAGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGA CCCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGT GTTTTGCGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTA CCATTTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCG AATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTT ACATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAA CTGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAA AGCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCA CGGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGAC TGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCT GTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGT GAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATAT ATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAG CTGTAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGA GCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCG AACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGC ACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAG CGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAG CGATCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAAC GTAACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCA CGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGC GGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAA ATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCG CGGTCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGA CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTT	This study

		TTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACC AACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA AACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC GGAGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATC CAGCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGA AATCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGT GTGACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCA AGCACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTT CGCGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAG CCAAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGC CCTCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGG GTTAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCC ATCGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGG CCAGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGT GAGGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACA ACAAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATT ACGTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACG TGGCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTAT CGCCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCA GCGGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAA CAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
$P_{J23101\_noRiboJ}$-sfGFP	Expression plasmid[a]	GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCCCCAATTATTGAACACCCTTCGGGGTGTTTTT TTGTTTCTGGTCTACCATCTCGTTGTGATAATAGACCTGAAGTGCCTACTCTGGAAAATCTTTTACAGCTAGCTCAGTCCTAG GTATTATGCTAGCAAAGAGGGGAAATAGTCCATGCGTAAAGGCGAAGAACTGTTCACGGGCGTAGTTCCGATTCTGGTCG AGCTGGACGGCGATGTGAACGGTCATAAGTTTAGCGTTCGCGGTGAAGGTGAGGGCGACGCGACCAACGGCAAACTGAC CCTGAAGTTCATCTGCACCACCGGTAAACTGCCGGTGCCTTGGCCGACCTTGGTGACGACGTTGACGTATGGCGTGCAGTG TTTTGCGCGTTATCCGGACCACATGAAACAACACGATTTCTTCAAATCTGCGATGCCGGAGGGTTACGTCCAGGAGCGTAC CATTTCCTTCAAGGATGATGGCACTTACAAAACTCGCGCAGAGGTTAAGTTTGAAGGTGACACGCTGGTCAATCGTATCGA ATTGAAGGGTATCGACTTTAAAGAGGATGGTAACATTCTGGGCCATAAACTGGAGTATAACTTCAACAGCCATAATGTTTA	This study

		CATTACGGCAGACAAGCAAAAGAACGGCATCAAGGCCAATTTCAAGATTCGCCACAATGTTGAGGACGGTAGCGTCCAAC TGGCCGACCATTACCAGCAGAACACCCCAATTGGTGACGGTCCGGTTTTGCTGCCGGATAATCACTATCTGAGCACCCAAA GCGTGCTGAGCAAAGATCCGAACGAAAAACGTGATCACATGGTCCTGCTGGAATTTGTGACCGCTGCGGGCATCACCCAC GGTATGGACGAGCTGTATAAGCGTCCGTAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACT GGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTG CGTTTATAGGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCCACTAGTCTTGGACTCCT GTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGT GAGAATCCAGGGGTCCCCAATAATTACGATTTAAATTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATAT ATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAGCGTGAAACGAG CTGTAGCCGTCCGCGTCTGAACAGCAACATGGATGCGGATCTGTATGGCTATAAATGGGCGCGTGATAACGTGGGTCAGA GCGGCGCGACCATTTATCGTCTGTATGGCAAACCGGATGCGCCGGAACTGTTTCTGAAACATGGCAAAGGCAGCGTGGCG AACGATGTGACCGATGAAATGGTGCGTCTGAACTGGCTGACCGAATTTATGCCGCTGCCGACCATTAAACATTTTATTCGC ACCCCGGATGATGCGTGGCTGCTGACCACCGCGATTCCGGGCAAAACCGCGTTTCAGGTGCTGGAAGAATATCCGGATAG CGGCGAAAACATTGTGGATGCGCTGGCCGTGTTTCTGCGTCGTCTGCATAGCATTCCGGTGTGCAACTGCCCGTTTAACAG CGATCGTGTGTTTCGTCTGGCCCAGGCGCAGAGCCGTATGAACAACGGCCTGGTGGATGCGAGCGATTTTGATGATGAAC GTAACGGCTGGCCGGTGGAACAGGTGTGGAAAGAAATGCATAAACTGCTGCCGTTTAGCCCGGATAGCGTGGTGACCCA CGGCGATTTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACTGATTGGCTGCATTGATGTGGGCCGTGTGGGCATTGC GGATCGTTATCAGGATCTGGCCATTCTGTGGAACTGCCTGGGCGAATTTAGCCCGAGCCTGCAAAAACGTCTGTTTCAGAA ATATGGCATTGATAATCCGGATATGAACAAACTGCAATTTCATCTGATGCTGGATGAATTTTTCTAATAATTAATTGGACCG CGGTCGGCTCGTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCCGGCCGGCCGATAATCTCATGA CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTT TTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACC AACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGCATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA AACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC GGAGCCTATGGAAAAACGCCAGCAACGCGGCCGTGAAAGGCAGGCCGGTCCGTGGTGGCCACGGCCTCTAGGCCAGATC CAGCGGCATCTGGGTTAGTCGAGCGCGGGCCGCTTCCCATGTCTCACCAGGGCGAGCCTGTTTCGCGATCTCAGCATCTGA AATCTTCCCGGCCTTGCGCTTCGCTGGGGCCTTACCCACCGCCTTGGCGGGCTTCTTCGGTCCAAAACTGAACAACAGATGT GTGACCTTGCGCCCGGTCTTTCGCTGCGCCCACTCCACCTGTAGCGGGCTGTGCTCGTTGATCTGCGTCACGGCTGGATCA AGCACTCGCAACTTGAAGTCCTTGATCGAGGGATACCGGCCTTCCAGTTGAAACCACTTTCGCAGCTGGTCAATTTCTATTT	

		CGCGCTGGCCGATGCTGTCCCATTGCATGAGCAGCTCGTAAAGCCTGATCGCGTGGGTGCTGTCCATCTTGGCCACGTCAG CCAAGGCGTATTTGGTGAACTGTTTGGTGAGTTCCGTCAGGTACGGCAGCATGTCTTTGGTGAACCTGAGTTCTACACGGC CCTCACCCTCCCGGTAGATGATTGTTTGCACCCAGCCGGTAATCATCACACTCGGTCTTTTCCCCTTGCCATTGGGCTCTTGG GTTAACCGGACTTCCCGCCGTTTCAGGCGCAGGGCCGCTTCTTTGAGCTGGTTGTAGGAAGATTCGATAGGGACACCCGCC ATCGTCGCTATGTCCTCCGCCGTCACTGAATACATCACTTCATCGGTGACAGGCTCGCTCCTCTTCACCTGGCTAATACAGG CCAGAACGATCCGCTGTTCCTGAACACTGAGGCGATACGCGGCCTCGACCAGGGCATTGCTTTTGTAAACCATTGGGGGT GAGGCCACGTTCGACATTCCTTGTGTATAAGGGGACACTGTATCTGCGTCCCACAATACAACAAATCCGTCCCTTTACAACA ACAAATCCGTCCCTTCTTAACAACAAATCCGTCCCTTAATGGCAACAAATCCGTCCCTTTTTAAACTCTACAGGCCACGGATT ACGTGGCCTGTAGACGTCCTAAAAGGTTTAAAAGGGAAAAGGAAGAAAAGGGTGGAAACGCAAAAAACGCACCACTACG TGGCCCCGTTGGGGCCGCATTTGTGCCCCTGAAGGGGCGGGGGAGGCGTCTGGGCAATCCCCGTTTTACCAGTCCCCTAT CGCCGCCTGAGAGGGCGCAGGAAGCGAGTAATCAGGGTATCGAGGCGGATTCACCCTTGGCGTCCAACCAGCGGCACCA GCGGCGCCTGAGAGGGGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAA CAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTTTAATTAAGGCTCG	
P_{htpG1}-sfGFP	Burden sensor	GGCCGCCTAGGCCACGAGAGCAGGATCACCTGCTCTCGCTTGAAATTATTCTCCCTTGTCCCCATCTCTCCCATTAATTAAG GATCTTAGCTACTAGAGAAAGAGGAGAAATACTAGATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTG GTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAATGGTAAACT GACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTTCCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCA GTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCTATGTGCAGGAACG CACGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGCA TTGAGCTGAAAGGCATTGACTTTAAAGAGGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAGCCACAAT GTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGT GCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAGCAC GCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCA CGCATGGTATGGATGAACTGTACAAATGACCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTT TTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAAGTC GGTCTCATACAGCGGCCGCGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCT AAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT GATACCGCGGGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAG TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTT	Ceroni *et al.*, 2018

		CTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGG ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGT TTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA CTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAA TAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTC CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAA CAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCA GCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTA CATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGAC GATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCG GTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCG GGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAAC GCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAAC CGTGCGGCCGC	

[a] These plasmids were assembled using BASIC assembly followed by site directed mutagenesis to excise the UTR within UTR-RBS linker sequence.

Appendix G: List of orthogonal BASIC linkers used in this study

BASIC linker	Type	Sequence[a] (5'→3')
UTR A-RBSa	UTR-RBS linker	ggctcgttgaacaccgtctcaggtaagtatcagttgtaa**aaagaggagaaa**tagtcc
UTR A-RBSb	UTR-RBS linker	ggctcgttgaacaccgtctcaggtaagtatcagttgtaa**atctaaggagg**tagtcc
UTR A-RBSc	UTR-RBS linker	ggctcgttgaacaccgtctcaggtaagtatcagttgtaa**aaagaggggaaa**tagtcc
UTR A-RBSd	UTR-RBS linker	ggctcgttgaacaccgtctcaggtaagtatcagttgtaa**atcccaggagg**tagtcc
UTR A-RBSe	UTR-RBS linker	ggctcgttgaacaccgtctcaggtaagtatcagttgtaa**atctcgggagg**tagtcc
UTR B-RBSa	UTR-RBS linker	ggctcggtatctcgtggtctgacggtaaaatctattgta**aaagaggagaaa**tagtcc
UTR B-RBSc	UTR-RBS linker	ggctcggtatctcgtggtctgacggtaaaatctattgta**aaagaggggaaa**tagtcc
UTR C-RBSa	UTR-RBS linker	ggctcgtgttactattggctgagataagggtagcagaaa**aaagaggagaaa**tagtcc
UTR C-RBSc	UTR-RBS linker	ggctcgtgttactattggctgagataagggtagcagaaa**aaagaggggaaa**tagtcc
UTR D-RBSc	UTR-RBS linker	ggctcgttataccctgtcccggctgtgtattctatgtaaa**aagaggggaaa**tagtcc
UTR E-RBSc	UTR-RBS linker	ggctcgttatacagcgtcttgacgataaaacatatgtaa**aaagaggggaaa**tagtcc
MLA	Methylated linker	ggctcggtaagaactcgcacttcgtggaaacactattatctggtgggtctctgtcc
MLB	Methylated linker	ggctcggagacctatcggtaataacagtccaatctggtgtaacttcggaatcgtcc
L2	Neutral linker	ggctcgatcggtgtgaaaagtcagtatccagtcgtgtagttcttattacctgtcc
L3	Neutral linker	ggctcgatcacggcactacactcgttgctttatcggtattgttattacagagtcc
L5	Neutral linker	ggctcgagaagtagtgccacagacagtattgcttacgagttgatttatcctgtcc
FL2	Fusion linker	ggctcgggctcgggctccggatctggttcaggttcaggatcgggctccgggtcc

[a] DNA sequence colours correspond to upstream scar (blue) and downstream scar (orange). The DNA sequence in underlined-bold indicates the RBS sequence.

Appendix H: List of bioparts and plasmids generated using PCR mutagenesis

Biopart/plasmid	Type	Template	Forward primer	Reverse primer
P_{J23119_RiboJ}-sfGFP	Expression plasmid[a]	P_{J23119_RiboJ}-UTR A-RBSc-sfGFP	P83	P84
P_{J23111_RiboJ}-sfGFP	Expression plasmid[a]	P_{J23111_RiboJ}-UTR A-RBSc -sfGFP	P83	P84
P_{J23104_RiboJ}-sfGFP	Expression plasmid[a]	P_{J23104_RiboJ}-UTR A-RBSc -sfGFP	P83	P84
P_{J23101_RiboJ}-sfGFP	Expression plasmid[a]	P_{J23101_RiboJ}-UTR A-RBSc -sfGFP	P83	P84
$P_{J23119\_noRiboJ}$-sfGFP	Expression plasmid[a]	$P_{J23119\_noRiboJ}$-UTR A-RBSc -sfGFP	P83	P48
$P_{J23111\_noRiboJ}$-sfGFP	Expression plasmid[a]	$P_{J23111\_noRiboJ}$-UTR A-RBSc -sfGFP	P83	P53
$P_{J23104\_noRiboJ}$-sfGFP	Expression plasmid[a]	$P_{J23104\_noRiboJ}$-UTR A-RBSc -sfGFP	P83	P54
$P_{J23101\_noRiboJ}$-sfGFP	Expression plasmid[a]	$P_{J23101\_noRiboJ}$-UTR A-RBSc -sfGFP	P83	P64
P_{J23105}-AraC[b]	Regulatory protein	pAN3945[36]	P35	P39
P_{J23105}-AraC	Regulatory protein	P_{J23105}-AraC[b]	P41	P40
P_{LacI}-LacI	Regulatory protein	pAN3945[36]	P37	P38
P_{J23105}-TetR	Regulatory protein	P_{J23116}-TetR	P35	P36
pSB1C3_BASIC	backbone	H2[82,211]	P175	P176
VioB-mCherry	GOI	H2[82,211]	P151	P152
LacZ	GOI	BL21(DE3) cells	P205	P206

[a] These plasmids were assembled using BASIC assembly followed by site directed mutagenesis to excise the UTR within UTR-RBS linker sequence. [b] This regulatory protein contains a BsaI site in the middle of the CDS.

Appendix I: List of primers used in this study

Primer	Phosphorylated	Direction	Sequence[a] (5'→3')	Used for
P07	No	Reverse	tgcagatgaacttcagggtcag	sfGFP sequencing confirmation
P35	No	Forward	tctggtgggtctctgtcctttacggctagctcagtcctaggtactatg	Regulatory protein extraction
P36	No	Reverse	cgataggtctcccgagccttattaggacccactttcacatttaagttgtttttctaatcc	Regulatory protein extraction
P37	No	Forward	tctggtgggtctctgtccgcggcgcgccatcg	Regulatory protein extraction
P38	No	Reverse	cgataggtctcccgagccttgtcactgcccgctttccag	Regulatory protein extraction
P39	No	Reverse	cgataggtctcccgagccccagaaacaaaaaaaggctgcgttagc	Regulatory protein extraction
P40	No	Reverse	gccgccattgggataccagaaac	Regulatory protein extraction
P41	No	Forward	gcgccatcgaatggcgcaaaa	Regulatory protein extraction
P48	Yes	Reverse	gctagcattatacctaggactgagctagctgtcaa	J23119 reverse
P53	Yes	Reverse	gctagcactatacctaggactgagctagccgtcaa	J23111 reverse
P54	Yes	Reverse	gctagcacaatacctaggactgagctagctgtcaa	J23104 reverse
P64	Yes	Reverse	gctagcataatacctaggactgagctagctgtaaa	J23101 reverse
P83	Yes	Forward	aaagaggggaaatagtccatg	UTR excision
P84	Yes	Forward	ttaaacaaaattatttgtagaggctgt	UTR excision
P151	No	Forward	tctggtgggtctctgtccatgagcattctggatttcccgcgtat	VioB extraction
P152	No	Reverse	cgataggtctcccgagccttaggtggagtggcggccctcg	VioB extraction
P167	No	Reverse	aggtgacggtggaactccgtaggat	VioB sequencing confirmation
P175	No	Reverse	tctggtgggtctctgtcctaatactagtagcggccgctgcagtccggcaaa	pSB1C3_BASIC
P176	No	Forward	cgataggtctcccgagccctctagaagcggccgcgaattccagaaatcatc	pSB1C3_BASIC
P205	No	Forward	tctggtgggtctctgtccatgaccatgattacggattcactggccgtc	LacZ extraction
P206	No	Reverse	cgataggtctcccgagccttttgacaccagaccaactggtaatggta	LacZ extraction
SP01	No	Forward	cgactcactatagggagagcggc	pJET sequencing confirmation
SP02	No	Reverse	aagaacatcgattttccatggcag	pJET sequencing confirmation
SP32	No	Forward	ggtgactgcgctcctccaagcc	Kan_outer sequencing confirmation
SP72	No	Reverse	ggacccctggattctcacc	SEVA-backbone sequencing confirmation
SP75	No	Forward	ggcggcggatttgtcctac	SEVA-backbone sequencing confirmation

Appendix J: The R script used for automated data processing

```
#A simple data processing for plate reader assay by Ari

#install packages
install.packages(c("tidyverse","readr","ggplot2","readxl",  "growthcurver",  "data.table",
"devtools", "tibbletime", "dplyr"))

#load library
library(tidyverse)
library(readr)
library(data.table)
library(devtools)
library(tibble)
library(ggplot2)
library(purrr)
library(tibbletime)
library(readxl)
library(dplyr)
library(growthcurver)

#NEED TO CHECK THIS SECTION
a<-80 #points, the number of time points measurement
d<-15 #min, the duration of each time points
type<-2 #pick the type of machine you used, 1 for Fluostar and 2 for Clariostar

#import platemap
platemap<-read_excel("platemap.xlsx")
Time= seq(0, (a-1)*d, d)
rawdata_length<-type+1+a
#Abs600 processing
```

```
Abs600_raw<-read_excel("Abs600.xlsx", col_names=F, skip=12)
Abs600_annotated<-bind_cols(Abs600_raw, platemap)
Abs600_subset<-select(Abs600_annotated, c((type+2):rawdata_length))
colnames(Abs600_subset)<-Time
Abs600_blank<-filter(Abs600_annotated, content=='medium')
Abs600_blank_mean<-select(Abs600_blank, c((type+2):rawdata_length))%>%
 summarise_all(.funs = (mean))
normalised_Abs600<-sweep(as.matrix(Abs600_subset),2,as.matrix(Abs600_blank_mean),  "-
")
Abs600_corrected<-bind_cols(platemap, as.data.frame(normalised_Abs600))
write.csv(Abs600_corrected, "Abs600_corrected.csv")

#growth parameterisation using logistic fitting by growthcurver
transformed<-as.data.frame(t(normalised_Abs600))
tTime=as.data.frame(Time)
combined<-bind_cols(tTime, transformed)
colnames(combined)<-c("time", platemap$well)
parameter_report<-SummarizeGrowthByPlate(combined)
write.csv(parameter_report,"growth_parameter.csv")
mplate<-growthdata
comb_new<-combined[,c(names(mplate))]
parameter<-SummarizeGrowthByPlate(comb_new, plot_fit = T, plot_file = "fit.pdf")

#GFP processing
GFP_raw<-read_excel("GFP.xlsx", col_names=F, skip=12)
GFP_annotated<-bind_cols(GFP_raw, platemap)
GFP_subset<-select(GFP_annotated, c((type+2):rawdata_length))
colnames(GFP_subset)<-Time
GFP_autofluorescence<-filter(GFP_annotated, content=='negative')
#GFP autofluorescence simple correction
GFP_blank_mean<-select(GFP_autofluorescence, c((type+2):rawdata_length))%>%
 summarise_all(.funs = (mean))
```

```
normalised_GFP<-sweep(as.matrix(GFP_subset),2,as.matrix(GFP_blank_mean), "-")
GFP_corrected<-bind_cols(platemap, as.data.frame(normalised_GFP))
write.csv(GFP_corrected, "GFP_corrected.csv")

#GFP/Abs600 corrected
FlperAbs600<-normalised_GFP/normalised_Abs600
colnames(FlperAbs600)<-Time
FlperAbs600<-bind_cols(platemap, as.data.frame(FlperAbs600))
write.csv(FlperAbs600, "FlperAbs600.csv")
```

Appendix K: Typical growth fitting curves