UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Engineering low-noise gene expression systems for single-molecule experiments

João Pedro Nunes Silva

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
PhD Zach Hensel
PhD Andreia Figueiredo

2018

Acknowledgments

Gostaria primeiro de agradecer ao meu orientador Zach Hensel pela paciencia para as perguntas estúpidas e pela grande ajuda durante todo o processo da tese.

Queria agradecer à Saroska por me ter afastado do caminho do vinho verde e por ser uma pessoa muito divertida que tornou a experiência no laboratório muito melhor com as suas cantorias, patas no balde e desastres naturais (como estragar-me os sapatos por me ter dado um banho de lixívia).

Queria agradecer também à Sohraia (a.k.a. bollycao) por me ter respondido sempre que lhe perguntava pela milionésima vez a mesma coisa e por me ter ensinado a andar de comboio sem pagar o bilhete mais caro.

E finalmente agradecer a alguma entidade transcendental por finalmente esta tese se encontrar escrita.

Agradeço a todas as pessoas deste lab por serem tão importantes na minha vida <3 (frase que surgiu da inspiração da Saroska)

Citação final: -Vais para o microscópio?
-Sim, mas talvez vá mais lá para as não sei.
Sohraia 2018

Resumo

A heterogeneidade observada na natureza é muitas vezes desfavorável para experiências de microbiologia ou de biologia sintética, altamente influenciadas pela plasticidade genética das populações. Para este tipo de experiências é importante minimizar a variabilidade, ou seja, diminuir o ruído entre células da mesma população. O ruído pode ser definido como a soma de fatores intrínsecos, que interferem na expressão de um gene de interesse como por exemplo a transcrição e tradução, e de fatores extrínsecos, como por exemplo as taxas de transcrição ou a quantidade de proteína expressa. Por outro lado, para experiências que envolvam expressão de proteínas é importante desenvolver um sistema que possibilite um controlo dessa mesma expressão, diminuindo o ruído. Apesar da identificação destes problemas, atualmente muitos investigadores preferem utilizar sistemas menos apropriados com altos níveis de ruído entre outros problemas facilmente solucionáveis. O objetivo principal deste trabalho é a caracterização e utilização de sistemas com baixo ruído e expressão controlada de proteínas de modo a possibilitar a criação de um sistema de dois plasmídeos capaz de expressar independentemente dois genes com níveis reduzidos de ruído.

Foram testados vários sistemas que contêm algumas características capazes de contornar estes problemas como a autorregulação negativa, transcrição bicistrónica e ainda a adição de diferentes promotores constitutivos de modo a diminuir a extensão da expressão proteica. Todos os plasmídeos foram testados em Escherichia coli utilizando microscopia e citometria de fluxo como ferramentas de caracterização do ruído e da expressão proteica. Indutores de expressão proteica como a tetraciclina e um análogo (ATc) foram testados de forma a testar os níveis de níveis de indução da expressão de proteína e o ruído. Com a tetraciclina, apesar de se esperar que o ruído induzido fosse menor uma vez que o ATc se liga mais fortemente ao promotor, resultou num maior ruído. Mais testes com concentrações de indutor menores têm de ser feitos. Para todos os sistemas com promotor PLLTetO-1 foi utilizado ATc como indutor uma vez que nas concentrações mais elevadas testadas a tetraciclina torna-se bastante tóxica para as células.

Foram também testados dois plasmídeos com número de cópias diferente (pZH509 e pJS101). Ambos os plasmídeos contêm o mesmo sistema de expressão proteica em que uma proteína de interesse e um repressor são traduzidos do mesmo mRNA, neste caso GFP e TetR foram testados. Os resultados permitiram concluir que mesmo com plasmídeos com número de cópias diferente as diferenças a nível de ruído são diminutas apesar da grande diferença nos níveis de expressão da proteína. As grandes variações ao nível da expressão proteica são muito provavelmente causadas pela diferença no número de cópias dos plasmídeos e o ruído mais elevado a elevados níveis de indução observado para pZH509 deve-se ao facto de o sistema, nesses níveis de indução, poder passar de um estado regulado para um estado não regulado. Contudo, o plasmídeo pJS101 pode substituir o plasmídeo pZH509 se forem usadas concentrações maiores de indutor.

Dois sistemas idênticos ao anterior (pJS102 e pJS103), mas com a variante de possuírem LacI enquanto repressor e com diferentes promotores entre si (PLLacO-1 e PLLacOsym respetivamente) foram também construídos e analisados. No caso do plasmídeo de pJS102 conclui-se que a curva de indução é próxima à do sistema Tet e que o ruído é também idêntico apesar de se verificar o mesmo problema acima referido de expressão proteica muito díspar. O outro sistema não funcionou como esperado uma vez que apresentou uma grande repressão do promotor levando a baixa expressão de GFP e a altos níveis de ruído. Os resultados obtidos sugerem que a energia de ligação do repressor a este promotor poderá ser mais forte impedindo que este se desassocie levando a crer que maior concentração de indutor não iria alterar os níveis de expressão. Contudo não existem modelos matemáticos que considerem dois locais de ligação de repressor, que é o caso de pJS103, e, portanto, mais estudos teriam

de ser feitos para podermos confirmar esta hipótese. Com esta experiênciaconfirma-se que sistemas de autorregulação bicistronica baseados em LacI podem substituir sistemas que utilizem TetR.

Foram testados ainda três promotores constitutivos (pJS23103, pJS23105 e pJS23106) inseridos anteriormente ao gene que codifica o repressor. Desta forma, seria expectável que uma expressão constitutiva do repressor diminuísse expressão basal de GFP. A hipótese inicial era a de que um promotor com 40% da força do promotor Tet seria ideal para que se obtivesse um ruído baixo a níveis intermédios o que não se observava para o pZH520 e ao mesmo tempo um intervalo de níveis de expressão maior que o observado para pZH509. As forças relativas dos promotores eram de 1%, 24% e 47% respetivamente, mas não foram comparados à força do promotor Tet não permitindo saber se algum corresponderia aos 40% acima referidos. Apenas pJS23103 mostrou potencial de melhoria de extensão dos níveis de expressão proteica, mas propõe-se que sejam testados mais promotores com forças compreendidas entre a do promotor pJS23103 e do promotor pJS23105.

O plasmídeo pZH713 (idêntico ao plasmídeo pJS102) apresenta a expressão da proteína PP7cp em fusão com SYFP2 (em detrimento da GFP)foi utilizado para que se obtivesse o intervalo de indução de IPTG que permite a ligação do PP7cp às repetições PP7, observável acima do ruído de fundo. Concluimos que o intervalo de indução mais adequado situa-se entre os 5 e os 40 µM. Este plasmídeo foi também utilizado para detetar moléculas de mRNA individuais com a técnica de hibridização in situ de fluorescência de moléculas individuais uma vez que esta permite deteção e quantificação de mRNAs específicos. Contudo, apresenta alguns problemas como o facto de apenas ser possível utilizá-la em células fixadas. A estirpe Escherichia coliMG1655, que contém um inserto (ISB0072) com um gene que codifica mVenus-Cro e ainda vinte e quatro repetições de PP7 foi utilizada nesta experiência. Este inserto foi testado em trabalhos anteriores e expressa mRNAs individuais. Assim, foi possível observar que a localização da proteína correspondia à localização das repetições tanto no mRNA expresso a partir do plasmídeo como do mRNA expresso a partir do cromossoma. Os dados obtidos apontam para que se tenham observado mRNAs individuais e uma boa correlação entre PP7cp e as repetições em PP7.

Baseados nos sistemas acima testados, dois plasmídeos foram criados (pZH740 e pZH742) e co transformados em Escherichia coliMG1655. O objetivo da criação deste sistema é que o mesmo permitisse exprimir controladamente dois genes, em simultâneo, em populações heterogéneas, com níveis de ruído iguais ou menores que o de genes constitutivamente expressos a partir de cromossomas (limite de ruído extrínseco) ao mesmo tempo reduzindo problemas de agregação de proteínas. O plasmídeo pZH740 permite a expressão da proteína Pf3 que possui um N-terminal que se liga à membrana e um C-terminal que fica sempre virado para o citoplasma enquanto pZH742 permite a expressão da proteína PP7. Com este sistema foi possível exprimir independentemente duas proteínas distintas. No futuro, a proteína Pf3 será usada para facilitar a contagem de proteínas uma vez que se liga à membrana celular permitindo assim comparar a expressão proteica ao número de mRNAs transcritos. O objectivo é obter a expressão de uma proteína por mRNA. Mais testes são necessários fazer ao sistema de dois plasmídeos, mas prevê-se a possibilidade de o estender para a expressão de três proteínas utilizando uma origem de replicação, um repressor e respetivo indutor e uma seleção de antibiótico diferentes.

Palavras-Chave

Autorregulação negativa, sistema bicistrónico, redução de ruído, sistema de dois plasmídeos

Abstract

Natural heterogeneity is often unfavorable in microbiological or synthetic biology experiments. Thus, lower noise in cells populations can be beneficial. For example, low noise can be useful in experiments involving protein expression. In this manuscript, we present systems capable of solving those problems. Several plasmids using bicistronic autoregulation of GFP expression were tested.

First, we found that noise was similar for compatible plasmids with different copy numbers, although they exhibited different protein expression levels.

Second, we analyzed two analogous systems with either lactose or tetracycline repressors controlling gene expression. We concluded that noise is low for both systems. A plasmid with stronger lactose repressor binding did not work well, showing that repression strength is a key parameter for this type of gene expression system.

Third, three constitutive promoters were tested to try to decrease uninduced GFP expression. One weak promoter showed potential for improving the dynamic range of protein expression. We propose that promoters with intermediate strengths should be tested.

Finally, our new plasmids were applied for imaging single mRNA molecules in living E. coli cells. The system based on the lactose repressor was used to express fluorescent proteins that bound mRNA molecules containing tandem repeats. Induction was varied to achieve single mRNA imaging, verified by an independent hybridization method for mRNA imaging. Correlation was observed between mRNAs detected by fluorescent protein binding and hybridization. An mRNA detection plasmid based on this lactose repressor system was co-transformed with another, compatible plasmid that is independently controlled by the tetracycline repressor.

With this two-plasmid system, we can now express two different genes independently with low noise. We anticipate the possibility of extending this principle to three proteins.

Keywords

Negative autoregulation, bicistronic system, noise reduction, two-plasmid system

INDEX

List of Figures

List of Tables

List of Abbreviations and Acronyms

2D - Two dimensions
ATc - Anhydrotetracycline
AU - Arbitrary units
bp - Base pairs
CV - Coefficient of variation
DNA - Deoxyribonucleic acid
FL1 - Fluorescence channel
FSC - Forward scatter
GFP - Green fluorescent protein
IPTG - Isopropyl β-D-1-thiogalactopyranoside
LacI - Lac repressor
LB media - Lysogeny broth media
mL - Milliliter
mRNA - Messenger ribonucleic acid
MS2cp - MS2 coat protein
ng - Nanograms
nm - Nanometer
nM - Nanomolar
OD600 - Optical density at 600 nm
ORF - Open reading frame
PBS - Phosphate buffered saline
PCR - Polymerase chain reaction
PP7cp - PP7 coat protein
RBS - Ribosome binding site
RNAP - RNA polymerase
smFISH - Single molecule fluorescence in situ hybridization
SOC - Super optimal broth
SSC - Side scatter
SYFP - Superfolder yellow fluorescent protein
TetR – Tet repressor
TF - Transcription factor
σ - Standard deviation
µg - Microgram
µL - Microliter
µM - micromolar
µ - GFP fluorescence mean

1. Introduction

As observed in natural environments, variation in gene expression within isogenic populations might prove beneficial and facilitate bet-hedging or division of labor strategies that permit an increased adaptation to the environment and to changes that may occur (temperature, nutrients, etc.). For example, in *Bacillus subtilis* two types of cells can be found in the same population, some that do not sporulate when nutrients become limiting and others, "sporulators", that do. Each of these pathways is a form of cell specialization and they are beneficial when occurring in the same population in circumstances of low nutrients since it works as a way of optimizing the use of resources in the long-term. This happens because when mother cells lyse to release endospores they also release cellular content that can be used by other cells as nutrients (Veening, J.-W. *et al.*, 2008).

Variations observed between same population cells can be explained by stochastic biochemical events that are intrinsic to gene expression, or arise from random separation of molecules during cell division or different cellular responses to environmental factors. However, at low expression levels, the main source of variability in these populations is usually considered to be the fluctuations observed in molecules associated with gene expression (Hooshangi, S. and Weiss, R., 2006).

This heterogeneity that occurs in isogenic populations of cells can be unfavorable in biotechnology and synthetic biology, for instance when a toxic protein is being expressed or the accumulation of a protein in a metabolic pathway can compromise cell viability. In this scenario it is useful to engineer phenotypically homogenous cells to obtain a more predictable and controllable target gene expression. There are some characteristics essential for processes where the formation of biomass interconnects with the accumulation of proteins while in an experimental environment, such as the rate of gene expression. Nevertheless, expression systems that normally display an all-or-none response to induction are widely used and have some disadvantages since they can interfere with metabolic pathways in the cell. In addition, the growth protocols and the optimizations needed for these systems are harder to obtain and take a longer time to achieve and sometimes only a small fraction of the cells have the level of proteins of interest needed for the experiments that are necessary to conduct (Binder, D. *et al.*, 2016). In order to evaluate methods to achieve less cell-to-cell variation in gene expression, expression systems need to be assessed at a single-cell level with respect to features that have important roles in the expression of proteins (Wang, Y.-H., Wei, K. Y. and Smolke, C. D., 2013). This information will be vital to acquire a better control over the expression of a gene of interest.

For microbiology experiments the need for recombinant expression of a protein with levels close to those of endogenous proteins is common and those levels may vary from 1 to 10000 molecules per cell and can be even higher for essential genes (Hensel, Z., 2017). Since recombinant expression systems used normally do not achieve low heterogeneity in same population cells, different expression levels of a protein of interest can be observed even at high levels of expression. This is not seen for most genes expressed from the chromosome (Taniguchi *et al*., 2010). Consequently, a system that minimizes these cell-to-cell variations would be useful. Such system should have low noise at low expression levels and would also be useful to improve yields in some cases where aggregation of proteins proves to be a problem making it easier to control expression and permitting cells to produce the maximum amount of protein possible without reaching levels that would trigger aggregation (Rosano, G. L. and Ceccarelli, E. A., 2014).

Noise can be defined as the sum of "intrinsic" factors that interfere in the expression of the gene of interest, such as randomness of binding, transcription, translation, and degradation, and "extrinsic" factors, for example gene dosage and rates of transcription/translation. Noise can be calculated as the standard deviation divided by the mean, in other words, the magnitude of fluctuations compared to the mean value (Paulsson, J., 2004). There are other ways to calculate noise but an advantage of this one is that it is dimensionless. In this manuscript, noise is calculated as the variation divided by the square mean as described in Taniguchi, Y. *et al*., 2010. Intrinsic noise can also be referred as the stochasticity in chemical reactions, meaning that because two molecules must find each other in the cell and overcome an energy barrier to interact, the process cannot occur continuously and deterministically, but only with a finite probability per unit time.

Noise in genetic networks can be explained by looking at DNA. DNA is the template for its own replication. Therefore, an unregulated replication in addition to first-order elimination (a constant proportion of some component is eliminated per unit time) is dynamically unstable. This means that if the average replication of every molecule is more than one per cell cycle the concentration increases until resources become scarce. On the contrary, if it is less than one, the copy number decreases until there are no more templates. In an intermediate state, where synthesis and elimination are balanced, there are still random fluctuations. This instability does not exist in case of unregulated transcription and translation when combined with first-order elimination given that mRNAs and proteins do not act as templates for their own synthesis. Influences from both "intrinsic noise" (resulting from random births and deaths of individual molecules) and "extrinsic noise" (arising from influences of other noisy components such as plasmid copy numbers) can be found in every cell component. (Paulsson, J., 2004). One way both types of noise can be reduced is with negative autoregulation (the process by which a protein regulates its own promoter resulting in a negative feedback loop repression either blocking the binding of RNAP or locking RNAP in an immobile state) (Swain, P. S., 2004 and Paulsson, J., 2004).

It was proposed that protein adjustments could possibly increase noise since they cause alterations on the mean over time but when considering proteins that have less than 10 copies per cell (low-abundance proteins) the noise decreases with higher levels of protein in the cell and for proteins with higher abundance the noise level remains stationary above a certain limit ("extrinsic" noise limit) (Tyagi, S., 2010). However, cells with more plasmid copies will accumulate more protein and this means that plasmid copy numbers can be a source of extrinsic noise in genes encoded by the plasmid. Previous works show that plasmid copy numbers can be compensated somewhat by autorepression causing a reduction of expression for each gene copy. In Paulsson, J., 2004 a model is described that explains how this process works. Nevertheless, because transcription factor lifetimes are long relative to transcription/translation rates, negative feedback primarily works to reduce relatively slowly varying sources of noise such as plasmid copy number variation. Negative autorepression is also beneficial to guarantee that plasmid replication proteins will not suffer from any disturbance in both plasmid copy numbers and the intracellular environment. It can also make gene response times quicker, minimize mRNA usage reducing the metabolic cost of protein production or even diminish ribosome variation and shift noise in gene expression to higher frequencies (Simpson, M. L. *et al.*, 2003 and Austin, D. W. *et al.*, 2006) (characteristic of intrinsic noise) poorly affecting the higher frequencies, resulting in a frequency range distribution closer to normal distributions (Paulsson, J., 2004 and Nevozhay, D. *et al*., 2009). In other words, an autoregulated gene will respond quicker to fluctuations. In prokaryotes where transcription factor decay and dilution are typically slow compared to transcription and translation autoregulation will have larger effects on slower fluctuations than on faster fluctuations. This explains why this type of autoregulation will mainly suppress extrinsic noise (Tsimring, L. S., 2014). An alternative regulatory motif to negative autoregulation is positive feedback which operates as a switch,

as a lower concentration of a given protein can activate its own gene. There are ways for a cell to maintain the gene in an inactive state such as the presence of cooperativity which establishes a limit that the protein must overcome to trigger the feedback, however, there is a possibility of random activation caused by occasional large fluctuations (Raj, A. and van Oudenaarden A., 2008). Noise can modify how information flows in a genetic network and is heavily affected by the network architecture. For instance, if a spontaneous change from inactive to active or vice-versa occurs in a given gene coding a transcriptional repressor this will subjugate the encoded protein spreading the fluctuations to the genes it represses. (Hooshangi, S. and Weiss, R., 2006).

In prokaryotes, autoregulation is common in gene regulatory networks (Shen-Orr, S. S. *et al.*, 2002) and in the last decades many inducible bacterial expression systems inspired by natural regulatory networks were created. The majority of these systems are based on catabolic regulatory networks such as those for lactose, arabinose or benzoate utilization and are used has tools for heterologous gene expression. They are normally constituted by a natural or mutagenized promoter and a transcriptional regulator that can repress, derepress or activate target gene expression when a specific inducer is added to the media. The inducers will enter the cell by passive diffusion or through an appropriate transport system (Binder, D. *et al.*, 2016).

It has been shown that negative autoregulation by a system based on the described above with the transcriptional repression made by the fusion protein TetR-EGFP can reduce gene expression noise in *Escherichia coli* (Dublanche, Y. *et al.*, 2006). The explanation proposed for this reduction was that it should result from a dosage compensation in the plasmid copy number. When the inducer anhydrotetracycline is added at very low concentration the coefficient of variation (measure of variability that is also known as noise) remains low. The proprieties of this autorepressive system can be changed when a weak repression factor is used or even by having a low concentration of the repressor. On the other hand, the autorepressive system will remain unaltered if the concentration of the inducer in the media is low. However at high inducer concentration the system can change from a state of regulation to a state of no regulation because a reduction on the feedback repression occurs. Nonetheless, an autoregulated system is normally superior, and only, under certain conditions, equal to the unregulated one (Becskei, A. and Serrano, L., 2000).

Although these systems seem to be useful for noise reduction, little attention has been paid to whether they could be a good alternative in the reduction of noise in downstream genes that are regulated by the same repressor. In *Saccharomyces cerevisiae* this system reduces noise but in *Escherichia coli* the noise could be relatively high, caused by a small cell volume or by a short mRNA lifetime and amplified downstream in transcriptional cascades (Hooshangi, S. and Weiss, R., 2006). So, to surpass this problem and to spread the noise reduction to a gene of interest a repressive system with the gene fused with the repressor followed by the cleavage of both proteins to achieve one-to-one expression could be used. Other alternative could be bicistronic expression letting the expression of the repressor and the gene of interest to occur at different levels to eliminate transcriptional noise (Hensel, Z., 2017).

In operons it is normal to find polycistronic transcription. Possibly reducing noise in genetic networks and to improve the efficiency of heteromeric protein complexes production. Both bicistronic (Sneppen, K. *et al.*, 2010) and polycistronic systems have been implemented in *in vitro* expression systems (Karig, D. K. *et al.*, 2012) and a compensation for the variation on plasmid copy numbers was observed with the addition of getting a relatively linearized inducer dose-response with noise below the “extrinsic noise limit" for chromosomal genes (Taniguchi, Y. *et al.*, 2010). In Hensel, Z., 2017, three systems were also tested as shown in figure 1.1. The figure 1.1c and d show the simulations of these systems for both response to induction and noise. As we can see the simulations indicate that the

constitutive system as wider dynamic range, but higher noise than the bicistronic system. Flow cytometry experiments (figure 1.1a and b) showed unexpected high noise at intermediate levels of induction for a plasmid in which the repressor is constitutively expressed. These simulations in figure 1.1c and d indicated that a system with both bicistronic autoregulation and weak constitutive expression could possibly allow both low gene expression noise and high dynamic range. In this work new regulatory systems based in the paper referred above were created and tested. Strong, promoters regulated by the lac or tet repressors (Lutz, R. and Bujard, H., 1997) were used since they were well characterized to have large differences between repressed and induced expression levels, reaching up to a 5000-fold range.

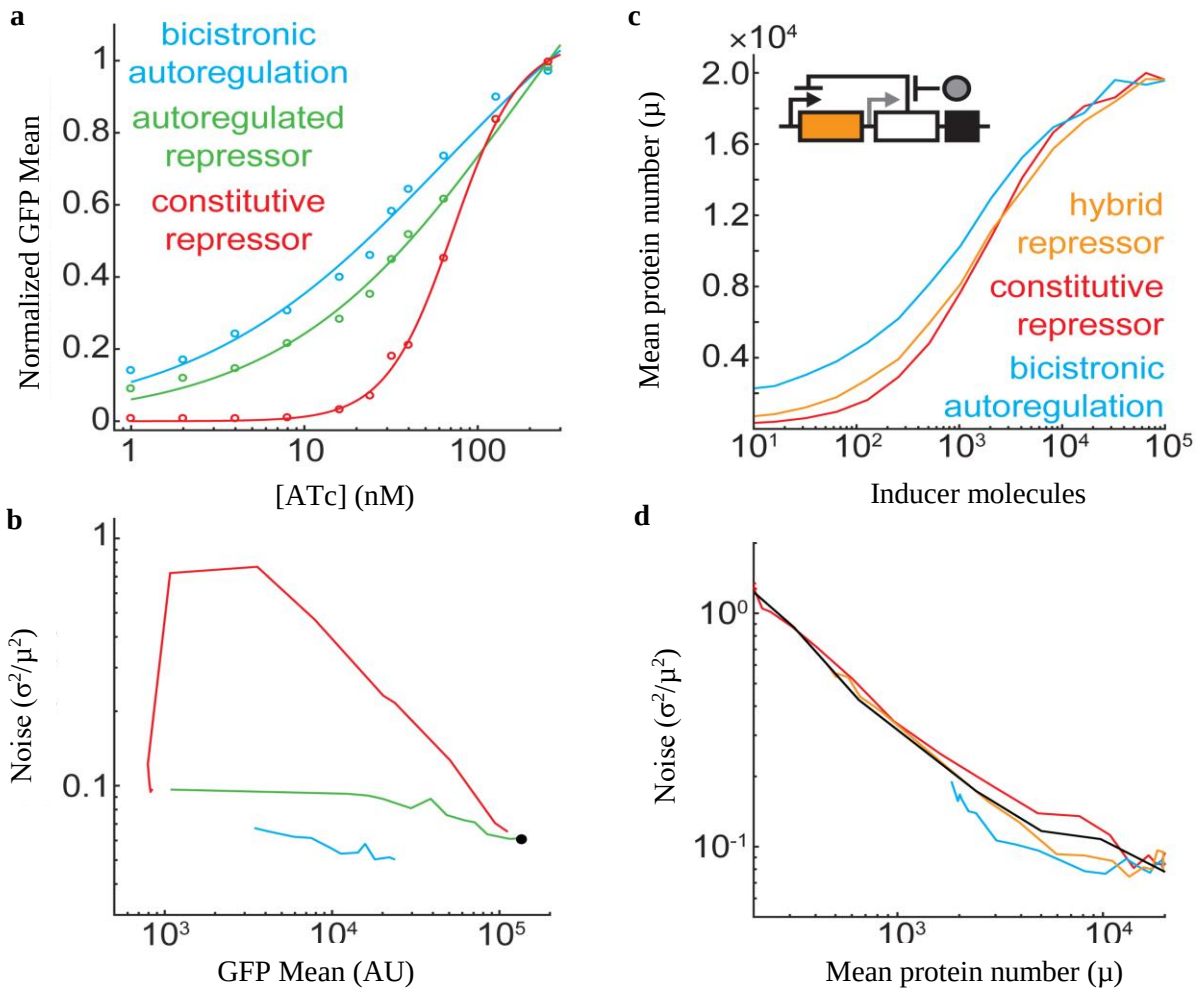

Figure 1.1. Comparison of regulatory circuits and increased dynamic range with a hybrid circuit. (**A**) GFP induction is measured by flow cytometry and fit using the Hill equation for pZH509 (blue, *nh* = 0.60 +/- 0.16), pZH517 (green, *nh* = 0.65 +/- 0.14) and pZH520 (red, *nh* = 2.24 +/- 0.22). Data and fit curves are normalized to the fit value at 256 nM ATc. (**B**) Noise dependence on mean expression level; coloring identical to **Fig 1.1.A**. Black dot, pZH514 at 32 nM ATc. Noise cannot be calculated for pZH509 at 0 nM ATc because of low expression. (**C**) A hybrid scheme is proposed (inset) in which repressor (white box) expression occurs both from autoregulated (black arrow) and relatively weak (gray arrow) promoters that share a transcription terminator (black box). This achieves an inducer dose-response in the gene of interested (orange) that is less steep than in the absence of autoregulation (red) while increasing the dynamic range relative to bicistronic autoregulatory circuit (blue). (**D**) The hybrid system reduces noise relative to the system with constitutively expressed repressor, with noise at or below the extrinsic limit (black). All simulations include extrinsic noise. Reproduced from Hensel, Z., 2017.

The transcripts contained both the repressor and the gene of interest (GFP) and transcription was induced using anhydrotetracycline (ATc) or isopropyl β-D-1-thiogalactopyranoside (IPTG), respectively. The

objective was to create and test new variants of this system, and to try and create a two-plasmid system that would allow for independent expression of two proteins that will be useful, for example, to combine single mRNA and single protein imaging in the same experiment.

Our preliminary experiments using these new plasmids focused on single mRNA detection. Earlier single-molecule mRNA detection in living bacterial cells used the bacteriophage MS2 coat protein fused with GFP to image RNA dynamics (Golding, I. and Cox, E. C., 2004) and transcriptional bursting (Golding, I. *et al.*, 2005). However, this presents some problems such as aggregation artifacts, as the coat protein is a structural protein that easily binds to other MS2 coat proteins. These artifacts are difficult to eliminate in bacteria because of their small size. In addition, mRNAs have short lifetimes in bacteria, further complicating usage of this system (Llopis, P. M. *et al.*, 2010). A way to overcome the problems for the MS2 system would be to use an mRNA-binding fluorescent protein with low-noise expression.

There are some alternative techniques that can be used in the detection and quantitative determination of a specific mRNA such as smFISH (Kwon, S., 2013) and quantitative real-time PCR (Reue, K., 1998).

A good alternative is smFISH because it is useful when one wants to measure single-cell heterogeneity or mRNA localization and it can be combined with immunofluorescent protein detection. It provides more information concerning transcription, cellular localization, and protein expression association in individual cells (Kwon, S., 2013), in addition to average mRNA levels. The drawback of this system is that it is limited to fixed cells, but it is a well characterized method that is a good test of the capabilities of our plasmid-based system for single mRNA detection.

The type of systems used in this work are still poorly characterized, and the purpose of this study was to further test the effects on reduction of noise, to extend dynamic range so it would be possible to analyze targets at or near the limits of detection and to implement such systems as a useful tool for single-molecule experiments.

Objectives:

- Create a two-plasmid system that would allow for independent expression of two proteins with low noise:
    1. Test new plasmid backbone
    2. Test different inducible system
- Apply the system to mRNA imaging

2. Methods

2.1. Plasmids engineering

Plasmid engineering began with plasmids previously described (Hensel, Z., 2017). Briefly, a plasmid was synthesized by Genewiz (New Jersey, USA) by inserting a synthetic sequence into the pUC57-Amp vector. The high-copy pUC origin of replication was replaced by p15a (estimated at 20–30 copies per cell (Lutz, R. and Bujard, H., 1997)). In the pZH501 plasmid, a non-fluorescent protein is expressed (a fusion of the bacteriophage lambda protein CI and SNAP-tag (Keppler, A. et al., 2003)). The CI-SNAP ORF in this plasmid was replaced by GFPmut2 (Cormack, B. P. *et al*., 1996) to create pZH509; GFPmut2 is referred to as "GFP" throughout this manuscript. The full DNA sequence of region encoding inducible GFP expression is included in the annexes. It includes the hybrid PLtetO-1 promoter (containing bacteriophage λ PL promoter overlapped by two copies of the tetO2 sequence) (Lutz, R. and Bujard, H., 1997), open reading frames with independent ribosome binding sites and double stop codons for GFP and tn10 TetR (Postle, K. *et al.*, 1984), and the rrnB T1 transcription terminator (Hartvig, L. and Christiansen, J., 1996). Plasmid pZH520 with constitutively expressed TetR was generated by isothermal assembly of one fragment amplified by inverse PCR of pZH509 and another fragment containing the moderate strength, constitutive proB promoter (Joseph, H. D. *et al.*, 2011). Insert DNAs were synthesized by Integrated DNA Technologies (Iowa, USA). PCR primer synthesis and DNA sequencing was performed by StabVida (Lisbon, Portugal).

The ISB0072 construct was integrated into Escherichia coliMG1655 chromosome by lambda red recombination (Datsenko, K. A. and Wanner, B. L., 2000) replacing the lac operon. This construct is similar to ZHX222 (Hensel, Z. and Marquez-Lago, T. T., 2015) except it has PR -35 and -10 sequences that were weakened to get lower PR transcription rate, the mVenus-Cro ribosome binding site was replaced with the strong RBS #136 (Espah Borujeni, A. *et al.*, 2014), and 24 tandem PP7stemloop repeats from pDZ251 (Larson, D. R. *et al.*,2011) were added to the 3' untranslated region of the PR transcript.

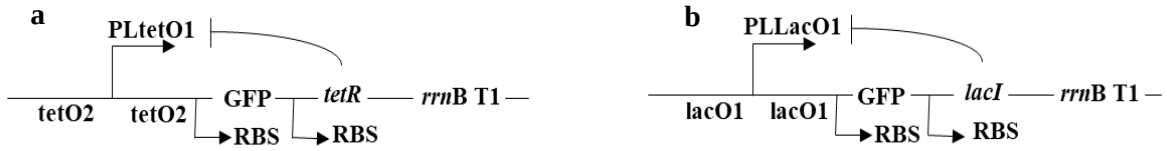

Figure 2.1. Schematic of autoregulatory constructs. (A) The PLtetO-1 promoter encodes the bicistronic transcript for GFP and TetR and is terminated at rrnB T1. TetR binding to either of two tetO2 sites represses transcription. (B) The PLLacO1 promoter encodes the bicistronic transcript for GFP and LacI and is terminated at rrnB T1. LacI binding to either of two lacO1 sites represses transcription. For pJS103 the only difference is the lacO1 sites are instead lacOsym.

For this work three plasmids, pJS101, pJS102 and pJS103, containing respectively *tetO*, *lacO* and *lacOsym* sites (Lutz, R. and Bujard, H., 1997 and Sadler, J. R. *et al.*, 1983) were constructed. pJS101 was created using pGB2 (containing a pSC101 origin of replication (Churchward G. *et al.*, 1984)) as template for the backbone and the inserts (PLTetO1 and TetR) were obtained using pZH509 as template (figure 2.1a). In pJS102, TetR was replaced by LacI and PLTetO1 replaced by PLLacO1 in the pZH509 (figure 2.1b). To create the pJS103 plasmid the same was done with the exception that PLTetO1 was replaced for PLLacOSym. Promoter PLlacOsym lacks the central G-C base pair in PLlacO1 LacI binding sites, which are symmetric with the exception that the second *lacOsym* sequence has one base changed to maintain the PLlacO1 promoter -10 sequence.

Plasmids JS23103, JS23105 and JS23106 were constructed using pZH509 as a template and inserting weak constitutive promoters with relative strengths of 1%, 24% and 47%, respectively, between the GFP gene and the repressor gene. Reported activities of the promoters are given as the relative fluorescence of these plasmids in strain TG1 grown in LB media to saturation (Anderson, J. C., 2006). All constructs were obtained using PCR and 1- or 2-fragment isothermal assembly (Gibson, D. G. *et al.*, 2009). PCR reactions were treated with DpnI for 30 minutes at 37°C and purified with Nzytech clean-up kit.

Plasmids were all carbenicillin resistant except for pJS101 and pZH742 that was spectinomycin resistant.

Table 2.1: Plasmids used in this study.

Plasmid	**Ori**	**Promoter**	**GOI**	**Repressor Expression**	**Reference**
pZH501	p15a	CI-SNAP-tag	$P_{LtetO-1}$	bicistronic	[Hensel Z., 2017]
pZH509	p15a	GFPmut2	$P_{LtetO-1}$	bicistronic	[Hensel Z., 2017]
pZH520	p15a	GFPmut2	$P_{LtetO-1}$	constitutive	[Hensel Z., 2017]
pJS101	pSC101	GFPmut2	$P_{LtetO-1}$	bicistronic	This work
pJS102	p15a	GFPmut2	$P_{LlacO-1}$	bicistronic	This work
pJS103	p15a	GFPmut2	$P_{LlacOsym}$	bicistronic	This work
pJS23103	p15a	GFPmut2	$P_{LtetO-1}$	constitutive/bicistronic	This work
pJS23105	p15a	GFPmut2	$P_{LtetO-1}$	constitutive/bicistronic	This work
pJS23106	p15a	GFPmut2	$P_{LtetO-1}$	constitutive/bicistronic	This work
pZH713	p15a	PP7cp-SYFP2	$P_{LlacO-1}$	bicstronic	This work
pZH742	pSC101	PP7cp-mNeonGreen	$P_{LlacO-1}$	bicistronic	This work
pZH740	p15a	mScarlet-I-24xPP7sl	$P_{LtetO-1}$	constitutive	This work

Purified plasmids were transformed into *Escherichia coli* TOP10 cells (Invitrogen) for flow cytometry experiments and into *Escherichia coli* MG1655 cells for microscopy by growing 2 mL of culture in SOB media at 37ºC to OD600=0.4, washing twice with 1 mL ice-cold water, resuspending in 40 µL water, electroporation of 1–10 ng plasmid with the EC1 setting of a Micropulser (BioRad), and recovering for 1 hour at 37ºC in SOC media before plating on selective LB-agar media.

Plasmid pZH713 is the same as pJS102 except that GFP was changed for PP7cp (Larson, D. R. *et al.*, 2011) fused with SYFP2 (Kremers, G. J. *et al.*, 2006).

pZH740 used pZH520 as backbone with the addition of Pf3 Coat protein (Kiefer, D. *et al.* 1997), mScarlet-I (Bindels, D. S. *et al.*, 2016) and 24 tandem PP7 stemloop repeats (Larson, D. R. *et al.*, 2011). For pZH742 the pGB2 was used as template for the backbone with the addition of a PLLacO1, a PP7 Coat Protein, mNeonGreen (Shaner, N. C. *et al.*, 2013) and a LacI gene. These plasmids were co-transformed into *Escherichia coli*MG1655 by electroporation following the above protocol except with 1 µL each undiluted plasmid (~20–40 ng) and selecting on plates with both spectinomycin and carbenicillin.

All plasmids were verified by sequencing and sequences are available in annexes.

2.2. Growth conditions

For microscopy experiments, cells were grown in 1 mL cultures of M9A (48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 19 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$) supplemented with 50 µg/ml carbenicillin or spectinomycin, 1 mL 50X MEM amino acids (Without L-Glutamine, SIGMA Life Science M5550) and 0.4% glucose ("M9A" media) at 32°C with shaking in 14-mL polypropylene culture tubes. Overnight cultures were diluted 1:100 or more (if multiple experiments were to be done sequentially over several hours) and maintained in exponential growth until observation in the microscope. TetR expression was induced by the addition of anhydrotetracycline hydrochloride (diluted from 100 µM stock in 50% ethanol) and LacI expression with the addition of Isopropyl β-D-1-thiogalactopyranoside both inducing GFP expression in their respective strains. In growth conditions with M9A with 1% SOB, doubling time was approximately 30 minutes for *Escherichia coli* MG1655 strains harboring these plasmids and 60 minutes when only using M9A.

For flow cytometry experiments, cells were grown in M9A media supplemented with 1% rich SOB media (2% tryptone, 0.5% yeast extract, 8.6 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$) and 50 µg/mL carbenicillin at 37°C except for pJS101 to which 50 µg/ml spectinomycin was added instead. Overnights were diluted 1:100 in same media with the addition of the inducer and maintained in exponential growth (grown for approximately 3 hours) until observation in the flow cytometer. After this they were diluted 1:100 in 1 mL PBS.

2.3. Microscopy

Cells growing exponentially were spotted onto M9A agarose gel pads (3% BP165-25, Fisher BioReagents) and kept at room temperature being imaged immediately after sample preparation to prevent that substantial cell growth was observed during data acquisition. The imaging was made using a DMI6000 SD microscope (Leica), using a Leica EL6000 external light source at the highest intensity (filter cube Leica GFP ET), 100x/1.46 a-plan apochromat oil immersion objective and an Evolve 512 EM-CCD camera to quantitatively compare the intensity of GFP molecules and the noise. This camera has a linear response to fluorescence intensity, which is proportional to the number of GFP molecules in the cell. The GFP mean fluorescence was obtained by analyzing microscopy images using ImageJ. The background of the pictures and also the pZH501 GFP fluorescence were subtracted from all strains

to determine the GFP mean. For smFISH imaging the microscope used was Zeiss Axio Observer.Z1, using colibri 2 or HXP 120 V lamps, a plan-apochromat 100/1.4 ph3 objective and a CoolSNAP hq camera (Photometrics), and filter cubes for YFP, CY5 (ATTO590), and TEXAS RED (QUASAR 670).

2.4. Flow Cytometry

Using the BioRad S3e cell sorter, GFP fluorescence (488 nm excitation with 525/30 nm bandpass filtered emission) was measured for all the strains. This device has a linear response to fluorescence intensity and is calibrated daily using fluorescent beads. The concentrations of inducer used were 0, 1, 8, 32, 64 and 128 nM ATc (the negative control, pZH501 at 0 and 128 nM ATc) or 0, 7.81, 62.5, 250, 500 and 1000 µM IPTG. Samples were collected at a target of 2,000 events per second for 30,000 events (FSC gain: 400; Threshold: 1.5; SSC gain: 280; FL1 gain: 800). Data was gated using the program FlowCal that is described in Tabor, J. J. et al., 2017. It used the forward-scattering height by side-scattering height density plot to take the one third of events with FSC and SSC values near the peak in a 2D histogram. The FlowCal software made the choice based on the region with the highest density of events. The gated events provided the mean GFP fluorescence and the standard deviation data.
Noise was calculated using the formula:

$$CV^2 = \frac{\sigma^2}{\mu^2}$$

Equation 2.1.

Noise equals the squared coefficient of variation (CV^2): ratio between variance and square mean (where σ is the standard deviation and µ the GFP fluorescence mean). Data collected for non-fluorescent pZH501 strain was used to subtract background fluorescence from the other strains.

2.5. smFISH

The protocol used is described in Skinner, S. O. *et al.*, 2013. The probe labelled with Atto590 was synthesized by IDT (Iowa, USA) and designed to have 50% GC content, with its sequence repeated 12 times in target mRNAs. Stellaris RNA FISH probes labelled with Quasar670 were synthesized by Biosearch Technologies and they hybridize with Venus and Cro genes. This protocol was started at step 12 and finished at step 31. The media used was M9A/1% SOB with 50 µg/ml carbenicillin and 100 µM of IPTG. At step 16 the direct method (B) was used with small changes in steps B (ii) (from 400 g at 8 min to 2000g, 5 min and 4ºC), B (iv) and 19 (from 600g to 1200 g for both). The centrifuge used was Micro Star 17R (VWR). Both the single anti-PP7 probe and the anti-VenusCro probe were added at the same time and they were added so the total concentration was the same (not the concentration of each individual probe). Probes sequences are available in annexes.

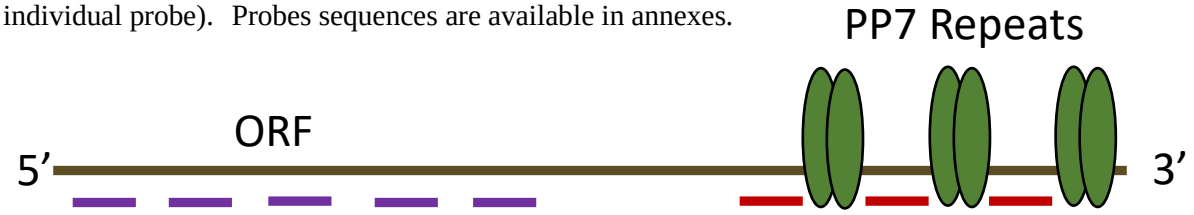

Figure 2.2. Representation of the binding of PP7cp to PP7 repeats and the binding of the probes to the ORF and to PP7 repeats.

3. Results

3.1. Effects of lower plasmid copy number on low-noise expression

A bicistronic system as shown in figure 2.1 was tested in two different backbones, pZH509 with a p15a origin of replication and pJS101 with pGB2 backbone and a pSC101 origin of replication, to show if this system would maintain the same low-noise behavior. They differ only in the resistance to antibiotics, spectinomycin for pJS101 and carbenicillin for pZH509, and in the plasmid copy number with pZH509 reaching 20 to 30 copies per cell and pJS101, 10 to 12 (Lutz, R. and Bujard, H., 1997).

The objective was to create a plasmid with this system with a compatible origin of replication to the original one (p15A); this was the first step towards our goal of making two compatible plasmids that can be used to independently modulate gene expression with low noise for two genes of interest.

But first, we tested two inducers for the pLTetO-1 promoter. ATc is an analog of tetracycline that binds TetR more tightly allowing for full induction without toxicity (Lederer, T. *et al.*, 1996), but it may lead to more noise at low induction levels because of its very strong binding and because low nanomolar concentrations correspond to only a few molecules per cell. Previous work was conducted with ATc,

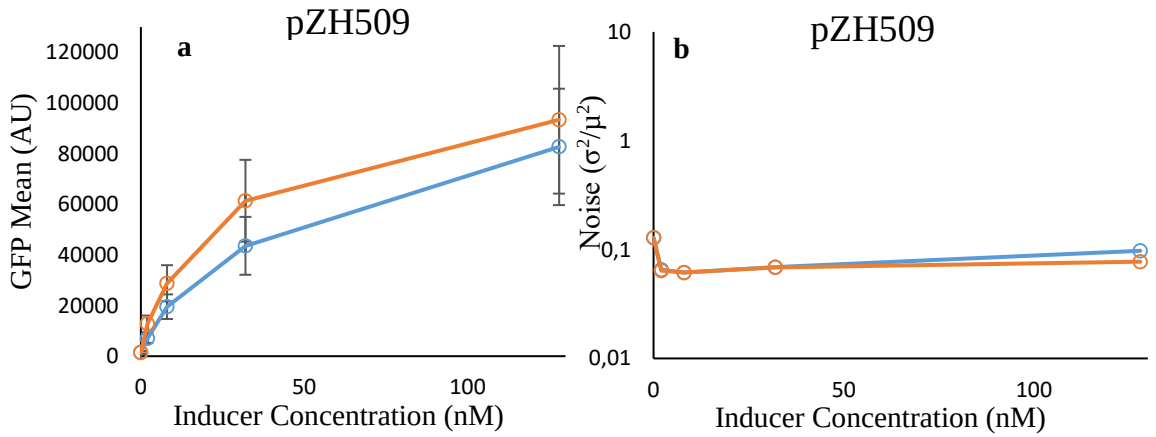

Figure 3.1. Response of TetR construct to strong-and weak-binding inducers. (A) Response of pZH509 to inducers ATc (orange) and tetracycline (blue). Error bars represent standard deviation of the mean from three experiments. (B) Dependence of gene expression noise on inducer concentration. Coloring identical in both graphics.

and we hypothesized that tetracycline would have less noise at lower levels of induction. As shown in figure 3.1a the curves are shaped similarly, with ATc permitting higher expression of GFP than tetracycline at a given concentration. No difference in noise was observed for ATc compared to tetracycline (figure 3.1b). For all other experiments ATc was used as the inducer of choice because with tetracycline we sometimes used concentrations that are close to the level where it becomes toxic.

Next, we analyzed the effect of moving the pZH509 construct (Hensel, Z., 2017) to a lower copy plasmid. GFP fluorescence data was collected using flow cytometry during exponential growth for different concentrations of the inducer. Noise was calculated using equation 2.1. As observed in figure 3.2d and e it is possible to see that the curves of induction are pretty similar between both strains with respect to ATc dependence for half induction. However, pJS101 shows a lower GFP fluorescence mean than pZH509: almost a 3-fold difference. A less drastic difference was expected, and this difference can be partially explained by the lower plasmid copy number (Lutz, R. and Bujard, H., 1997). There is also the possibility that leaky transcription from promoters upstream of pLTetO-1 in the pZH509 backbone would give higher GFP expression (Rosano, G. L. and Ceccarelli, E. A., 2014). However, we

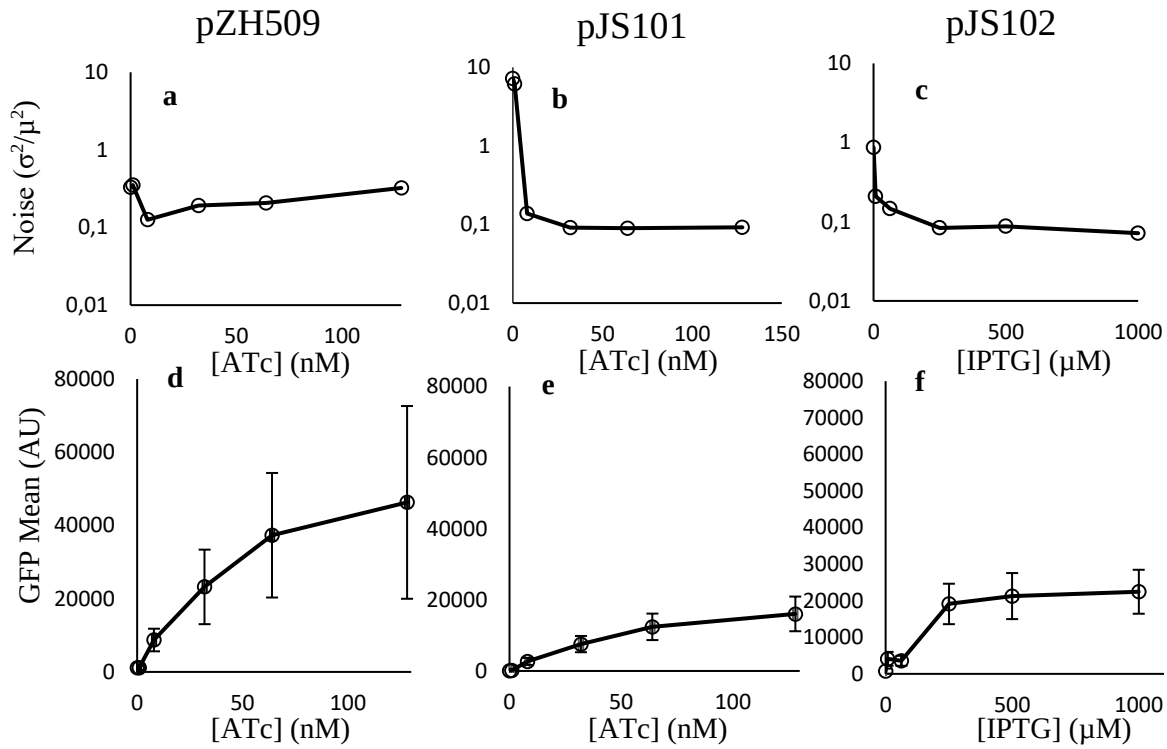

Figure 3.2. Comparison of different repressor systems. (A), (B) and (C) Dependence of gene expression noise on inducer concentration for pZH509, pJS101 and pJS102 plasmids respectively. (D), (E) and (F) Response to inducer for pZH509, pJS101 and pJS102 plasmids respectively. Error bars represent standard deviation of the mean from three experiments. Data collected using flow cytometry.

see that negative autoregulation does not fully compensate for plasmid copy number differences (Becskei, A. and Serrano, L., 2000).

When we compare noise for both strains (figure 3.2a and b) we see that noise is high for low induction levels, but this is because the cell sorter is less sensitive at lower levels of fluorescence which can have affected the GFP mean values and consequently the calculation of noise (Hensel, Z., 2017). At 32 nM of ATc it reaches a plateau for pJS101 where for pZH509 the noise varies constantly decreasing at intermediate levels of induction and then increasing for higher levels. This can be explained because at higher inducer concentrations the system can change from a state of regulation to a state of no regulation that is caused by a reduction on the feedback repression (Becskei, A. and Serrano, L., 2000). This does not occur for pJS101 maybe because the level of protein expressed is lower, obscuring the u-shaped noise curve because of relatively high measurement noise.

These results show that this system has a different behavior for noise and for GFP expression in different backbones but maintains the important characteristics of noise reduction and a relatively linearized dose-response curve. So, pJS101-type plasmids can replace pZH509-type plasmids by performing experiments at a higher ATc concentration.

3.2. A new low-noise expression system using the Lac repressor

The second step towards our goal of creating two, compatible plasmids for low-noise gene expression was to replace the tetracycline repressor and its promoter with another repressor/promoter pair that also could achieve low expression noise through bicistronic autoregulation. For this experiment two strains, pJS102 and pJS103, containing respectively a lac promoter (with two lacO1) and a lacOsym promoter (with two lacOsym that is a symmetrical 20 bp synthetic operator of lacO1) were constructed

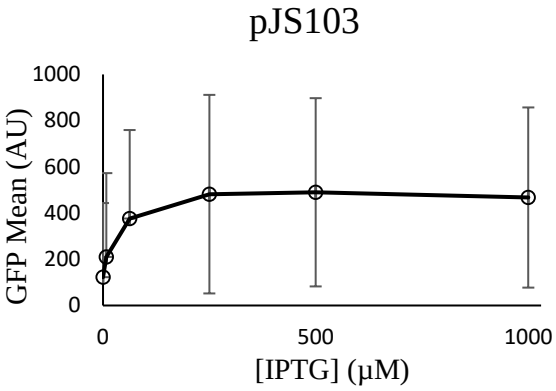

Figure 3.3. Response to inducer for plasmid pJS103. Error bars represent standard deviation of the mean from three experiments. Data collected using flow cytometry.

using the same backbone as pZH509, differing only in the promoter and repressor (LacI vs TetR). This way we compared both LacI-dependent promoters to the tet TetR-dependent one. The objective was to create a plasmid with a promoter induced by a different inducer, and also to explore how the strength of repressor binding affected gene expression noise. Combining with changing the backbone, this will allow for independent control of two genes.

As shown in figure 3.3 pJS103 was poorly induced even at the maximum concentration of inducer in the media reaching a plateau at above 100 μM IPTG concentration. The response to IPTG occurs over a similar range as pJS102, but with lower maximum induction and reduced dynamic range.

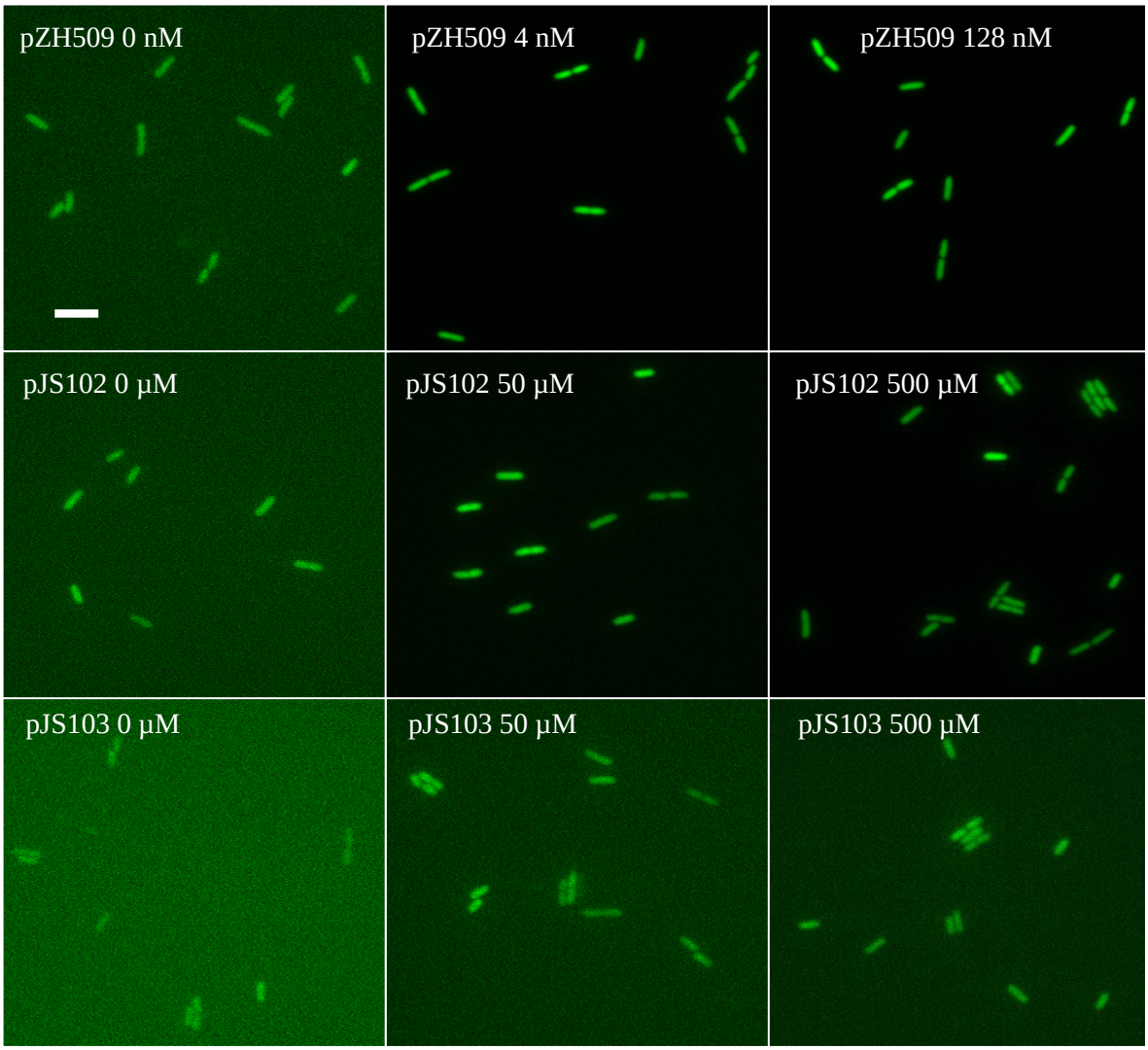

Figure 3.4. Induction of TetR construct compared to LacI constructs with and without very strong binding sites. Fluorescence of plasmid pJS102 is compared to the plasmid pJS103 and to pZH509 for different induction concentrations (0, 50, 500 μM IPTG and 0,4 128 nM ATc for pZH509). Scale bar 4 μm.

Taking in account this data, it seems that this strain is being heavily repressed even at maximum induction. In Brewster, R. C. et al., 2014, a model was created and predicts the fold-change taking only in account the fold-change for simple repression where a transcription factor represses the gene of interest. We hypothesize because in our case, the binding energy is stronger, that long lived LacI:lacOsym binding makes it impossible to achieve similarly high levels of expression for pJS103 and pJS102. Even with more inducer concentration the repression is still maintained, possibly because the repressor is not unbinding the promoter, so the inducer cannot act to bind LacI and derepress the promoter. So, even very high IPTG concentrations, there is very little GFP expression. This data is consistent with binding of the repressor LacI to PLLacOsym being stronger than binding to PLLacO1. All the plasmids were sequenced, and no mutation was found so there is not possibility for a mutation that would interfere in the binding of the repressor. Despite reduced dynamic range, pJS103-type plasmids may be useful for low level expression control.

As for pJS102, the induction curve has a similar shape to the one observed for pZH509 (figure 3.2d and f) but the GFP mean is lower reaching a plateau around 250 µM IPTG. Tet has a higher regulatory range than the Lac promoter and is stronger, so this difference was expected (Lutz, R. and Bujard, H., 1997). The noise is lower for pJS102 at higher levels of induction and the same or slightly higher at lower levels (figure 3.2a and c). Thus, LacI-based bicistronic autoregulation can replace or complement TetR-based expression in future experiments.

Microscopy data corroborated flow cytometry data, with the analyses showing that pJS103 has lower fluorescence and higher noise than the other strains (figure 3.4).

Both TetR and Lac-I-dependent promoters work similarly even though pZH509 has higher expression of protein and noise at high her levels of induction. However, we attribute pZH509 noise at high induction partially to difficulties in obtaining reproducible data by flow cytometry. This was not seen in earlier work by flow cytometry or microscopy (Hensel, Z., 2017), and further microscopy experiments will be used to measure gene expression noise with less experimental error.

3.3. Inserting constitutive promoters to increase dynamic range

One problem pZH509 showed was a limited dynamic range as shown in figure 1.1b (Hensel, Z., 2017). To overcome this issue, getting a range closer to pZH520 and at the same time maintaining the noise response and induction curve observed for pZH509 (figure 1.1.a and b), different constitutive promoters (23103, 23105 and 23106) were inserted between the GFP gene and the repressor gene, with different relative strengths (1%, 24% and 47%). This way TetR expression is increased in the absence of ATc permitting the system to get to lower levels of protein expression. The hypothesis was that a constitutive promoter with a rate of transcription equivalent to 40% of the rate of the bicistronic promoter would be perfect to achieve this objective (Hensel, Z., 2017). However, it is difficult to accurately predict promoter strengths, so various constitutive promoters described in Anderson, J. C., 2006 were tested and compared to the plasmid without any promoter in that region (pZH509).

Flow cytometry data was collected for all of the strains and microscope data was only collected to fill some gaps in respect to low levels of induction where, as said before, the cell sorter is less sensitive.

As observed in figure 3.5a, the strain with the induction curve closer to the pZH509 one was pJS23103. The other strains at low levels of induction have low expression because the promoters are too strong and at higher induction levels get to only half of the mean fluorescence of both pZH509 and

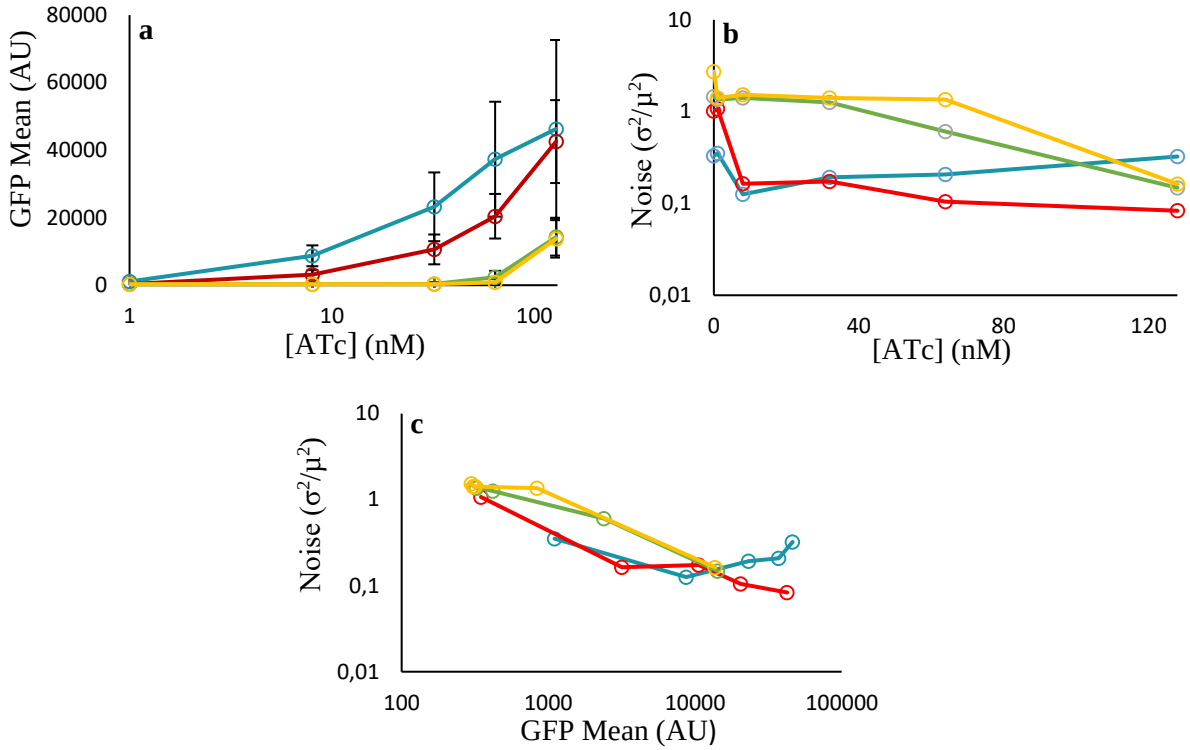

Figure 3.5. Bicistronic system compared to bicistronic/constitutive systems. (A) Response to inducer for pZH509 (blue), pJS23103 (red), pJS23105 (green), pJS23106 (yellow). (B) Dependence of gene expression noise on inducer concentration. (C) Dependence of gene expression noise on average expression. Coloring identical in all graphics. Error bars represent standard deviation of the mean from three experiments. Data collected using flow cytometry.

pJS23103. This is similar to what was observed for pZH520 in which TetR is only expressed constitutively (Hensel, Z., 2017). These strains need more inducer to reach the same level of expression, and, like pZH520, exhibit high noise at intermediate induction levels. As for the noise of pJS23103 (figure 3.5b), it shows again a more similar behavior to pZH509 than the other strains, with even lower noise at high levels of induction. In addition, the dynamic range was somewhat wider than pZH509with a lower uninduced expression level, meaning it can get to lower expression of proteins while maintaining high levels of expression with low noise at high inducer concentration (figure 3.5c).

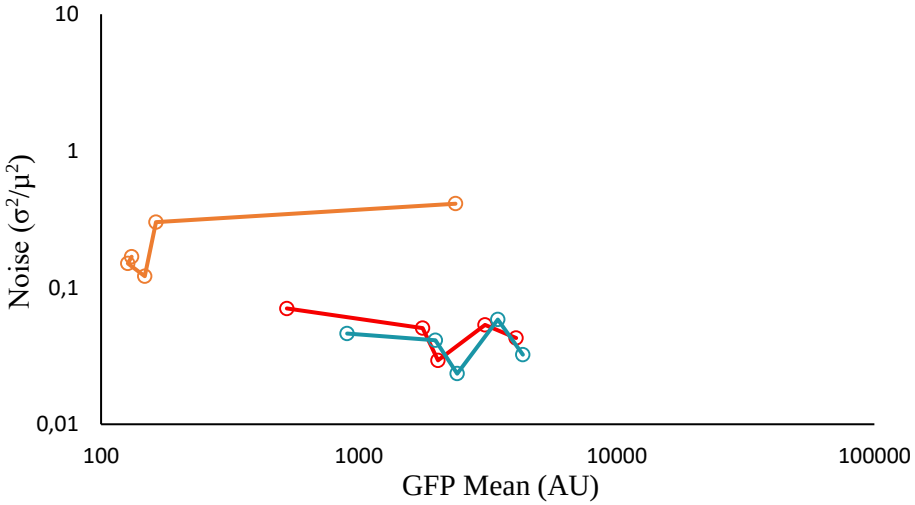

Figure 3.6. Dependence of gene expression noise on average expression. Microscopy data for pZH509 (blue), pJS23103 (red) and pzH520 (orange).

Microscopy data corroborates the results for pJS23103; no data was collected for the promoters from the other strains because they showed worst behavior than pZH509 by flow cytometry. It too was compared to pZH520 because this plasmid shows high noise at intermediate induction.

The microscope data (figure 3.6) shows less noise but the same curve as the flow cytometry one, with a decrease at 4 nM for both strains. An explanation for such a big difference in the noise calculated using these techniques is probably because for microscopy data we obtained the mean fluorescence of the cell while for flow cytometry the total intensity of GFP is measured thus, variation in cell size contributes more to noise in flow cytometry experiments. This could be addressed using more microscopy experiments or by using a flow cytometer other than the BioRad S3; scattering data from other flow cytometers can more accurately reflect cell size (Hensel 2017).

Although pJS23103 seems better than pZH509 it should be further tested because some incongruences were found between the no induction data from flow cytometry and microscopy data.

3.4. Applications of new plasmids for single-molecule imaging

3.4.1. Utilizing the LacI-based low-noise expression system

In this experiment plasmid pZH713, similar to pJS102 except instead of GFP it produces PP7cp-SYFP2, was used. In figure 3.7 we can see different levels of induction of the protein PP7cp-SYFP2. At 40 µM of IPTG the expression of the protein is too high making it difficult to distinguish the spots that correspond to PP7cp bound to the PP7 stemloop repeats from the background and at levels of induction lower than 5 µM the expression is too low to any spots to be observed. This way we could find the most

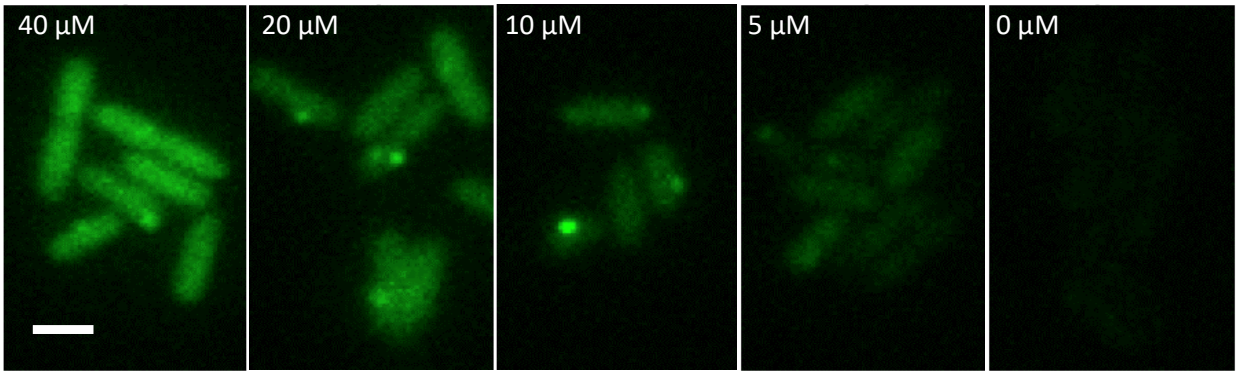

Figure 3.7. PP7cp-SYFP2 induction. Different concentrations of IPTG (0, 5, 10, 20 and 40 µM) were used for plasmid pZH713. Intensity scaling identical for all images. Scale bar 2 µm.

suited IPTG concentration for single-molecule experiments. The next experiment involved the use of smFISH technique since it is an independent method for counting and localizing single mRNAs. For this we used the strain MG1655 that contains the ISB0072 construct in the chromosome and that we know from unpublished work expresses single mRNAs encoding mVenus-Cro and containing the 24 tandem PP7stemloop repeats (figure 2.2) since it exhibits bursts of fluorescence from mVenus-Cro expression that are separated in time (Yu, J. *et al.*, 2006).

The objective was to make sure that the PP7cp-SYFP2 spots corresponded with the localization of the PP7 stemloop repeats and the ORF because to have a good PP7 mRNA reporter we wanted a 3-color colocalization of the ORF to happen most of the time, to show that single mRNAs detected by PP7cp-SYFP2 accurately reflected the copy number of the ORF.

In figure 3.8c, the spots correspond to the probe bond to the mRNA containing the repeats while in figure 3.8b we have the transcription of the mRNA from the chromosome. Figure 3.8d corresponds to

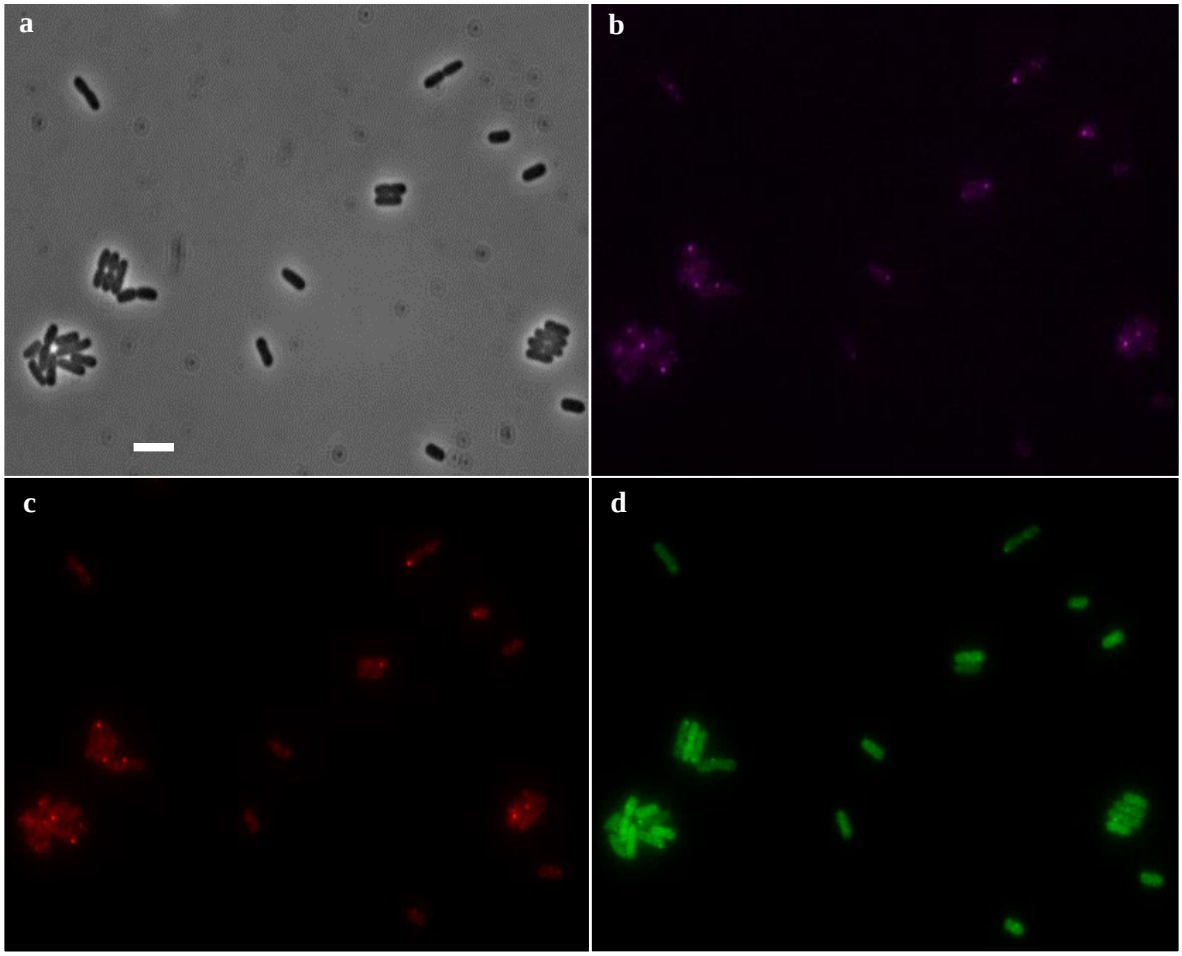

Figure 3.8. Plasmid pZH713 smFISH. (A) Phase contrast (B) Cro-Venus ORF with smFISH probes labeled with Quasar670. (C) PP7 mRNA repeats with single smFISH probe labeled with Atto590. (D) PP7cp-SYFP2 expression. Scale bar 4 µm.

PP7cp-SYFP2 expression induced by IPTG. When we compare figure 3.8b, c and d it is possible to observe that the localization of the protein PP7cp-SYFP is almost the same as the repeats. In addition, the almost perfect correlation with what we observed for the Cro-Venus ORF means that the spots in figure c are probably single mRNAs.

The conclusion is that this plasmid exhibited low gene expression noise for PP7cp-SYFP2, making it possible to tune the protein expression level to detect spots of localized, bound PP7cp-SYFP2 over the background of unbound protein and that it was possible to observe single mRNAs.

3.4.2. Two-plasmid system

The systems tested above permitted the creation of a two-plasmid system. These plasmids (pZH740 and pZH742) were co-transformed in Escherichia coli to allow for an independent expression of two genes. This was possible because the plasmids have different antibiotics selection (carbenicillin and spectinomycin) and origins of replications that are compatible (the p15a and pSC101 origins, respectively), which were used together in previous synthetic biology experiments (Lee, T.S. *et al.*, 2011). The plasmid pZH740 was based on pZH520, has 24 tandem PP7 stemloop repeats, a constitutive promoter for TetR (proB promoter), and a TetR-regulated promoter that allows the expression of the protein Pf3 fused with mScarlet-I when ATc is added to the media. The constitutive promoter leads to a continuous expression of TetR that strongly represses the expression of Pf3-mScarlet-I in the absence of inducer. Plasmid pZH742 is based on both pJS101 (pSC101 origin of replication and spectinomycin resistance)

and pJS102 (expression regulated by LacI), with a bicistronic promoter and expresses a protein (PP7cp that is fused to mNeonGreen), when IPTG is added. The PP7cp-mNeonGreen fusion protein binds to the PP7 stemloops.

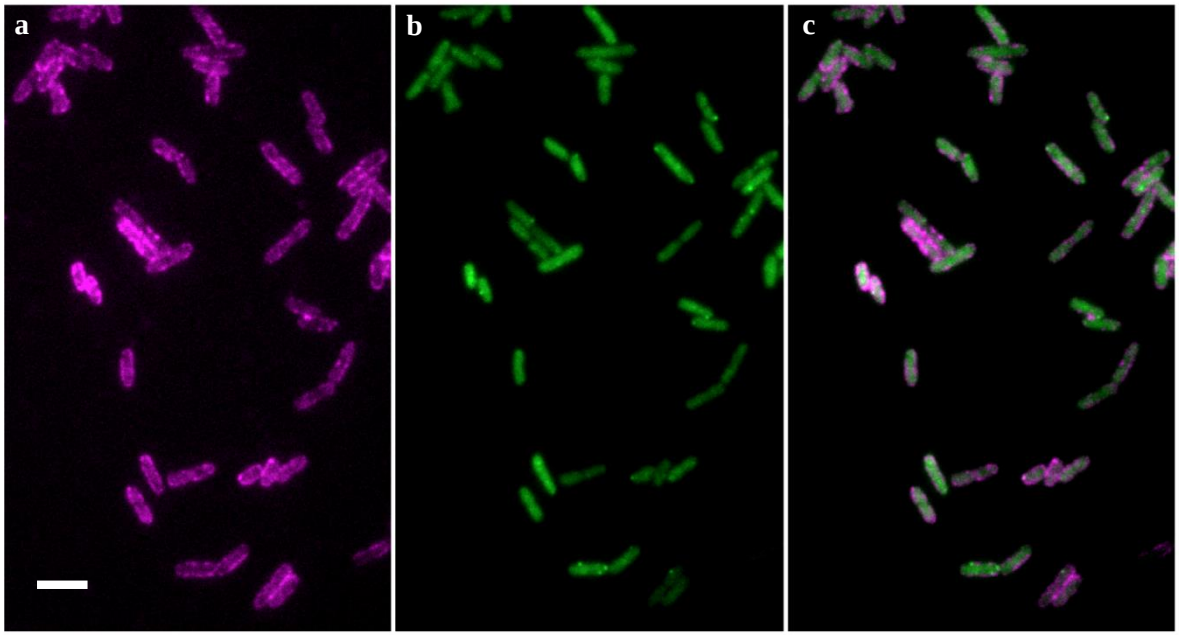

Figure 3.9. Two-plasmids strain. (A) Pf3-mScarlet-I from pZH740. (B) PP7cp-mNeonGreen from pZH742. (C) Overlay of both images (a and b) Scale bar 5 μm.

Figure 3.9a shows the expression of Pf3-mScarlet-I protein. This protein has an N-terminus that binds to the cell membrane with mScarlet-I oriented into the cytoplasm (Kiefer *et al.*, 1997). In figure 3.9b we can see again the expression of PP7cp, but this time fused with mNeonGreen. In figure 3.9c we have the overlay of both images. With this we show that a two-plasmid system can be used for independent expression of two distinct genes with different inducers. They do not seem to show incompatibilities and seem to be a useful tool in future experiments. Preliminary data shows that PP7cp-mNeonGreen exhibits less aggregation than PP7cp-SYFP2 (see cells with multiple spots in fig 3.9b), suggesting that fluorescent proteins with less tendency to dimerize, such as mNeonGreen, may be useful for imaging strains with high numbers of mRNAs.

4. Discussion

pJS101, pJS102 and pJS103 were constructed with the objective of creating a plasmid that would be used in a two-plasmid system. We could characterize the behavior of these systems with respect to mean and noise during induction and are now using them in single-molecule microscopy experiments.

We tested two inducers (tetracycline and ATc) of the Tet promoter and unexpectedly we found out that tetracycline does not have lower noise at lower induction levels, but more experiments need to be made so we can further characterize the behavior of both inducers relatively to noise and protein expression.

The second experiment was to test different backbones where we found that the plasmid pJS101 had a similar low noise, but unexpectedly low levels of GFP mean expression. It was found that TetR has a more inefficient repression when the plasmid is a high copy one and this happens since there is a change in the operators to repressors ratio in addition to an increase in unspecific binding sites that will affect the concentration of free repressor (Lutz, R. and Bujard, H., 1997) and this would explain the difference we found. There is even the possibility of different leaky transcription from other promoters in the pZH509 backbone that would justify this data (Rosano, G. L. and Ceccarelli, E. A., 2014), but we have yet to test it.

In the third experiment, we tested different promoters so we could have expression of two proteins in the same cell with the objective of using it for the two-plasmid system.

The plasmid pJS102 worked really well but, pJS103 showed the unexpected behavior of extremely low GFP mean expression at maximum induction. In Brewster, R. C. *et al*., 2014 a model that can predict the fold-change based on the ratio of expression of a gene in the presence of a transcription factor by the expression of the same gene in the absence of this factor was created and this particular formula takes only in account the fold-change for simple repression where the TF only binds to the gene of interest. Our constructs are more complex, with autoregulated repressor expression and two binding sites.

The formula is:

$$fold - change = \frac{1}{1 + \frac{R}{N_{NS}} \mathrm{e}^{-\Delta\varepsilon/k_B\mathrm{T}}}$$

Equation 4.1.

Where R is the number of repressors present in the cell, N_{NS} is the size of the nonspecific binding reservoir, and $\Delta\varepsilon$ is the binding energy of repressor to its operator. This way the parameters $\Delta\varepsilon$ and R are the only that can be different for both plasmids. The pJS103 promoter is constituted by two *lacOsym* and the energy of binding for this promoter could be, taking the formula in account, much more negative making it harder for the repressor to unbind and preventing the transcription of the gene to occur. All the plasmids were sequenced, and no mutation was found so there is not a possibility for a mutation that would interfere in the binding of the repressor. To our knowledge, an equivalent fold change formula has not been derived for an autorepressive system with two binding sites. More work is required to see if this is consistent with gene regulation models, and to test how effective binding sites of intermediate strength work will be for low noise gene expression.

The constitutive promoters tested with the objective of extending the dynamic range of the pZH509 system did not worked as well as expected, although pJS23103 seems to have improved a little bit the dynamic range. Our hypothesis was that a promoter of 40% the strength of the tetO one would work well, but we only tested one sufficiently weak promoter. Nonetheless, we now hypothesize that a promoter with a strength comprised between the 23103 and the 23105 promoters strengths may give us a better dynamic range and maintain the pZH509 proprieties.

The two-plasmids system was created with the intention of controlling expression of two proteins at once with low gene expression and reproducible dose response to induction using repressed promoters (negative autoregulation). This way we tested first if LacI would allow for the detection of a protein that would bound repeats on the mRNAs over the background of unbound proteins. In the second experiment smFISH was used so we could see if the spots corresponded to the binding of the PP7cp-SYFP2 to the repeats in the mRNA. What we expected was a 3-color colocalization of the ORF to happen most of the time. Since the reading frame is transcribed before the repeats, and the two halves might degrade at different times we cannot have 100% accurate colocalization. Considering the data collected we can assume that we achieved our objective. However, it is difficult to know exactly if the ORF and the PP7 repeats are present at the same time (figure 3.8b) since PP7 repeats could aggregate and have a longer lifetime than the reading frame. This would not allow for an accurate report of the transcription that actually occurred in the cell. This system needs further testing with live-cell microscopy, and the ultimate proof will be correlating mRNA spots with bursts of protein expression from that mRNA.

In the last experiment we have shown the expression of two genes from different plasmids in the same cell. This proves that the two-plasmid system works but more tests need to be made. In the future, the protein Pf3 fused with mScarlet-I will be used to facilitate the counting of the fluorescence spots since this protein is a bacteriophage coat protein that in other organisms assembles into a filamentous bacteriophage that pierces the cell wall, but that does not occur in *Escherichia coli* so it will work just as a membrane anchor. In this case, will anchor the FP to the membrane as it has an N-terminus that binds to the cell membrane and a C-terminus that always faces towards the cytoplasm making it easier to count proteins (Kiefer, D. *et al.* 1997). The gene for this protein is transcribed to an mRNA that has the PP7 repeats. In addition, the protein PP7cp fused with SYFP2 can bind to those repeats. This way we can count both mRNAs and the proteins expressed, what will allow us to find out if we have a protein expression of 1-to-1 relatively to mRNAs. For this a fluorescent protein that would avoid photobleaching PP7cp-SYFP2 protein that is expressed at low levels was necessary. The preference was a red fluorescent protein. This way, mScarlet-I was chosen because, although it is not very photostable, it matures quickly in comparison to other RFPs (Balleza, E. *et al.*, 2018) and it is very bright.

With these experiments concluded and if we get the expected data we can even try to expand the system to express 3 genes. This would be possible if we create a new plasmid using a different repressor that can have the same characteristics as the ones tested in this work (simple repressors that bind strongly to two sites overlapping the promoter) and the corresponding inducer , a compatible origin of replication for example pMB1 that has a similar copy number plasmid as p15A (Patrick, M. *et al*.,2017), a different fluorescent protein such as orange fluorescent protein that has a compatible wavelength with the other proteins used, meaning it will not photobleach, and a different antibiotic selection, for example kanamycin. The expansion of this system to express more genes is possible but there are many limitations such as, the existence of a limited number of compatible plasmids and antibiotics selections and the fact that having many plasmids can stress out the cell. A way to reduce the number of plasmids needed to expand the system could be inserting more than one gene per plasmid the drawback would be the size of the plasmids that would be way bigger than with only one gene.

5. Bibliography

Anderson, J. C. (2006, August 24). Promoters Catalog. Retrieved September 10, 2018, from http://parts.igem.org/Promoters/Catalog/Anderson

Austin, D. W. *et al.* (2006) 'Gene Network Shaping of Inherent Noise Spectra', Nature, 439(7076), pp. 608–611. doi: 10.1038/nature04194.

Balleza, E. *et al.* (2018) 'Systematic Characterization of Maturation Time of Fluorescent Proteins in Living Cells', *Nature Methods*, 15(1), pp. 47–51. doi: 10.1038/nmeth.4509.

Becskei, A. and Serrano, L. (2000) 'Engineering Stability in Gene Networks by Autoregulation', *Nature*, 405(6786), pp. 590–593. doi: 10.1038/35014651.

Bindels, D. S. *et al.* (2016) 'mScarlet: A Bright Monomeric Red Fluorescent Protein for Cellular Imaging', *Nature Methods*, 14(1), pp. 53–56. doi: 10.1038/nmeth.4074

Binder, D. *et al.* (2016) 'Comparative Single-cell Analysis of Different E. Coli Expression Systems During Microfluidic Cultivation', *PLoS ONE*, 11(8), pp. 1–19. doi: 10.1371/journal.pone.0160711.

Brewster, R. C. *et al*. (2014) 'The Transcription Factor Titration Effect Dictates Level of Gene Expression', *Cell*, 156(6), pp. 1312–1323. doi: 10.1016/j.cell.2014.02.022.

Churchward G. *et al.* (1984) 'A pSC101-derived Plasmid which Shows no Sequence Homology to Other Commonly used Cloning Vectors', *Gene*, 31, pp. 165–171. doi: 10.1016/0378-1119(84)90207-5

Cormack, B. P. *et al.* (1996) 'FACS-optimized Mutants of the Green Fluorescent Protein (GFP)', Gene, 173(1), pp. 33–38. doi: 10.1016/0378-1119(95)00685-0.

Datsenko, K. A. and Wanner, B. L. (2000) 'One-step Inactivation of Chromosomal Genes in Escherichia coli K-12 using PCR Products.', *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), pp. 6640–5. doi: 10.1073/pnas.120163297.

Dublanche, Y. *et al.* (2006) 'Noise in Transcription Negative Feedback Loops: Simulation and Experimental Analysis', *Molecular Systems Biology*, 2, pp. 1–12. doi: 10.1038/msb4100081.

Espah Borujeni, A. *et al.* (2014) 'Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites', *Nucleic Acids Research*, 42(4), pp. 2646–2659. doi: 10.1093/nar/gkt1139.

Gibson, D. G. *et al.* (2009) 'Enzymatic Assembly of DNA Molecules up to several Hundred Kilobases', *Nature Methods*, 6(5), pp. 343–345. doi: 10.1038/nmeth.1318.

Golding, I. and Cox, E. C. (2004) 'RNA Dynamics in live Escherichia coli Cells', *Proceedings of the National Academy of Sciences*, 101(31), pp. 11310–11315. doi: 10.1073/pnas.0404443101.

Golding, I. *et al.* (2005) 'Real-time Kinetics of Gene Activity in Individual Bacteria', *Cell*, 123(6), pp. 1025–1036. doi: 10.1016/j.cell.2005.09.031.

Hartvig, L. and Christiansen, J. (1996) 'Intrinsic Termination of T7 RNA Polymerase Mediated by either RNA or DNA', *The EMBO Journal*, 15(17), pp. 4767–4774. doi: 10.1002/j.1460-2075.1996.tb00854.x

Hensel, Z. and Marquez-Lago, T. T. (2015) 'Cell-cycle-synchronized, oscillatory expression of a negatively autoregulated gene in E. coli', *arXiv:1506.08596 [q-bio]*, pp. 1–52.

Hensel, Z. (2017) 'A Plasmid-based Escherichia coli Gene Expression System with cell-to-cell Variation below the Extrinsic Noise Limit', *PLOS ONE*, 12(10). doi: 10.1371/journal.pone.0187259.

Hooshangi, S. and Weiss, R. (2006) 'The Effect of Negative Feedback on Noise Propagation in Transcriptional Gene Networks', *Chaos*, 16(2). doi: 10.1063/1.2208927.

Joseph, H. D. *et al.* (2011) 'Design, Construction and Characterization of a set of Insulated Bacterial Promoters', *Nucleic Acids Research*, 39(3), pp. 1131–1141. doi: 10.1093/nar/gkq810

Karig, D. K. *et al.* (2012) 'Expression Optimization and Synthetic Gene Networks in Cell-free Systems', *Nucleic Acids Research*, 40(8), pp. 3763–3774. doi: 10.1093/nar/gkr1191.

Keppler A. *et al.* (2003) 'A General Method for the Covalent Labeling of Fusion Proteins with small Molecules in Vivo', *Nature Biotechnology*, 21(1), pp. 86–89. Doi: 10.1038/nbt765

Kremers, G. J. *et al.* (2006) 'Cyan and Yellow Super Fluorescent Proteins with improved Brightness, Protein Folding, and FRET Förster Radius', *Biochemistry*, 45(21), pp. 6570–6580. doi: 10.1021/bi0516273.

Kwon, S. (2013) 'Single-molecule Fluorescence in Situ Hybridization: Quantitative Imaging of Single RNA Molecules', *BMB Reports*, 46(2), pp. 65–72. doi: 10.5483/BMBRep.2013.46.2.016.

Larson, D. R. *et al.* (2011) 'Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene', *Science*, 332(6028), pp. 475–478. doi: 10.1126/science.1202142.Real-Time.

Lederer, T. *et al.* (1996) 'Tetracycline Analogs Affecting Binding to Tn10-encoded Tet Repressor Trigger the same Mechanism of Induction', *Biochemistry*, 35(23), pp. 7439–7446. doi: 10.1021/bi952683e.

Lee, T.S. *et al.* (2011) 'BglBrick Vectors and Datasheets: A Synthetic Biology Platform for Gene Expression', Journal of Biology, 5 (12). doi: 10.1186/1754-1611-5-12.

Llopis, P. M. *et al.* (2010) 'Spatial Organization of the Flow of Genetic Information in Bacteria', *Nature*. Nature Publishing Group, 466(7302), pp. 77–81. doi: 10.1038/nature09152.

Lutz, R. and Bujard, H. (1997) 'Independent and Tight Regulation of Transcriptional Units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements', *Nucleic Acids Research*, 25(6), pp. 1203–1210. doi: 10.1093/nar/25.6.1203.

Nevozhay, D. *et al.* (2009) 'Negative Autoregulation Linearizes the Dose-response and Suppresses the Heterogeneity of Gene Expression', *Proceedings of the National Academy of Sciences*, 106(13), pp. 5123–5128. doi: 10.1073/pnas.0809901106.

Patrick, M. *et al.* (2017) *Plasmids 101: A Desktop Resource* Retrieved from https://blog.addgene.org/plasmid-101-origin-of-replication

Paulsson, J. (2004) 'Summing up the Noise in Gene Networks', *Nature*, 427(6973), pp. 415–418. doi: 10.1038/nature02257

Postle, K. *et al.* (1984) 'Nucleotide Sequence of the Repressor Gene of the TN10 Tetracycline Resistance Determinant', *Nucleic Acids Research*, 12(12), pp. 4849–4863.

Raj, A. and van Oudenaarden, A. (2008) 'Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences', *Cell*, 135(2), pp. 216–226. doi: 10.1016/j.cell.2008.09.050.

Reue, K. (1998) 'Issues and Opinions in Nutrition mRNA Quantitation Techniques: Considerations for Experimental', *The Journal of Nutrition*, pp. 2038–2044.

Rosano, G. L. and Ceccarelli, E. A. (2014) 'Recombinant Protein Expression in Escherichia coli: Advances and Challenges', *Frontiers in Microbiology*, 5(APR), pp. 1–17. doi: 10.3389/fmicb.2014.00172.

Sadler, J. R. *et al.* (1983) 'A Perfectly Symmetric lac Operator Binds the lac Repressor very Tightly.', *Proceedings of the National Academy of Sciences of the United States of America*, 80(22), pp. 6785–9. doi: 10.1073/pnas.80.22.6785.

Shaner, N. C. *et al.* (2013) 'A Bright Monomeric Green Fluorescent Protein derived from Branchiostoma lanceolatum', *Nature Methods*, 10(5), pp. 407–409. doi: 10.1038/nmeth.2413.A.

Shen-Orr, S. S. *et al.* (2002) 'Network Motifs in the Transcriptional Regulation Network of Escherichia coli', *Nature Genetics*, 31(1), pp. 64–68. doi: 10.1038/ng881.

Simpson, M. L. *et al.* (2003) 'Frequency Domain Analysis of Noise in Autoregulated Gene Circuits', *Proceedings of the National Academy of Sciences*, 100(8), pp. 4551–4556. doi: 10.1073/pnas.0736140100.

Skinner, S. O. *et al.* (2013) 'Measuring mRNA Copy Number in Individual Escherichia coli Cells using Single-molecule Fluorescent in situ Hybridization', *Nature Protocols*, 8(6), pp. 1100–1113. doi: 10.1038/nprot.2013.066.

Sneppen, K. *et al.* (2010) 'Economy of Operon Formation: Cotranscription Minimizes Shortfall in Protein Complexes', *mBio*, 1(4), pp. 3–5. doi: 10.1128/mBio.00177-10.

Swain, P. S. (2004) 'Efficient Attenuation of Stochasticity in Gene Expression through Post-transcriptional Control', *Journal of Molecular Biology*, 344(4), pp. 965–976. doi: 10.1016/j.jmb.2004.09.073.

Tabor, J. J. *et al.* (2017) 'FlowCal: A User-Friendly, Open Source Software Tool for Automatically Converting Flow Cytometry Data from Arbitrary to Calibrated Units', *ACS Synthetic Biology*, 5(7), pp. 774–780. doi: 10.1021/acssynbio.5b00284.FlowCal.

Taniguchi, Y. *et al.* (2010) 'NIH Public Access', *Most*, 329(5991), pp. 533–538. doi: 10.1126/science.1188308.Qu

Tsanov, N. *et al.* (2016) 'SmiFISH and FISH-quant - A Flexible Single RNA Detection Approach with Super-resolution Capability', *Nucleic Acids Research*, 44(22), pp. 1–11. doi: 10.1093/nar/gkw784.

Tsimring, L. S. (2014) 'Noise in biology', *Reports on Progress in Physics*, 77(2), pp. 1–62. doi: 10.1088/0034-4885/77/2/026601.antifying.

Tyagi, S. (2010) ‘E. coli, What a Noisy Bug’, *Science*, 329(5991), pp. 518–519. doi: 10.1126/science.1194036.

Veening, J.-W. *et al.* (2008) ‘Bet-hedging and epigenetic inheritance in bacterial cell development’, *Proceedings of the National Academy of Sciences USA*, 105(11), pp. 4393–4398. doi: 10.1073/pnas.0700463105.

Wang, Y.-H., Wei, K. Y. and Smolke, C. D. (2013) ‘Synthetic Biology: Advancing the Design of Diverse Genetic Systems’, *Annual Review of Chemical and Biomolecular Engineering*, 4(1), pp. 69–102. doi: 10.1146/annurev-chembioeng-061312-103351.

Yu, J. *et al.* (2006) ‘Probing Gene Expression in Live Single Escherichia coli Cells – One Molecule at a Time’, *Science*, 311(5767), pp. 1600–3. doi: 10.1126/science.1119623.

6. Annexes

The synthetic sequence incorporated into pZH509 is shown below. Transcription start and translation start/stop signals are capitalized and bolded and other features are marked as indicated:

tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcac**A**tcagcaggacgcactgaccgaattcattaaagaggagaaag
gtaccgc**ATG**agtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaattttctgtcagtg
gagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacctgttccatggccaacacttgtcactactt
tcgcgtatggtcttcaatgctttgcgagatacccagatcatatgaaacagcatgacttttcaagagtgccatgcccgcgaaggttatgtacaggaaag
aactatatttttcaaagatgacgggaactacaagacacgtgctgaagtcaagtttgaaggtgatacccttgttaatagaatcgagttaaaaggtattga
ttttaaagaagatggaaacattcttggacacaaattggaatacaactataactcacacaatgtatacatcatggcagacaaacaaaagaatggaatcaa
agttaacttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttta
ccagacaaccattacctgtccacacaatctgccctttcgaaagatcccaacgaaaagagagaccacatggtccttcttgagtttgtaacagctgctgg
gattacacatggcatggatgaactatacaa**TAATAA**tctagcaggaggatttcacc**ATG**tctagattagataaaagtaaagtgattaacagc
gcattagagctgcttaatgaggtcggaatcgaaggtttaacaacccgtaaactcgcccagaagctaggtgtagagcagcctacattgtattggcatgt
aaaaaataagcgggctttgctcgacgccttagccattgagatgttagataggcaccatactcacttttgccctttagaaggggaaagctggcaagattt
tttacgtaataacgctaaaagttttagatgtgctttactaagtcatcgcgatggagcaaaagtacatttaggtacacggcctacagaaaaacagtatgaa
actctcgaaaatcaattagccttttttatgccaacaaggtttttcactagagaatgcattatatgcactcagcgctgtggggcattttactttaggttgcgtatt
ggaagatcaagagcatcaagtcgctaaagaagaaagggaaacacctactactgatagtatgccgccattattacgacaagctatcgaattatttgatc
accaaggtgcagagccagccttcttattcggccttgaattgatcatctgcggattagaaaaacaacttaaatgtgaaagtgggtct**TAATAA**ctg
cagcccgggggatcccatggtacgcgtgctagaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttat

TetO2 GFPmut2 TetR rrnB T1

The sequence for pZH520 is shown below and annotated as above:

tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcac**A**tcagcaggacgcactgaccgaattcattaaagaggagaaag
gtaccgc**ATG**agtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaattttctgtcagtgg
agagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacctgttccatggccaacacttgtcactactttc
gcgtatggtcttcaatgctttgcgagatacccagatcatatgaaacagcatgacttttcaagagtgccatgcccgaaggttatgtacaggaaagaact
atatttttcaaagatgacgggaactacaagacacgtgctgaagtcaagtttgaaggtgatacccttgttaatagaatcgagttaaaaggtattgattttaa
agaagatggaaacattcttggacacaaattggaatacaactataactcacacaatgtatacatcatggcagacaaacaaaagaatggaatcaaagtta
acttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccag
acaaccattacctgtccacacaatctgccctttcgaaagatcccaacgaaaagagagaccacatggtccttcttgagtttgtaacagctgctgggatta
cacatggcatggatgaactatacaa**TAATAA**tctagcataaaacgaaaggctcagtcgaaagactgggcctttcgttttatcacagctaacac
cacgtcgtccctatctgctgccctaggtctatgagtggttgctggataactttacgggcatgcataaggctcgtaatatatattcagggagaccacaac
ggtttccctctacaaataattttgtttaacttt**aggaggatttcacc**ATG**tctagattagataaaagtaaagtgattaacagcgcattagagctgcttaat
gaggtcggaatcgaaggtttaacaacccgtaaactcgcccagaagctaggtgtagagcagcctacattgtattggcatgtaaaaaataagcgggctt
tgctcgacgccttagccattgagatgttagataggcaccatactcacttttgccctttagaaggggaaagctggcaagatttttacgtaataacgctaa
aagttttagatgtgctttactaagtcatcgcgatggagcaaaagtacatttaggtacacggcctacagaaaaacagtatgaaactctcgaaaatcaatta
gccttttatgccaacaaggtttttcactagagaatgcattatatgcactcagcgctgtggggcattttactttaggttgcgtattggaagatcaagagcat
caagtcgctaaagaagaaagggaaacacctactactgatagtatgccgccattattacgacaagctatcgaattatttgatcaccaaggtgcagagcc
agccttcttattcggccttgaattgatcatctgcggattagaaaaacaacttaaatgtgaaagtgggtct**TAATAA**ctgcagcccgggggatccc
atggtacgcgtgctagaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttat

proB

The sequences for pJS101 that differ from pZH509:

Spectinomycin resistance:

cgtcggcttgaacgaattgttagacattatttgccgactaccttggtgatctcgcctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggc
caagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggccggcaggcgctccattgcccagtcggcagcga
catccttcggcgcgattttgccggttactgcgctgtaccaaatgcgggacaacgtaagcactacatttcgctcatcgccagcccagtcgggcggcga
gttccatagcgttaaggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctccgccgctggacctaccaaggcaacgct
atgttctcttgcttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaatt
gcagttcgcgcttagctggataacgccacggaatgatgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccagggga
agccgaagtttccaaaaggtcgttgatcaaagctcgccgcgttgtttcatcaagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggctt
caggccgccatccactgcggagccgtacaaatgtacggccagcaacgtcggttcgagatggcgctcgatgacgccaactacctctgatagttgagt
cgatacttcggcgatcaccgcttccctcatgatgtttaactttgttttagggcgactgccctgctgcgtaacatcgttgct

pSC101 origin of replication:

gttacattgtcgatctgttcatggtgaacagctttgaatgcaccaaaaactcgtaaaagctctgatgtatctatctttttacaccgttttcatctgtgcatatg
gacagttttccctttgatatgtaacggtgaacagttgttctacttttgtttgttagtcttgatgcttcactgatagatacaagagccataagaacctcagatcc
ttccgtatttagccagtatgttctctagtgtggttcgttgtttttgcgtgagccatgagaacgaaccattgagatcatacttactttgcatgtcactcaaaaat
tttgcctcaaaactggtgagctgaatttttgcagttaaagcatcgtgtagtgtttttcttagtccgttatgtaggtaggaatctgatgtaatggttgttggtatt
ttgtcaccattcatttttatctggttgttctcaagttcggttacgagatccatttgtctatctagttcaacttggaaaatcaacgtatcagtcgggcggcctcg
cttatcaaccaccaatttcatattgctgtaagtgtttaaatctttacttattggtttcaaaacccattggttaagccttttaaactcatggtagttattttcaagca
ttaacatgaacttaaattcatcaaggctaatctctatatttgccttgtgagtttttcttttgtgttagttcttttaataaccactcataaatcctcatagagtatttgtt
ttcaaaagacttaacatgttccagattatattttatgaatttttttaactggaaaagataaggcaatatctcttcactaaaaactaattctaatttttcgcttgag
aacttggcatagtttgtccactggaaaatctcaaagcctttaaccaaaggattcctgatttccacagttctcgtcatcagctctctggttgctttagctaata
caccataagcattttccctactgatgttcatcatctgagcgtattggttataagtgaacgataccgtccgttctttccttgtagggttttcaatcgtggggttg
agtagtgccacacagcataaaattagcttggtttcatgctccgttaagtcatagcgactaatcgctagttcatttgctttgaaaacaactaattcagacata
catctcaattggtctaggtgattttaatcactataccaattgagatgggctagtcaatgataattactagtccttttcctttgagttgtgggtatctgtaaattct
gctagacctttgctggaaaacttgtaaattctgctagaccctctgtaaattccgctagacctttgtgtgttttttttgtttatattcaagtggttataatttataga
ataaagaaagaataaaaaaagataaaaagaatagatcccagccctgtgtataactcactactttagtcagttccgcagtattacaaaaggatgtcgcaa
acgctgtttgctcctctacaaaacagaccttaaaaccctaaaggcttaagtagcaccctcgcaagctcgggcaaatcgctgaatattccttttgtctccg
accatcaggcacctgagtcgctgtctttttcgtgacattcagttcgctgcgctcacggctctggcagtgaatgggggtaaatggcactacaggcgcctt
ttatggattcatgcaaggaaactacccataatacaagaaaagcccgtcacgggcttctcagggcgttttatggcgggtctgctatgtggtgctatctga
ctttttgctgttcagcagttcctgccctctgattttccagtctgaccacttcggattatcccgtgacaggtcattcagactggctaatgcacccagtaaggc
agcggtatcatcaacaggcttacccgtcttactgtc

pJS102 and pJS103 sequences that differ from pZH509:

PLlacO1 (pJS102): ataaatgtgagcggataacattgacattgtgagcggataacaagatactgagcacatcagcaggacgcactgacc

LacOSym promoter (pJS103): aattgtgagcgctcacaatttgacaattgtgagcgctcacaatgatactgagcacatcag-caggacgcactgacc

LacI:

atgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttct
gcgaaaacgcgggaaaaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggcgggcaaacagtcgttg
ctgattggcgttgccacctccagtctggccctgcacgcgccgtcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtgg
tggtgtcgatggtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatc
cgctggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattatt
ttctcccatgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaagttctgtctc
ggcgcgtctgcgtctggctggctggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtccgg
ttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccatt
accgagtccgggctgcgcgttggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccgccgttaaccaccatca
aacaggattttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtct
cactggtgaaaagaaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggttt
cccgactggaaagcgggcagtaa

The sequence for pZH713 is shown below and annotated as above:

ataaatgtgagcggataacattgacattgtgagcggataacaagatactgagcacAtcagcaggacgcactgaccgaattcattaaattcgagaaa
ggtaccgcatgtctaaaactatcgtgctgtctgttggtgaagcaacccgcaccctgactgaaattcagtctactgctgatcgccaaatcttcgaagaaa
aagttggtccgctggttggccgtctgcgtctgactgcttccctgcgtcagaacggcgcgaagaccgcgtatcgcgttaatctgaaactggatcaagc
ggacgtagtggactctggtctgccgaaagttcgctacacgcaggtgtggtctcacgacgttaccatcgttgctaattccaccgaagccagccgtaagt
ccctctacgacctcaccaagagcctggttgcaacctctcaggttgaggatctggtggtcaacctggtaccactgggccgcccgccggttgccacca
gtaaaggagaagagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagg
gcgagggcgatgccacctacggcaagctgaccctgaagctgatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccct
gggctacggcgtgcagtgcttcgcccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagc
gcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaaggg
catcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcaccgccgacaagcagaag
aacggcatcaaggccaacttcaagatccgccacaacatcgaggacggcggcgtgcagctcgccgaccactaccagcagaacaccccccatcggc
gacggccccgtgctgctgcccgacaaccactacctgagctaccagtccaagctgagcaaagaccccaacgagaagcgcgatcacatggtcctgc
tggagttcgtgaccgccgccgggatcactctcggcatggacgaactatacaaataatctagcaggaggatttcaccatgaaaccagtaacgttat
acgatgtcgcagagtatgccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtgg
aagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacctccagtc
tggccctgcacgcgccgtcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagc
ggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattg
ctgtggaagctgcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacggtacgcga
ctgggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggct
ggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtccggttttcaacaaaccatgcaaatgct
gaatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgcgttg
gtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccgccgttaaccaccatcaaacaggattttcgcctgctgggg
caaaccagcgtggaccgcttgctgcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaaaaccac
cctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagt
aataactgcagcccgggggatcccatggtacgcgtgctagaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttat

PLlacO1 SYFP PP7 Coat Protein lacI

The sequence for pZH740 is shown below and annotated as above:

tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcac**A**tcagcaggacgcactgaccgaattcattaaagaggagaaag
gtaccgcATGcagtccgttattactgatgtgactggccagctgaccgcagtgcaggcggacatcaccactatcggtggcgctattatcgttctggc
tgccgtggtactgggcatccgctggatcaaagcacagtttttcagtaaaggagaagcagtgatcaaggagttcatgcggttcaaggtgcacatggag
ggctccatgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgacca
agggtggccccctgcccttctcctgggacatcctgtcccctcagttcatgtacggctccagggccttcatcaagcaccccgccgacatccccgacta
ctataagcagtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgccgtgaccgtgacccaggacacctccctg
gaggacggcaccctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccccgtaatgcagaagaagacaatgggctgggaag
cgtccaccgagcggttgtaccccgaggacggcgtgctgaagggcgacattaagatggccctgcgcctgaaggacggcggccgctacctggcgg
acttcaagaccacctacaaggccaagaagcccgtgcagatgcccggcgcctacaacgtcgaccgcaagttggacatcacctcccacaacgagga
ctacaccgtggtggaacagtacgaacgctccgagggccgccactccaccggcggcatggacgaactatacaaataactaaggtacctaattgcct
agaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcat
tagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcg
ctggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctaga
aaccagcagagcatatgggctcgctggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgt
cgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcattagatcctaaggtacctaattgccta
gaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcatt
agatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgc
tggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaa
accagcagagcatatgggctcgctggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtc
gctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcattagatcctaaggtacctaattgcctag
aaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcatta
gatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaaccagcagagcatatgggctcgct
ggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcgctccctgcaggtcgactctagaaa
ccagcagagcatatgggctcgctggctgcagtattcccgggttcattagatcctaaggtacctaattgcctagaaaggagcagacgatatggcgtcg
ctccctgcaggtcgactctagaaaccagcagagcatatgggctcgctggctgcagtattcccgggttcatagatccccgggctgcagtaataatct
agcataaaacgaaaggctcagtcgaaagactgggcctttcgttttat

Pf3 Coat Protein mScarlet-I PP7 Repeats

Promoter sequence for pJS23103: ctgatagctagctcagtcctagggattatgctagc

Promoter sequence for pJS23105: tttacggctagctcagtcctaggtactatgctagc

Promoter sequence for pJS23106: tttacggctagctcagtcctaggtatagtgctagc

The sequence for pZH742 is shown below and annotated as above:

gatcatatgataaatgtgagcggataacattgacattgtgagcggataacaagatactgagcac**A**tcagcaggacgcactgaccgaattcattaaatt
cgagaaaggtaccgcatgtctaaaactatcgtgctgtctgttggtgaagcaacccgcaccctgactgaaattcagtctactgctgatcgccaaatcttc
gaagaaaaagttggtccgctggttggccgtctgcgtctgactgcttccctgcgtcagaacggcgcgaagaccgcgtatcgcgttaatctgaaactgg
atcaagcggacgtagtggactctggtctgccgaaagttcgctacacgcaggtgtggtctcacgacgttaccatcgttgctaattccaccgaagccag
ccgtaagtccctctacgacctcaccaagagcctggttgcaacctctcaggttgaggatctggtggtcaacctggtaccactgggccgccgccggctt
gccaccgtatcaaaaggtgaagaagataatatggcaagcttaccagcaacacacgaattacatattttcggttctataaacggagtggattttgatatg
gttggtcaaggcacaggtaacccaaatgatggttatgaagaacttaatcttaaaagtacgaaaggtgacttacaatttagtccttggattttagttccgca
tattggttatgggtttcatcaatacttgccatatccagatggtatgtctccttttcaagcagcgatggttgacggatcgggataccaagtacatagaacga
tgcagtttgaagatggcgcgtcattaacagtcaattacagatatacttatgaagggtcacatattaaaggtgaagctcaagttaaaggtactggcttccc
agctgatggaccagtgatgacaaatagtttaactgcagccgactggtgtcgatcaaagaaaacatatcctaatgataaaacaataatctctacgtttaa
gtggtcatatactacgggaaatggtaaaagatatcgtagcacagctcgcacaacatatacatttgcaaaacctatggcagcaaattatttaaagaatca
accgatgtatgtttttcgtaaaacagaattgaaacatagtaaaactgagctaaactttaaagaatggcaaaaagctttcactgatgtaatgggcatggat
gagttatacaaataataatctagcaggaggatttcaccatgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtctcttatcagaccgtttc
ccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtg
gcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccgtcgcaaattgtcgcggcgattaaa
tctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgc
gcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcctgcactaatgttccggcgttatttcttgat
gtctctgaccagacacccatcaacagtattattttctcccatgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaat
cgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggctggcataaatatctcactcgcaatcaaattcagccgatagcgg
aacgggaaggcgactggagtgccatgtccggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgctggttgccaacg
atcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgcgttggtgcggatatctcggtagtgggatacgacgataccgaagac
agctcatgttatatcccgccgttaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggcca
ggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttgg
ccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtaaTAActgcagcccgggggatcccatggtacgcgtgcta
gaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttat

mNeonGreen

Carbenicillin resistance:

gtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgta
gataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaa
ccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagt
tcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacg
atcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatc
actcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaata
gtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgtt
cttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcacc
agcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttc
aatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggggttccgcgcacatttcccc
gaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc

ssDNA oligo with 5' Atto590 label (probe) and sequence: aatgaacccgggaatactgc

ssDNA oligos with 5’ Quasar670 label and sequences:

taataacctccttagtactt
tgaacagttcttcgcctttg
ttcgaccaggatcggaacta
tttatgaccgttaacgtcgc
cttcgccttcaccagaaacg
gtggtacagatcagcttcag
taacccagagtagtgaccag
aacgtgcgaagcactggagg
tgctgtttcatgtgatcagg
catcgcgcttttgaagaagt
gttcttgaacgtagccttcc
tcgtccttgaagaagatggt
cgcgcgggtcttatagttac
tgtcgccctcaaatttaact
agttcgatgcggtttacgag
cttccttaaaatctatgcct
tttgtgaccgaggatgttac
gagagttgtagttgtactcg
tttgtccgcggtgatgtaaa
gtttgctttaataccgtttt
gtcttcgatgttgtgacgaa
gtgttctgctgatagtgatc
tagctcaggtagtgattgtc
gttcggatctttagacagtt
gaaccatgtgatcgcgtttt
gcagcggtaacgaactcgag
ttatacgctgctccatgtag
aagcgcattgcgtaatcttt
tctttagctgtcttggtttg
ttgttgatcgcgctttgata
aaaaatctttcggcctgcgt
ttccgtcagcgtttatagtt
tttacctcttccgcgtaaac
gctgctgttgtttttttgtt
aatcagcgagagcgtagttt

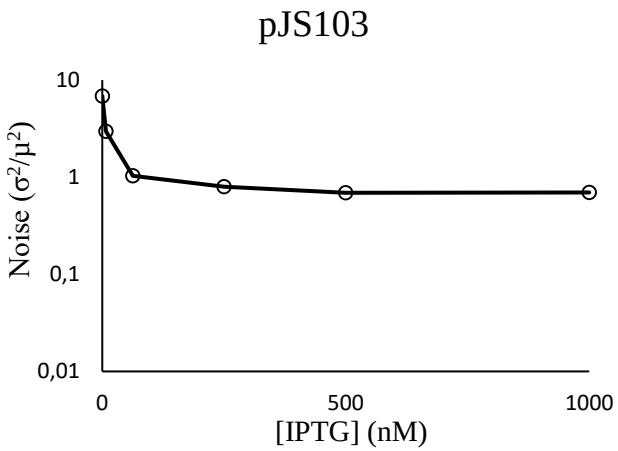

Figure 6.1. Response to inducer for pJS103.

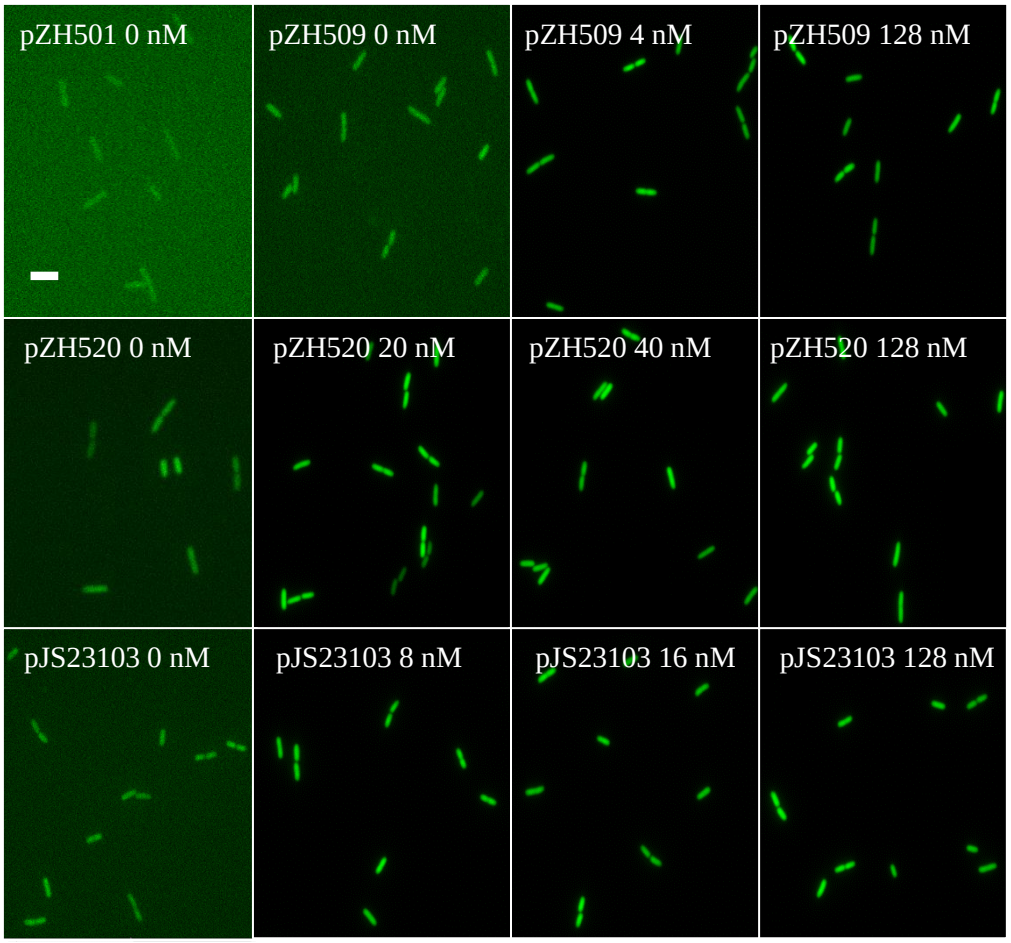

Figure 6.2. Induction of bicistronic system compared to a constitutive and a hybrid system. Comparison of fluorescence of plasmids pZH509, pZH520 and pJS23103 for different induction concentrations that reflect the same expression level for the different strains (0, 4, 8, 16, 20, 40 and 128 nM ATc). Scale bar 4 μm.

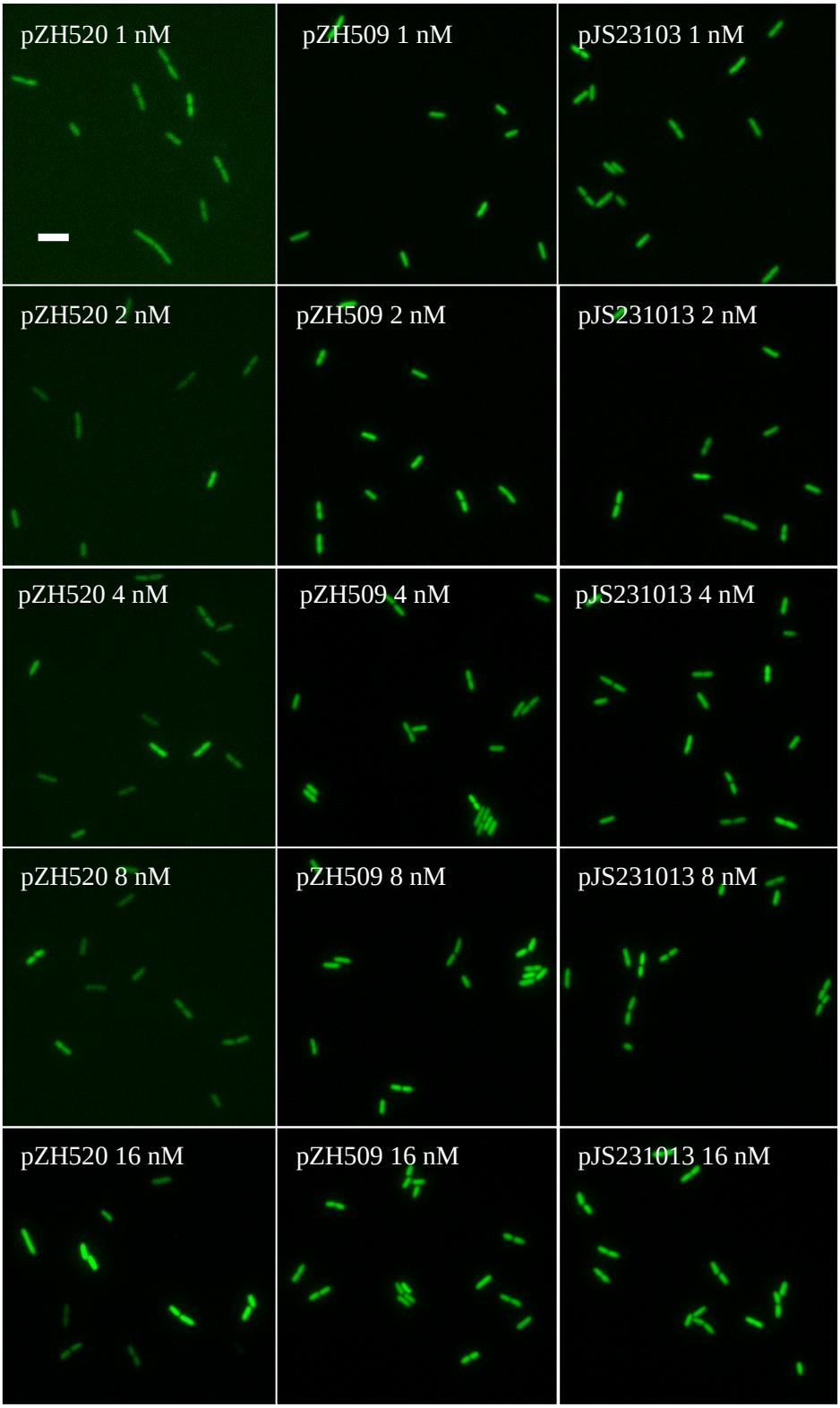

Figure 6.3. Induction of bicistronic system compared to a constitutive and a hybrid system. Comparison of fluorescence of plasmids pZH509, pZH520 and pJS23103 for different induction concentrations (0, 1, 2, 4, 8 and 16 nM ATc). Scale bar 4 μm.