

A Dynamic Model of Resource Allocation in Response to the Presence of a Synthetic Construct

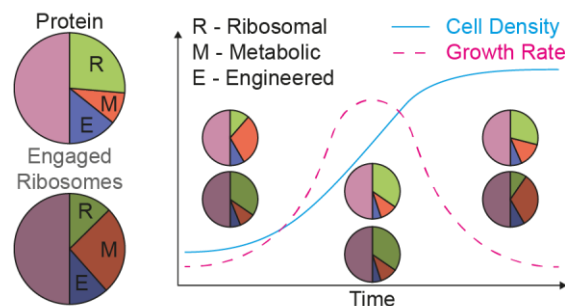
Axel Nyström^{†,‡}, Antonis Papachristodoulou^{*‡} and Andrew Angel^{*†}

[†]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, U.K.

[‡]Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K.

Abstract

Introducing synthetic constructs into bacteria often carries a burden that leads to reduced fitness and selective pressure for organisms to mutate their constructs and hence to a reduced functional lifetime. Understanding burden requires suitable methods for accurate measurement and quantification. We develop a Dynamic Growth Model from physiologically-relevant first principles that allows parameters relevant to burden to be extracted from standard growth curves. We test several possibilities for the response of a bacterium to a new environment in terms of resource allocation. We find that burden manifests in the time taken to respond to new conditions as well as the rate of growth in exponential phase. Furthermore, we see that the presence of a synthetic construct hastens the reduction of ribosomes when approaching stationary phase, altering memory effects from previous periods of growth.



Keywords: synthetic biology, burden, growth, mathematical model

The reliable engineering of synthetic genetic circuits can have a significant impact on a number of areas where Synthetic Biology has application, such as bioremediation or biomedicines¹⁻³. A major current problem in synthetic biology concerns the lifetime of engineered constructs^{4,5}. The insertion of a synthetic construct into a bacterial host can introduce fitness costs, either through toxicity of the product⁶, unwanted interaction with metabolic pathways^{7,8} or burden from diverting cellular resources from growth to the production of exogenous protein⁹. In this work we focus specifically on the fitness cost due to the burden involved in producing excess protein that is non-toxic and metabolically inert. Fitness costs from synthetic constructs leads to a selective pressure which favours hosts with mutations that reduce or remove the expression of constructs¹⁰. Typical synthetic circuits comprising only a few parts have been observed to have a mean half-life on the order of 100 generations¹¹⁻¹³ which is sufficiently short to preclude even growing a population to a volume suitable for industrial application. With increasing size and complexity of circuits, this problem can only worsen. It is thus crucial to understand the fitness costs that synthetic constructs impose so

that they can be properly evaluated and controlled. The ability to measure and/or predict burden early in the design stage could also greatly reduce issues with developing a synthetic circuit with perfect function only to find that it does not last long enough to be of practical use.

One method of enhancing synthetic construct lifetimes is to design circuit components that have reduced potential for mutation by, for example, limiting the inclusion of DNA sequences that are more prone to mutation¹²⁻¹⁴. Another method is to control the burden imposed by synthetic constructs. This can reduce the chance for construct-disrupting mutations to fix in a population or extend the time it takes to fix allowing for an increased overall lifetime and window in which an intervention can take place.

In order to control burden, the effect that a synthetic construct has on the fitness of an organism must first be accurately measured and understood. This can theoretically go beyond simply reducing the maximum growth rate observed in exponential phase, which is the measurement that is currently most frequently utilised. The length of the lag phase following dilution to fresh media or an environmental change could also be affected as could the approach to stationary phase. Other less straightforward aspects of growth can also be affected, for example, the length of time spent in exponential phase has been observed to correlate with a reduction in the separation of growth rates between wild-type bacteria and those carrying constructs¹⁵.

A common experimental setup in synthetic biology research utilises high-throughput well-plates which allows for optical density measurements to produce a growth curve with high temporal resolution. As a measure of fitness costs, the maximum growth rate is often extracted from the data. However, more information about growth can be inferred by fitting the curves to theoretical models of growth. There are a great number of such models in common use and care has to be taken in choosing the model that has the right balance between simplicity and the effective extraction of relevant parameters of growth^{16, 17}. Some of the most widely used models were not constructed with bacterial growth in mind. For the purposes of understanding burden, it is crucial to extract physiologically relevant parameters from growth. Some of the most recent growth models centre on the relative fractions of the transcriptome and proteome that are devoted to growth, metabolism and those that are due to any synthetic circuits¹⁸⁻²⁰. However, the most developed versions of these models are currently static in nature.

In this work, we derive a Dynamic Growth Model, DGM, that takes into account how fractions of the proteome and engaged ribosomes change over time, focussing on the exit from lag phase and entry to exponential growth following a dilution into fresh media. We also discuss how this model can be utilised to accurately measure the burden imposed by synthetic constructs via exogenous protein production and detail its effects on different aspects of growth.

Results

Model Derivation – Dynamic Growth

Recent growth models have centred on the allocation of resources to cellular growth and metabolic function to explain several empirical observations of bacterial growth^{15, 18-23}. A common approach is to divide the proteome into three fractions based on cellular function: Φ_R is the protein fraction of ribosome related proteins, Φ_M is the protein fraction of enzymes for nutrient import and amino acid metabolism and Φ_Q is the remaining fraction which is not related to growth and assumed to be constant. When considering the effect of synthetic constructs a fourth fraction is considered, Φ_E , the protein fraction composed of protein produced by synthetic constructs in the bacteria (a potential, and likely, source of burden). The sum of the protein fractions is equal to unity (Figure 1).

$$\Phi_R + \Phi_M + \Phi_E + \Phi_Q = 1$$

Now, $\Phi_x = \frac{P_x}{P}$, where P is the total protein mass in the system and P_x is the total mass of protein of type x where $x \in \{R, M, E, Q\}$.

The specific growth rate, μ , is defined as the rate of increase of total protein mass per protein mass

$$\mu = \frac{1}{P} \frac{dP}{dt}$$

We focus on the ribosome-limited ramp-up period of growth, which we define as the portion of the growth curve prior to deceleration which occurs later in the growth cycle. As the growth rate drops and the population approaches stationary phase, the ribosomes are no longer the limiting factor for growth and nutrient availability comes into play. We denote the growth rate in the ramp-up period as μ_L . The rate at which protein mass is accumulated depends on the total amount of ribosomes and the rate of translation, γ , $\frac{dP}{dt} = P\gamma\Phi_R$ and so

$$\mu_L = \gamma\Phi_R$$

Under normal conditions, *Escherichia coli* (*E. coli*) is expected to maintain an optimal rate of translation, γ_0 .

Considering stationary solutions of these fractions in regards to influx of nutrients and amino acids and balance of the ribosomal fraction can explain several empirically observed growth laws. Here we consider a dynamic growth model that requires the introduction of fractions of the ribosomes engaged in the production of a protein of a particular type, denoted R_x where $x \in \{R, M, E, Q\}$ as for the protein fractions.

$$\frac{dP_x}{dt} = R_x \frac{dP}{dt} = R_x P \mu$$

Then,

$$\begin{aligned} \frac{d\Phi_x}{dt} &= \frac{d(P_x/P)}{dt} \\ &= (R_x - \Phi_x)\mu \end{aligned}$$

In the ramp-up period, the equation most directly relevant to growth is then

$$\frac{d\Phi_R}{dt} = \gamma_0(R_R - \Phi_R)\Phi_R$$

which, assuming R_R has a constant value, has the solution

$$\Phi_R = \frac{R_R}{C e^{-R_R \gamma_0 t} + 1}$$

where C is a constant of integration, $C = (R_R(0) - \Phi_R(0))/\Phi_R(0)$. Thus, the specific growth rate in the ramp-up period is

$$\mu_L = \frac{\gamma_0 R_R}{C e^{-R_R \gamma_0 t} + 1}$$

We note that this is functionally equivalent to the model of Baranyi and Roberts²⁴ without an inhibition term governing how the growth slows down when approaching stationary phase.

However, here we have derived the equation based on known physiological properties that can be tested and with a view to determining levels of burden.

Essentially the specific growth rate approaches a constant value that depends on the translational efficiency and fraction of ribosomes engaged in creating further ribosomes, R_R . The fraction of proteins that are ribosomes or ribosome related, Φ_R , approaches R_R and thus the specific growth rate tends towards a constant value as observed for the exponential phase.

In order to fit the DGM to growth curves derived from well-plate optical density measurements we note that $\mu_L = \frac{d \ln OD}{dt}$ leading to

$$\frac{d \ln OD}{dt} = \frac{\gamma_0 R_R}{C e^{-R_R \gamma_0 t} + 1}$$

with solution

$$OD = C_2 (1 + C^{-1} \exp[R_R \gamma_0 t])$$

where C_2 is a further constant of integration, $C_2 = OD(0)(R_R(0) - \Phi_R(0))/R_R(0)$ for $C \neq 0$. This is then fitted to our growth data (see Methods for further details).

A constant level of ribosomes engaged in making a particular kind of protein in time is likely to be an approximation. However, it has been hypothesised that bacteria respond to sudden changes in the environment with a rapid up-regulation of the relevant genes, potentially with a rapid down-regulation afterwards, in order to effect the most efficient response²⁵⁻²⁷. In the following sections we consider a number of simple responses by fitting to well-plate growth data for bacteria with and without plasmid-borne synthetic constructs expressing green fluorescent protein, GFP, and how the presence of an artificial construct affects the response.

Model Verification – Fitting to Growth Curves

To test our model, we collected growth data from *E. coli* cells carrying plasmids with a construct created from a pool of strong constitutive promoters and ribosomal binding sites (RBSs) that would allow high production of GFP. The primary plasmid used was a medium copy number variant containing the ProD promoter and B0034 RBS. Additional plasmids with medium and high copy number with constructs comprising promoters ProD and ProB²⁸, RBSs B0034, B0035²⁹ and 630K (designed for this study) were used for additional validation.

The data for low optical density values can be quite noisy. The fits can be improved by getting a high quality estimate of the starting OD from the final OD of the previous growth period, which being higher is less susceptible to measurement noise, and the known dilution factor. Using this ‘anchoring’ method is of quite general use with well-plate data and along with our dynamic growth model allows more use of the growth data.

There are many ways in which the number of ribosomes engaged in creating further ribosomes and associated proteins, R_R , could respond to a new period of growth in fresh media. Here we try some of the simplest possibilities showing that they are consistent with our data and indicating how the presence of a construct affects all the possibilities we tried. The simplest possible response is the first we consider and is that R_R is set to a constant value instantly upon dilution into fresh media. A more realistic response is that there is initially a low level for R_R which is increased after a delay for the bacterium to respond to the fresh media and improved conditions for growth. The final response we consider is similar to the second but it has an additional step down of R_R ; this kind of response

has been hypothesised to be the most efficient, rapidly attaining a target value by initially acting to overshoot before correcting.

Flat Response

The first situation we consider is the simplest. We assume that the proportion of ribosomes engaged in creating further ribosomes, R_R , changes instantly upon experiencing a new environment and remains constant during the ramp-up period of growth. Figure 2 shows the result of fitting a flat R_R response (Figure 2a) to growth curves (Figure 2d,g) from bacterial populations, one set with empty vectors and another with a medium copy number plasmid expressing Green Fluorescent Protein, GFP, with a ProD promoter and B0034 Ribosomal Binding Site strength, RBS. It can be seen that the presence of an expressed artificial construct causes the fraction of ribosomes engaged in translating ribosome-associated proteins to be significantly reduced (Figure 2b,f) along with the fraction of proteins that are ribosomes at the beginning of growth (Figure 2c). This is also reflected in the specific growth rate (Figure 2e).

Flat Response Following a Delay

Until now, we have proceeded under the assumption that R_R is a fixed constant throughout the entire ramp-up phase, the simplest possible initial assumption, Figure 2a. However, in reality, cells require time to process signals and adjust their transcription and translation in a new environment, the early part of lag phase. The next simplest assumption is that there is a delay before R_R attains its constant value for the remaining duration of the ramp-up phase, Figure 3a. To properly fit this value, further assumptions on the initial condition must be made: After long periods of growth in an environment, Φ_R and R_R are assumed to reach equilibrium. Figure 3 shows the results of fitting a step function for R_R activation during the early ramp-up phase of growth (Figure 3f), which produces an improved fit, 1/least square value 7.3e3 and 8.5e3 compared with 6.4e3 and 7.7e3 previously for the sets with and without the construct, respectively. We see that the presence of an engineered construct, identical to that described in the previous section, causes a longer delay, t_s , (Figure 3d,e) before the increase in R_R and a lower maximum R_R level (Figure 3b,e). The latter being consistent with what was observed previously when assuming no delay in R_R attaining its constant value. Also shown (Figure 3c) is the fit $R_R(0)$ parameter indicating that the initial levels of ribosomes engaged in making more ribosomes are lower in the case of a construct, further contributing to the overall levels of burden. In order to verify that these results are not unique to a single construct, we repeated the fitting with 3 further constructs, Supporting Information Figure 1. In all four constructs tested the maximum R_R level is significantly lower in the presence of the construct; for three of the four tested, t_s is significantly longer in the presence of a construct, highlighting that the trend remains the same but can manifest differently for different constructs.

Bang-Bang Response Following a Delay

Recently, it has been hypothesised that the most efficient way to respond to an environmental change is via a bang-bang response²⁵⁻²⁷ whereby the organism initially overcompensates before relaxing to a lower level in order to effect a rapid change (Figure 4a). Taking additional data, in the form of two levels of dilution from the same sample when transferring to fresh media (Figure 4c top and bottom curves), requires us to fit a 'step-up step-down' function for R_R which represents a bang-bang response with a delay. Figure 4 shows the results of fitting this response to a data series for which the sample was diluted 25 times into one set of fresh media and a further 49 times into another set of fresh media (Figure 4c). Due to the higher-dilution starting from an OD that is below detectable levels, the step-up part of the fit uses only the lower-dilution data. The growth of the

higher-dilution data is assumed to follow that of the lower-dilution data from the same population to the point that it exits the ramp-up phase, beyond that point the DGM is fitted to the higher-dilution data only. We observe that after step-down the levels of R_R between populations with and without a construct are not significantly different (Figure 4b,d), in agreement with previous observations that construct carrying populations approach a similar maximum growth rates to those without constructs after sufficiently long in exponential phase¹⁵. This further highlights the need to consider the early period of growth when quantifying the fitness burden imposed by artificial constructs.

Behaviour Linked Before and After Environmental Change (Dilution into Fresh Media)

Thus far, the models presented are all assumed to be valid only for the initial ramp-up period of growth where the availability of ribosomes is the limiting factor. A full treatment of the growth curves including stationary phase and approach to it is beyond the scope of this work. However, it is possible to extract some information from our data about what happens as growth slows down. As detailed above, the initial levels of ribosomes available are markedly reduced at the start of a growth period due to the presence of an engineered construct (Figure 2c). However, when one relates $\Phi_R(0)$ back to the growth history of individual samples, a further relationship emerges (Figure 5). To arrive at this relationship a number of steps were taken. It is common to measure at fewer time points for the earliest period of growth which is typically used to get all samples to similar population densities prior to dilution into fresh media and a more intensive measurement regime. Thus, to measure growth in the first period we define a time to initiation, t_d , as the point at which a sample hits $OD_{600} = 0.2$, using linear interpolation to provide data between measurement points (Figure 5a). The correlations between this and the readings at 2 and 4 hours in the growth following dilution to fresh media (Figure 5b) are shown in Figure 5 c,d. The longer a sample takes to initiate growth in the first period the greater its density in the second period at a fixed point shortly after growth. Similar correlations were calculated for multiple time points, 0 to 5 hours in increments of 0.1 hours, in the second period of growth and used to extrapolate growth curves associated with t_d values at 0.5 hour intervals from 4.5 to 9.5 hours (see Methods for details). The DGM was fit to these extrapolated growth curves with t_s and R_R fixed at their respective mean values as measured for the corresponding data sets (Figure 5e). These fits were used to extract the initial ribosome levels, $\Phi_R(0)$, for the different t_d values, Figure 5f. It is known that the growth state of a population before transfer to a new environment or fresh media can affect the subsequent growth³⁰. Here we observe that not only does earlier initiation of growth in a previous period cause a lower level of starting ribosomes but the addition of a construct enhances this effect (Figure 5f). To test that this effect and the others observed in this study were not due to loss of the construct due to negative evolutionary pressure, we analysed the GFP signal from all the used growth curves, Supporting Information Figure 2. We observe that there is no significant loss of construct.

Discussion

There are multiple potential effects on the fitness of an organism when an exogenous construct is introduced. These include toxicity of the construct's product, interference of the product with normal metabolism and the creation of the product overutilising an organism's carefully balanced resources. Even when focussing on only the final aspect, burden, it is clear that there is still much to learn about how the effect manifests itself. As burden due to exogenous protein has previously been observed to reduce after several generations of exponential growth¹⁵, an obvious way to limit burden would be to maintain bacteria in exponential growth indefinitely, perhaps through the use of

a chemostat system. However, if a bacterium carrying a synthetic construct with the intent of performing a useful function is expected to reside in a natural or less artificial environment, it is likely that it will have to endure imperfect and fluctuating conditions. In addition, it is often easier to test the behaviour of synthetic constructs in high-throughput well-plate systems which necessitates multiple rounds of growth and dilution into fresh media. It therefore becomes important to understand the burden induced by a synthetic construct in varying and imperfect environments for two key reasons. One is to reduce the effect of burden and extend the lifetime of a construct in the environment in which it is intended to be deployed. The other is to better predict the long-term behaviour of burden in the design and early testing phases of a construct or circuit of constructs. Another way to limit burden is to supply the cell with orthogonal ribosomes that have specificity for engineered circuits^{31, 32}. For this and other approaches for reducing burden, simple models such as the one presented here will be useful for elucidation of the degree of alleviation of burden.

Study of burden induced by exogenous protein has previously often centred on maximum growth rates as a measure of fitness without redress to the remainder of the measured growth curves or fitting of growth curves to a model. Our research shows that it is vital to consider a greater portion of the growth curve and also to develop models that relate the form of the growth curve to physiological parameters that can aid in the explanation of how burden manifests during growth. In order to achieve this, it becomes necessary to consider more dynamical models of growth that are capable of capturing the effects of burden in nonmaximal growth phases. Our model allows for quantification of burden purely from growth curves which may be more applicable in areas where more involved quantification³³ is too time-consuming or costly.

Intriguingly, the memory effects whereby growth in fresh media correlates with previous growth is affected by the presence of a synthetic construct. Our data indicate that this is due to the level to which the ribosomes have been reduced as growth slows down in the later phases of the growth curve. Thus, further investigation of the latter parts of the growth curve, where the ribosomes can no longer be assumed to be the limiting factor for growth, will also be essential to properly understand the effects of burden. Moreover, even without a deeper understanding of the later phases of growth, it is important that differences in previous growth environments be taken into account when trying to measure burden especially when using well-plate systems for early evaluation of constructs. Over longer time scales, the effect of potential adaptation of the host to a construct must also be taken into account. Here, our experiments were performed in time scales where such adaptation is unlikely.

In addition to an increased understanding of the entire growth curve, there are multiple ways in which the findings of this study could be extended. The amount of protein produced and nutrients available could be measured and compared with the predicted ribosomal protein fraction to gain a greater understanding of the different protein fractions and how they change over time. The model could also be compared with more complex models of resource allocation²² to find the best compromise between detail and simplicity for quantifying fitness cost in synthetic biology design. With larger data sets it also becomes possible to apply improved statistical methods to differentiate growth curves³⁴. Finally, as more data becomes available it will be worth pursuing the investigation of a more-detailed functional form for the way in which an organism adapts to a new environment in terms of the fraction of ribosomes engaged in translating further ribosomes and associated proteins than the rapid change between fixed levels that we have hypothesised here. This could allow for burden to be separated from the effects of other fitness costs in further studies.

Conclusions

In this work, we propose a new dynamic growth model that treats the early phases of growth. This model is based on physiological considerations of the amount of ribosomes available for growth and how their levels change over time. This model is ideal for studying the burden on fitness that manifests when an exogenous synthetic construct is introduced into bacteria.

A large part of the burden due to the synthetic construct we tested appeared during the lag phase when growth is increased from low levels after a period of stationary growth and transfer to fresh media. Bacteria with constructs were observed to start more slowly and achieve a lower maximum fraction of ribosomes available for growth. Thus, our model led us to a novel measure of burden that includes its effect on the early stages of growth as the bacteria ramp up their levels of ribosomes. We tested several simple responses for ribosome production in a new environment: an instantaneous flat response, a flat response after a delay and a bang-bang response after a delay. All showed a significant effect due to the presence of an engineered construct in a way that is relatable back to the physiology of the cell, thus providing an avenue for future investigation.

A further part of the effect of a synthetic construct appears to be due to the rate at which ribosomes are reduced as an organism approaches stationary phase after a period of faster growth. The rate of loss appears greater with a construct and this leads to a slower ramp up of growth in a following growth period. Thus burden also manifests in the memory effect in which growth in a new environment is affected by the growth history in a previous environment. We cannot rule out that this is in part due to changes in translational efficiency but recent findings support the idea that ribosome reduction is a major player in this process ³⁵.

It is clear that much work is needed to fully elucidate the effect of an exogenous synthetic construct, or multiple such constructs, on the growth of an organism, even when only considering burden due to nontoxic but unnecessary protein. By moving to simple models of growth that relate protein and RNA fractions dynamically to growth we have moved towards a better understanding of this burden on the fitness. With further research, this will hopefully lead to a greater control over burden, especially during the design and early-testing phases and, ultimately, longer lived constructs that can perform functions in realistic environments.

Methods

Bacterial Strains, Growth and Transformation

E. coli strain DH5 α (Zymo Research) was used for all transformations and growth experiments. Experiments were performed using cloud-accessible robot facilities at Transcriptic Inc., Menlo Park, CA, USA.

Competent cells were transformed with 10ng of plasmid in 96-well plates and after recovery in SOC media, cells were plated on LB agar plates supplemented with ampicillin at a final concentration of 100ng/mL and the plates incubated for 16 hours at 37° C.

Growth Assays to Determine Relationship Between Pre- and Postdilution Behaviour (Figure 5)

A single colony of each transformant was used to inoculate 240 μ L of LB media supplemented with 100ng/mL of ampicillin in a 96-well plate. Plates were incubated for 24 hours at 37° C with vigorous shaking. Optical density (measured at 600 nm) readings and fluorescence readings (plate reader, excitation 475nm and emission 509nm) were taken every 2 hours. From this overnight culture, 5 μ L from each well was used to inoculate 230 μ L of fresh, prewarmed LB supplemented with ampicillin (i.e. a 47 fold dilution), in a new 96-well plate. These cultures were then incubated at 37° C with shaking for 12 hours with optical density and fluorescence measurements taken every 6 minutes.

Growth assays to validate the dynamic growth model (Figures 2, 3 and 4)

Transformations were performed as described for Figure 5. The initial culture was incubated for 26 hours at 37° C with vigorous shaking with optical density and fluorescence measurements taken every 6 minutes. Subsequently, 10 µL of this culture was used to inoculate 240 µL of prewarmed LB supplemented with ampicillin (i.e. a 25 fold dilution). Immediately after mixing, 5 µL from each of the new cultures was removed and used to inoculate another 240 µL of prewarmed LB supplemented with ampicillin on the same 96-well plate, resulting in a total growth volume was 245 µL for each sample and a1225 fold dilution of the initial preculture. These samples were subsequently incubated for 18 hours at 37° C with vigorous shaking, and optical density and fluorescence measurements taken every 6 minutes.

Key Sequences

B0034:
AAAGAGGAGAAA

B0035:
ATTAAAGAGGAGAA

ProD:
ttctagagCACAGCTAACACCGTCGTCCTATCTGCTGCCCTAGGTCTATGAGTGGTTGCTGGATAACTTTAC
GGGCATGCATAAGGCTCGTATAATATATTCAGGGAGACCACAACGGTTTCCCTCTACAAATAATTTTGTAA
CTTTtactagag

sfGFP:
ATGAGCAAAGGAGAAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGC
ACAAATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCTTAAATTTATTTGCAC
TACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGT
TATCCGGATCACATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACT
ATATCTTTCAAAGATGACGGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAAT
CGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAACTCGAGTACAACCTTA
ACTCACACAATGTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAA
CGTTGAAGATGGTTCCGTTCACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCT
TTTACCAGACAACCATTACCTGTCGACACAATCTGTCTTTTGGAAAGATCCCAACGAAAAGCGTGACCACAT
GGTCCTTCTTGAGTTTGTAAGTGTGCTGGGATTACACATGGCATGGATGAGCTCTACAAA

For full sequences, see Supporting Information.

Extrapolation

To generate the extrapolated growth curves for determination of the effect on growth following dilution into fresh media by previous growth, we used the following procedure. First, the time to intercept $OD_{600} = 0.2$, t_d , was obtained for each sample that grew in the first period of growth, 22 with a construct and 19 without a construct. Due to the relatively large spacing for measurements in this period, linear interpolation was used to improve the time resolution. The OD_{600} value for the corresponding sample in the second period of growth was obtained at even time intervals from 0.1 to 5 hours in intervals of 0.1 hours. Linear interpolation was also utilised here but the measurement intervals were much tighter. Scatter plots showing t_d against the OD_{600} values at a given time point in the second period of growth (Figure 5 c,d) indicate a linear relationship between the two. Thus, linear regression was used to extrapolate the relationship between t_d and the level of growth achieved at evenly spaced time points in the second period of growth. Thus, for each t_d a growth curve for the second period of growth could be made for a fixed t_d . The DGM was fit to these growth

curves, with fixed t_s and R_R values, in order to obtain the starting ribosome levels in the second period of growth, $\Phi_R(0)$, for given values of the time to initiate growth in the first period, Figure 5f.

Data Fitting

Data for the initial growth phase is fitted up to the time point of the smoothed maximum growth rate. The smoothed maximum growth rate is calculated by using linear fitting with a length of five time points over the cell density data. The derivative of the smoothed density data divided by the smoothed density data is taken as the growth rate.

The data fitted to models is the time span from the first measurements up until the smoothed maximum growth rate.

Least square fitting was used to fit the models to the data. Fitting was made with the Levenberg-Marquardt algorithm, from the `curve_fit` function in the Python Scipy library. For testing of statistical significance Welch's t-test was used.

Supporting Information

Two supplementary figures, plasmid maps and sequences of our constructs.

Author Contributions

A.N., A.P. and A.A. designed the research; A.N. designed and performed experiments, modelling and computational work; all authors analysed data; all authors wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgements

We thank George Wadhams, Wei Huang, Alessandro Abate and Judith Armitage for critical discussions and support. This work was supported by a Royal Society University Research Fellowship (UF120327 to A.A.) and an Engineering and Physical Sciences Research Council Fellowship (EP/M002454/1 to A.P.).

References

- [1] Purnick, P. E. M., and Weiss, R. (2009) The second wave of synthetic biology: from modules to systems, *Nat Rev Mol Cell Biol* 10, 410-422.
- [2] Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A., and Weiss, R. (2014) Realizing the potential of synthetic biology, *Nat Rev Mol Cell Biol* 15, 289-294.
- [3] Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014) A brief history of synthetic biology, *Nat Rev Micro* 12, 381-390.
- [4] You, L., Cox Iii, R. S., Weiss, R., and Arnold, F. H. (2004) Programmed population control by cell-cell communication and regulated killing, *Nature* 428, 868.

- [5] Balagaddé, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake, S. R., and You, L. (2008) A synthetic *Escherichia coli* predator–prey ecosystem, *Molecular systems biology* 4.
- [6] Withers, S. T., Gottlieb, S. S., Lieu, B., Newman, J. D., and Keasling, J. D. (2007) Identification of Isopentenol Biosynthetic Genes from *Bacillus subtilis* by a Screening Method Based on Isoprenoid Precursor Toxicity, *Applied and environmental microbiology* 73, 6277-6283.
- [7] Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids, *Nature biotechnology* 21, 796.
- [8] Kizer, L., Pitera, D. J., Pfleger, B. F., and Keasling, J. D. (2008) Application of Functional Genomics to Pathway Optimization for Increased Isoprenoid Production, *Applied and environmental microbiology* 74, 3229-3241.
- [9] Price, M. N., Wetmore, K. M., Deutschbauer, A. M., and Arkin, A. P. (2016) A Comparison of the Costs and Benefits of Bacterial Gene Expression, *PLoS One* 11, e0164314.
- [10] Sleight, S. C., and Sauro, H. M. (2013) Visualization of Evolutionary Stability Dynamics and Competitive Fitness of *Escherichia coli* Engineered with Randomized Multigene Circuits, *ACS synthetic biology* 2, 519-528.
- [11] Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices, *Nature biotechnology* 26, 787-793.
- [12] Sleight, S. C., Bartley, B. A., Lieviant, J. A., and Sauro, H. M. (2010) Designing and engineering evolutionary robust genetic circuits, *Journal of Biological Engineering* 4, 12.
- [13] Renda, B. A., Hammerling, M. J., and Barrick, J. E. (2014) Engineering reduced evolutionary potential for synthetic biology, *Molecular bioSystems* 10, 1668-1678.
- [14] Jack, B. R., Leonard, S. P., Mishler, D. M., Renda, B. A., Leon, D., Suárez, G. A., and Barrick, J. E. (2015) Predicting the Genetic Stability of Engineered DNA Sequences with the EFM Calculator, *ACS Synthetic Biology* 4, 939-943.
- [15] Shachrai, I., Zaslaver, A., Alon, U., and Dekel, E. (2010) Cost of unneeded proteins in *E. coli* is reduced after several generations in exponential growth, *Mol Cell* 38, 758-767.
- [16] Zwietering, M., Jongenburger, I., Rombouts, F., and van 't Riet, K. (1990) Modeling of the bacterial growth curve, *Appl Environ Microb* 56.
- [17] Peleg, M., and Corradini, M. G. (2011) Microbial growth curves: what the models tell us and what they cannot, *Critical reviews in food science and nutrition* 51, 917-945.
- [18] Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010) Interdependence of cell growth and gene expression: origins and consequences, *Science* 330, 1099-1102.
- [19] Klumpp, S., and Hwa, T. (2014) Bacterial growth: global effects on gene expression, growth feedback and proteome partition, *Current opinion in biotechnology* 28, 96-102.
- [20] Scott, M., Klumpp, S., Mateescu, E. M., and Hwa, T. (2014) Emergence of robust growth laws from optimal regulation of ribosome synthesis, *Molecular systems biology* 10, 747.
- [21] Maitra, A., and Dill, K. A. (2015) Bacterial growth laws reflect the evolutionary importance of energy efficiency, *Proceedings of the National Academy of Sciences of the United States of America* 112, 406-411.
- [22] Weisse, A. Y., Oyarzun, D. A., Danos, V., and Swain, P. S. (2015) Mechanistic links between cellular trade-offs, gene expression, and growth, *Proceedings of the National Academy of Sciences of the United States of America* 112, E1038-1047.
- [23] de Jong, H., Geiselmann, J., and Ropers, D. (2017) Resource Reallocation in Bacteria by Reengineering the Gene Expression Machinery, *Trends in microbiology* 25, 480-493.
- [24] Baranyi, J., and Roberts, T. A. (1994) A dynamic approach to predicting bacterial growth in food, *International journal of food microbiology* 23, 277-294.
- [25] Giordano, N., Mairet, F., Gouze, J. L., Geiselmann, J., and de Jong, H. (2016) Dynamical Allocation of Cellular Resources as an Optimal Control Problem: Novel Insights into Microbial Growth Strategies, *PLoS Comput Biol* 12, e1004802.

- [26] Pavlov, M. Y., and Ehrenberg, M. (2013) Optimal control of gene expression for fast proteome adaptation to environmental change, *Proceedings of the National Academy of Sciences of the United States of America* 110, 20527-20532.
- [27] Madar, D., Dekel, E., Bren, A., Zimmer, A., Porat, Z., and Alon, U. (2013) Promoter activity dynamics in the lag phase of *Escherichia coli*, *BMC Syst Biol* 7, 136.
- [28] Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011) Design, construction and characterization of a set of insulated bacterial promoters, *Nucleic Acids Research* 39, 1131-1141.
- [29] Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K., Gao, Y., Arkin, A. P., Endy, D., and Church, G. M. (2013) Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*, *Proceedings of the National Academy of Sciences* 110, 14024-14029.
- [30] Joers, A., and Tenson, T. (2016) Growth resumption from stationary phase reveals memory in *Escherichia coli* cultures, *Scientific reports* 6, 24055.
- [31] Darlington, A. P. S., Kim, J., Jiménez, J. I., and Bates, D. G. (2018) Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes, *Nature communications* 9, 695.
- [32] Darlington, A. P. S., Kim, J., Jimenez, J. I., and Bates, D. (2018) Engineering Translational Resource Allocation Controllers: Mechanistic Models, Design Guidelines, and Potential Biological Implementations, *bioRxiv*.
- [33] Ceroni, F., Algar, R., Stan, G. B., and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden, *Nature methods* 12, 415-418.
- [34] Rickett, L. M., Pullen, N., Hartley, M., Zipfel, C., Kamoun, S., Baranyi, J., and Morris, R. J. (2015) Incorporating prior knowledge improves detection of differences in bacterial growth rate, *BMC Systems Biology* 9, 60.
- [35] Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y. P., and Hwa, T. (2016) Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth, *Nature microbiology* 2, 16231.

Figures

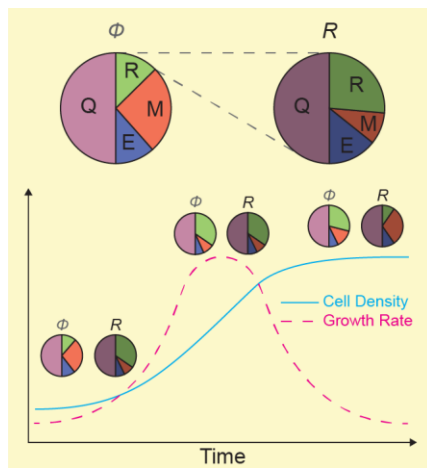


Figure 1 – Schematic of protein fractions, Φ_x , engaged ribosome fractions, R_x , top panels, and how they change during growth, bottom panel. The subscript x runs over the different classifications of protein: R denotes ribosome related proteins, M denotes metabolic related proteins, E denotes engineered proteins from an artificial construct and Q denotes the remaining fraction which is not related to growth and assumed to remain constant.

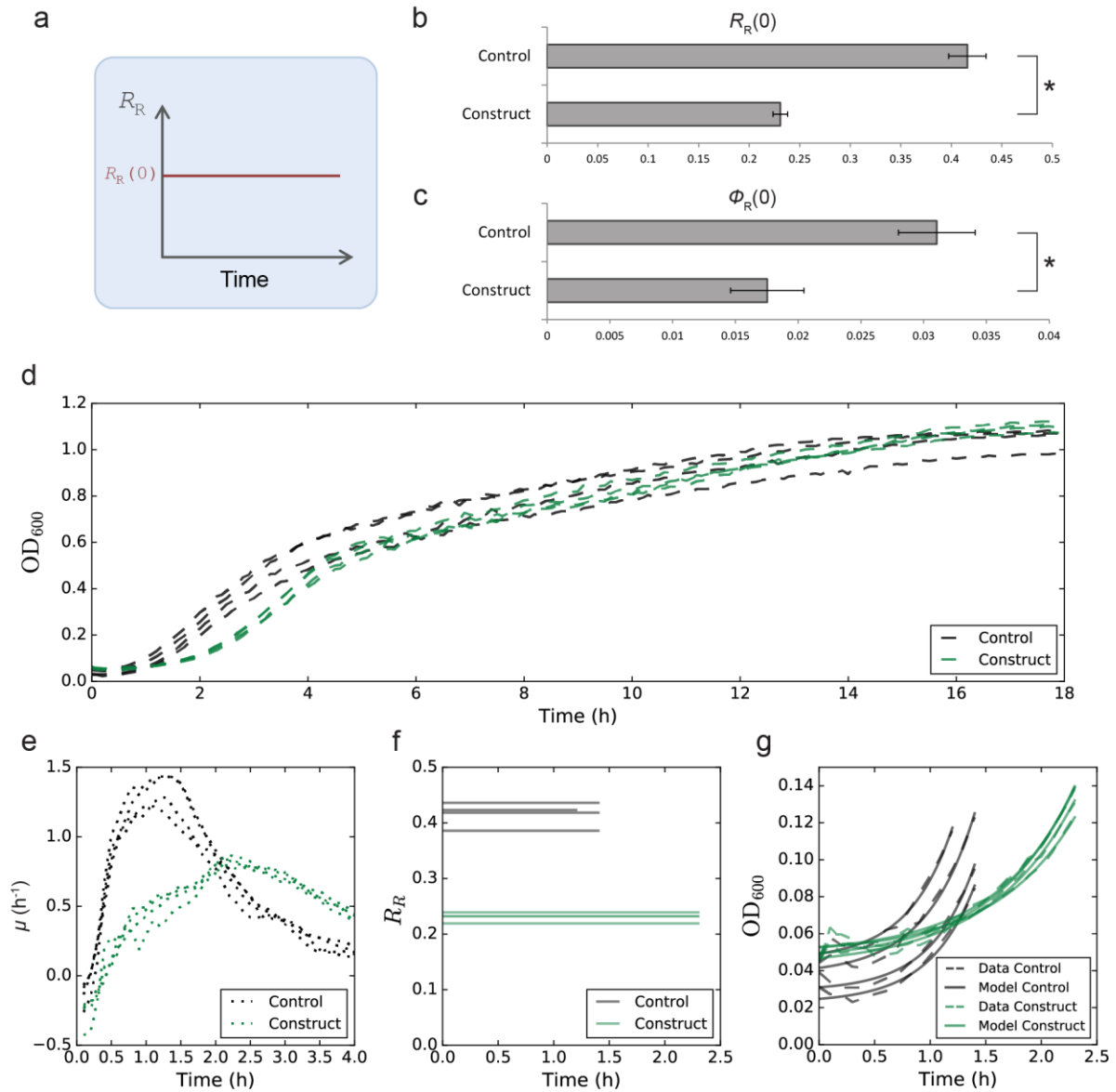


Figure 2 – Fits of the dynamic growth model to growth data in the case of a constant and instantaneous response of the fraction of ribosomes engaged in making further ribosomes to dilution to fresh media following pregrowth. (a) Schematic of the flat response showing the fraction of ribosomes engaged in translating ribosome-associated proteins, R_R , over time. (b,c) Mean of fits of R_R and Φ_R , the fraction of ribosome-associated proteins, respectively, p-values $1.04e-04$ and $1.47e-03$. (d) Growth curves for populations. (e) Specific growth rate, μ , for populations. (f) Fit values of R_R for populations over time. (g) Growth curves (dashed lines) and model fits (solid lines) for populations. Green lines are populations carrying a construct (medium-copy ProD-B0034 expressing GFP) and black lines are populations without the construct (control).

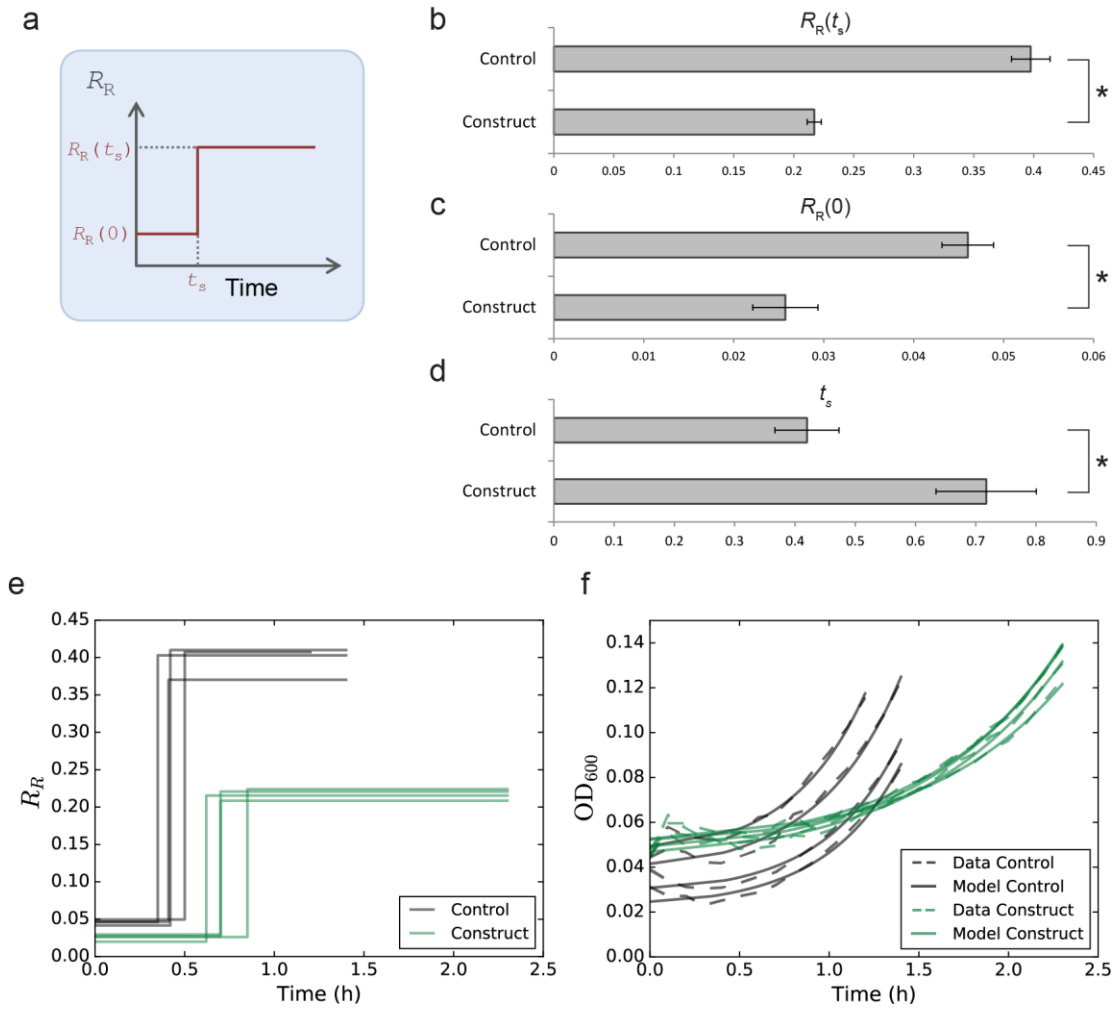


Figure 3 – Fits of the dynamic growth model to growth data in the case of a step response of the fraction of ribosomes engaged in making further ribosomes to dilution to fresh media following pregrowth. (a) Schematic of the step response showing the fraction of ribosomes engaged in translating ribosome-associated proteins, R_R . (b,c,d) Mean of fits of $R_R(t_s)$, the fraction of ribosomes engaged in making further ribosomes after the step up, $R_R(0)$, the fraction of ribosomes engaged in making further ribosomes at the time of dilution, and t_s , the time until the step of the response post dilution, respectively; p-values $7.62e-05$, $3.43e-04$ and $3.21e-03$. (e) Fit values of R_R over time. (f) Growth curves (dashed lines) and model fits (solid lines) for populations. Green lines are populations carrying a construct (medium-copy ProD-B0034 expressing GFP) and black lines are populations without the construct (control).

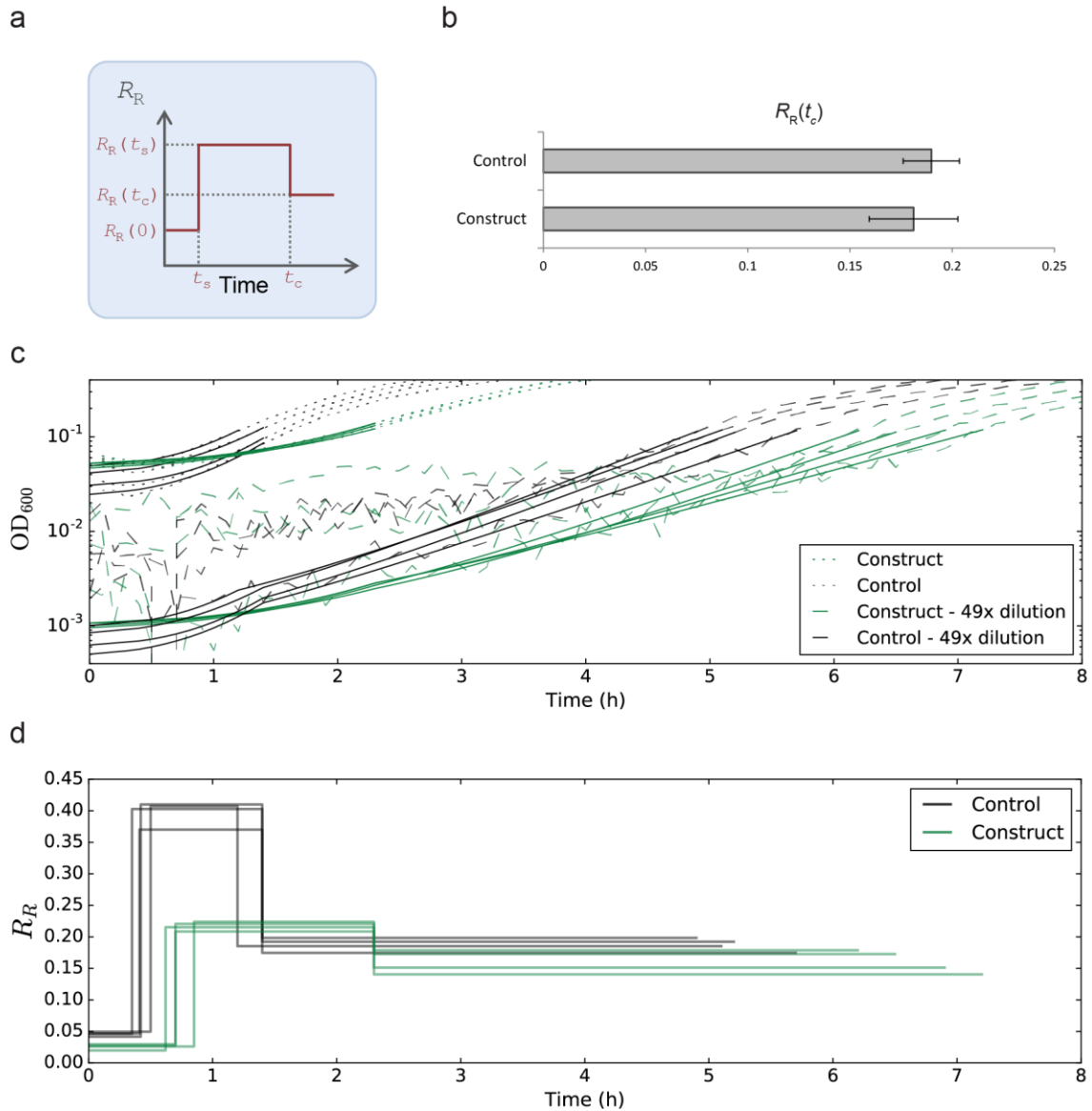


Figure 4 – Fits of the dynamic growth model to growth data in the case of a step up followed by a step down (bang-bang with delay) response of the fraction of ribosomes engaged in making further ribosomes to dilution to fresh media following pregrowth. (a) Schematic of the response showing the fraction of ribosomes engaged in translating ribosome-associated proteins, R_R over time. (b) Mean of fits of $R_R(t_c)$, the fraction of ribosome-associated proteins following a step down in the later stages of the response, p -value > 0.05. (c) Growth curves (dashed lines) and model fits (solid lines) for populations with dilutions of 25 times and 1225 times from the previous period of growth. (d) Fit values of R_R for populations over time. Green lines are populations carrying a construct (medium-copy ProD-B0034 expressing GFP) and black lines are populations without the construct (control).

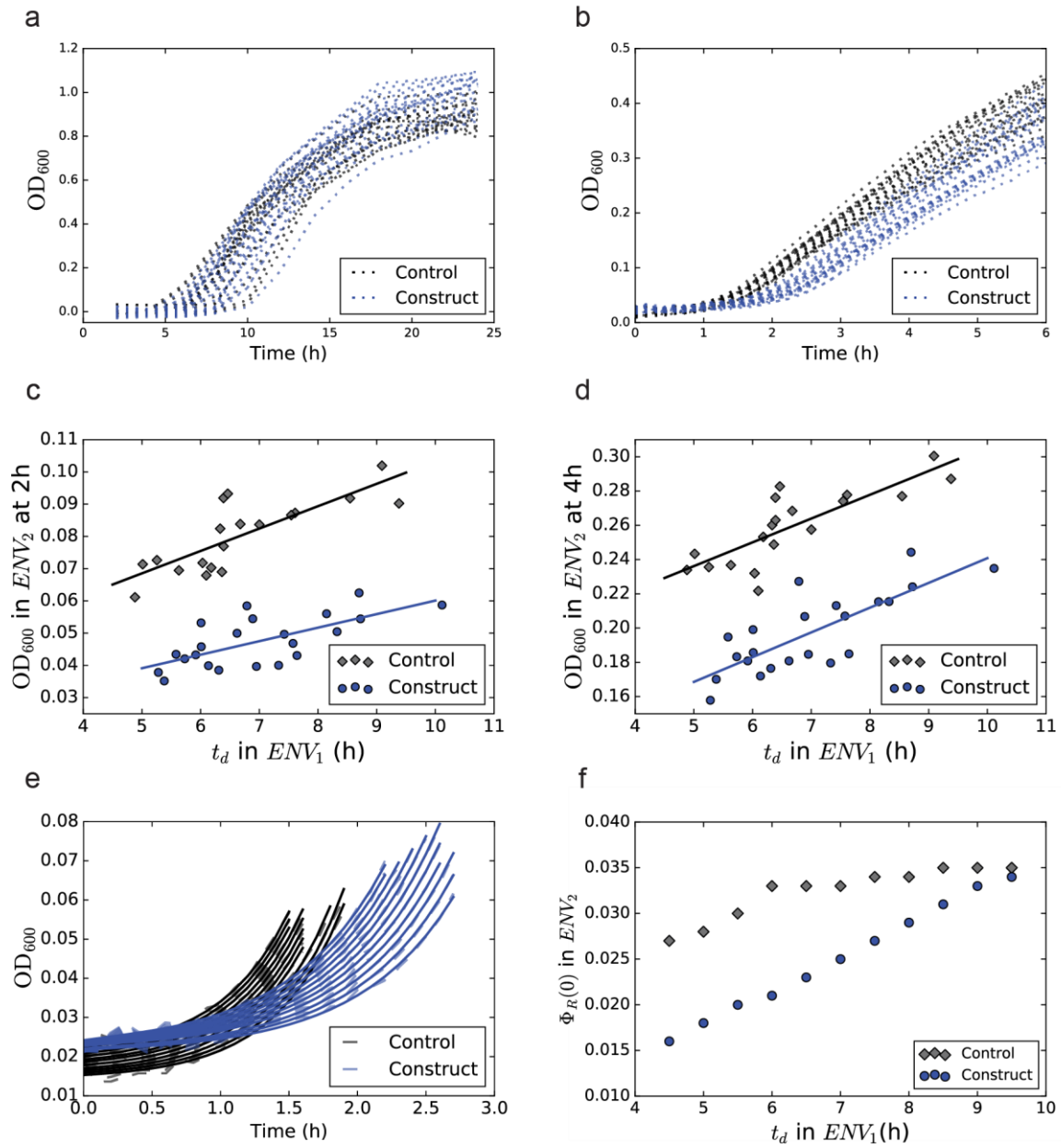
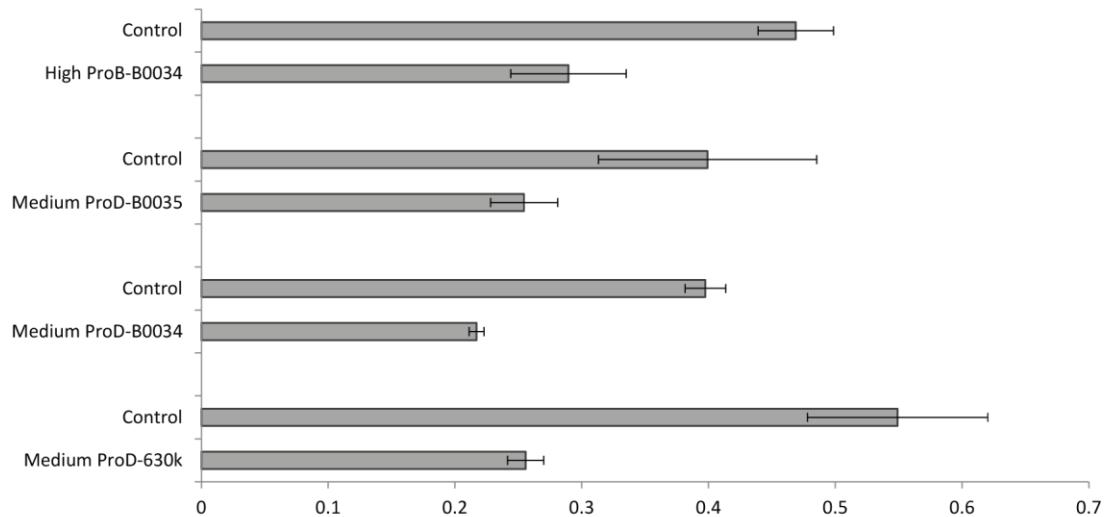


Figure 5 – Relationship of growth following dilution into fresh media to growth prior to dilution showing that constructs can increase the lower initial levels of ribosomes available for growth due to earlier initiation in a previous period of growth. (a) Growth curves for populations during the first period of growth (ENV₁). (b) Matching growth curves for the second period of growth (ENV₂). (c,d) Scatter plots of the measured time-to-initiation, t_d , in the first period of growth and the OD attained in the second period of growth after growth of 2 and 4 hours respectively (e) Fit (solid lines) of the DGM to extrapolated growth curves (dashed lines) for populations with and without a construct. (f) Initial ribosome fractions, $\Phi_R(0)$, at the start of growth in the second period of growth against t_d from the first period of growth. Black lines indicate populations without a construct and blue lines indicate populations with a medium-copy plasmid with ProD-B0035 expressing GFP.

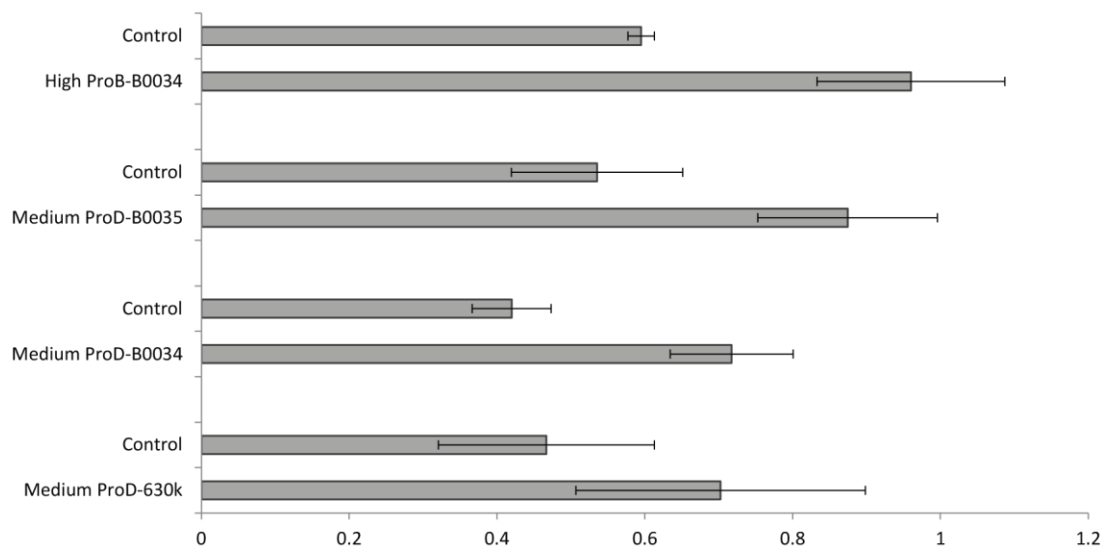
Supporting Information for: A Dynamic Model of Resource Allocation in Response to the Presence of a Synthetic Construct

Supplementary Figures

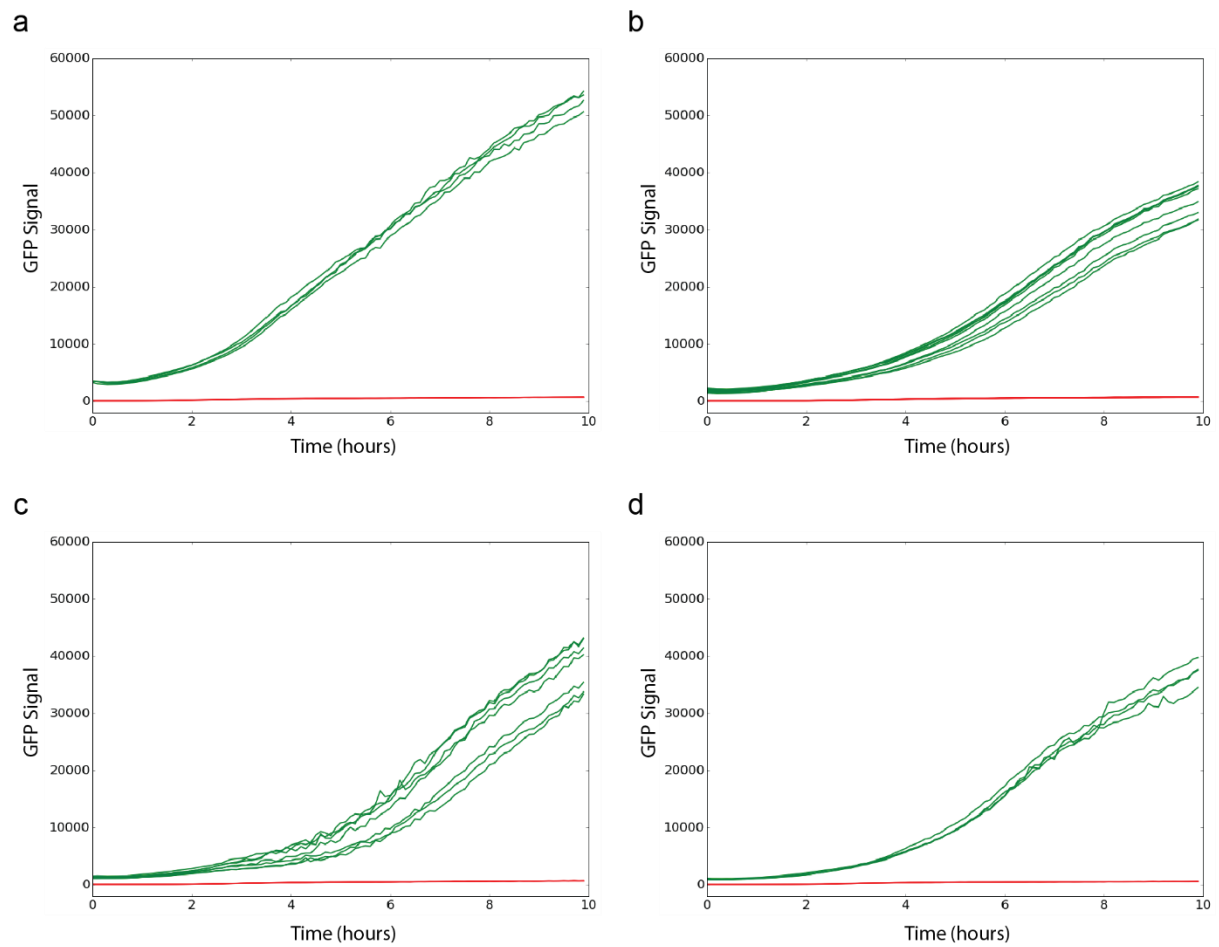
a



b



Supporting Information Figure S1 – Fits of the dynamic growth model to growth data in the case of step response of the fraction of ribosomes engaged in making further ribosomes to dilution to fresh media following pre-growth. (a) Mean of fits of $R_R(t_s)$, the fraction of ribosomes engaged in making further ribosomes after the step up, p-values 5.35e-05, 3.81e-05, 7.62e-05 and 2.55e-02, respectively. (b) Mean of fits of t_s , the time of the step of the response post dilution, p-values 3.31e-04, 4.25e-05, 3.21e-03 and 1.85e-01, respectively. Constructs used are High Copy ProB-B0034, Medium Copy ProD-B0035, Medium Copy ProD-B0034, and Medium Copy ProD-630K; each is paired with a control from the same experiment

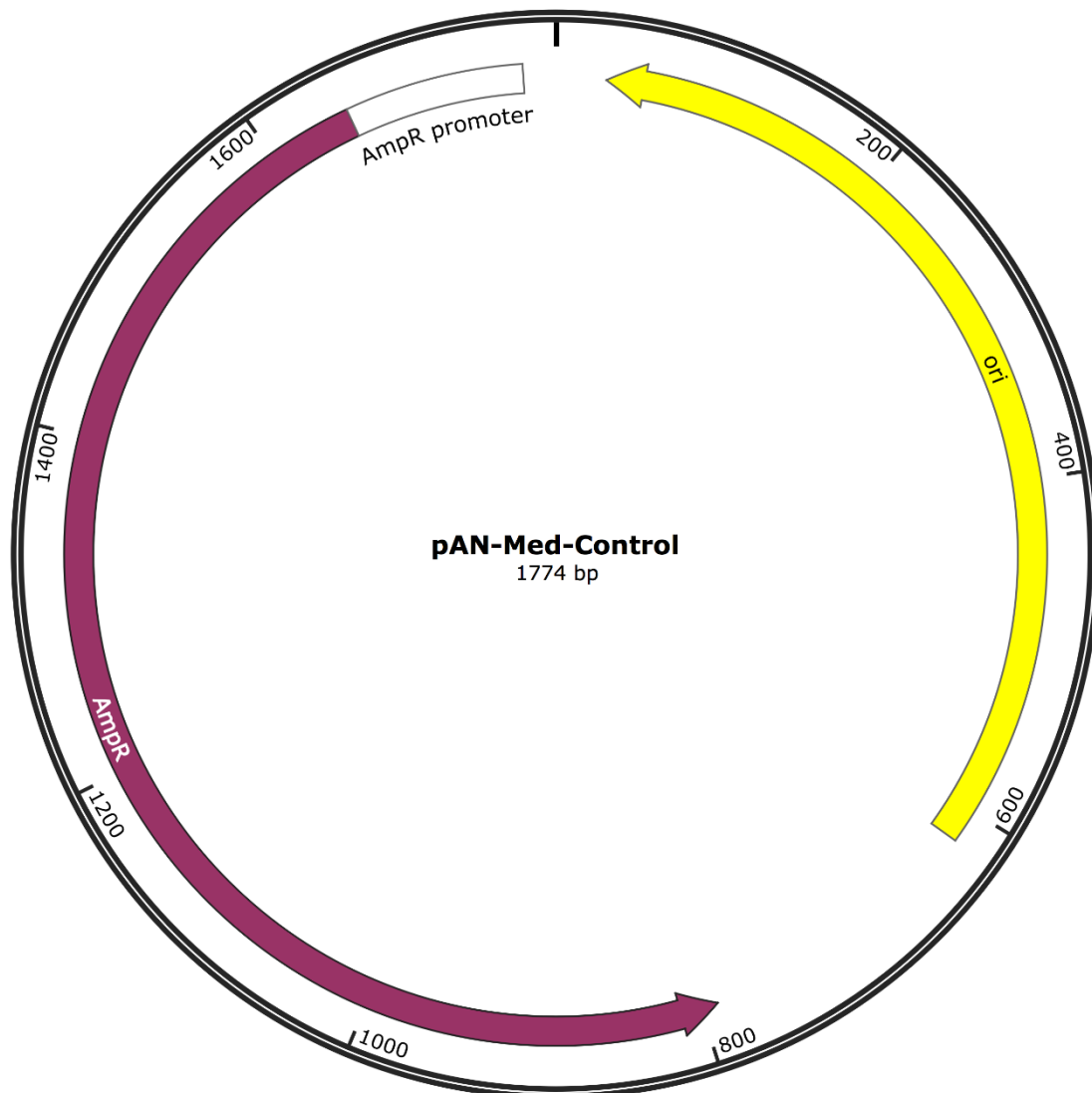


Supporting Information Figure S2 – GFP traces from the four constructs as in Supporting Information Figure S1 showing that the function of the construct is not impaired by mutation throughout the duration of the experiment. (a) GFP signal for Medium Copy ProD-B0034. (b) GFP signal for Medium Copy ProD-B0035. (c) GFP signal for High Copy ProD-B0034. (d) GFP signal for Medium Copy ProD-630K. In all panels green represents GFP from bacteria containing the construct-carrying plasmid and red represents the GFP signal from the negative controls (lacking the construct).

Full Sequences for Plasmids

Negative Control Medium Copy Number

Created with SnapGene®



```
gaaccgtaaaaaggccgcgttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgct
caagtcagaggtggcgaaacccgacaggactataaagataaccaggcggttccccctggaagctccctcgtgcgctct
cctgtccgacctgccgcttacggataacctgtccgcctttctcccttcgggaagcgtaggcgtttctcatagctcacgct
gtaggtatctcagttcgggtgtaggtcgctccaagctgggctgtgtgcacgaaccccccggttcagcccgaccgctg
cgccttaccggtaactatcgcttgagtccaacccggtgaagacacgacttatcgccactggcagcagccactggtaa
caggattagcagagcgaggtatgtaggcggtgctacagagttctgaagtggcctaactacggctacactagaa
ggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca
aaccaccgctggttagcgggtggtttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcc
ttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggttttggtcatgagattatcaaaaag
gatcttcacctagatccttttaattaaaaatgaagttttaaatacaatctaaagtatatatgagtaaacttggtctgacagtt
accaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcggtcatccatagttgctgactccccgctcgtgtag
```

ataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgagagacccacgctcaccggctc
cagatttatcagcaataaaccagccagccggaagggccgagcgagcagaagtggctctgcaactttatccgcctccatc
cagtcctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttggtgccattgctgcag
gcatcgtgggtgcacgctcgtcgtttggatggcttcattcagctccggttccaacgatcaaggcgagttacatgatccc
ccatgttggtgcaaaaaagcggtagctccttcggtcctccgatcgttgcagaagtaagtggccgcagtggttatcactca
tggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgctttctgtgactggtgagtactcaaccaag
tcattctgagaatagtgtatgcggcgaccgagtgctcttgccggcgctcaatacgggataataccgcgccacatagc
agaactttaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatcca
gttcgatgtaaccactcgtgcacccaactgatcttcagcatctttactttcaccagcgttctgggtgagcaaaaacag
gaaggcaaaatgccgcaaaaaaggggaataagggcgacacggaaatgttgaatactcatactcttcttttcaatatt
attgaagcatttatcagggtattgtctcatgagcggatacatattgaatgtatttagaaaaataaacaatataggggtcc
gcgcacatttccccgaaaagtg



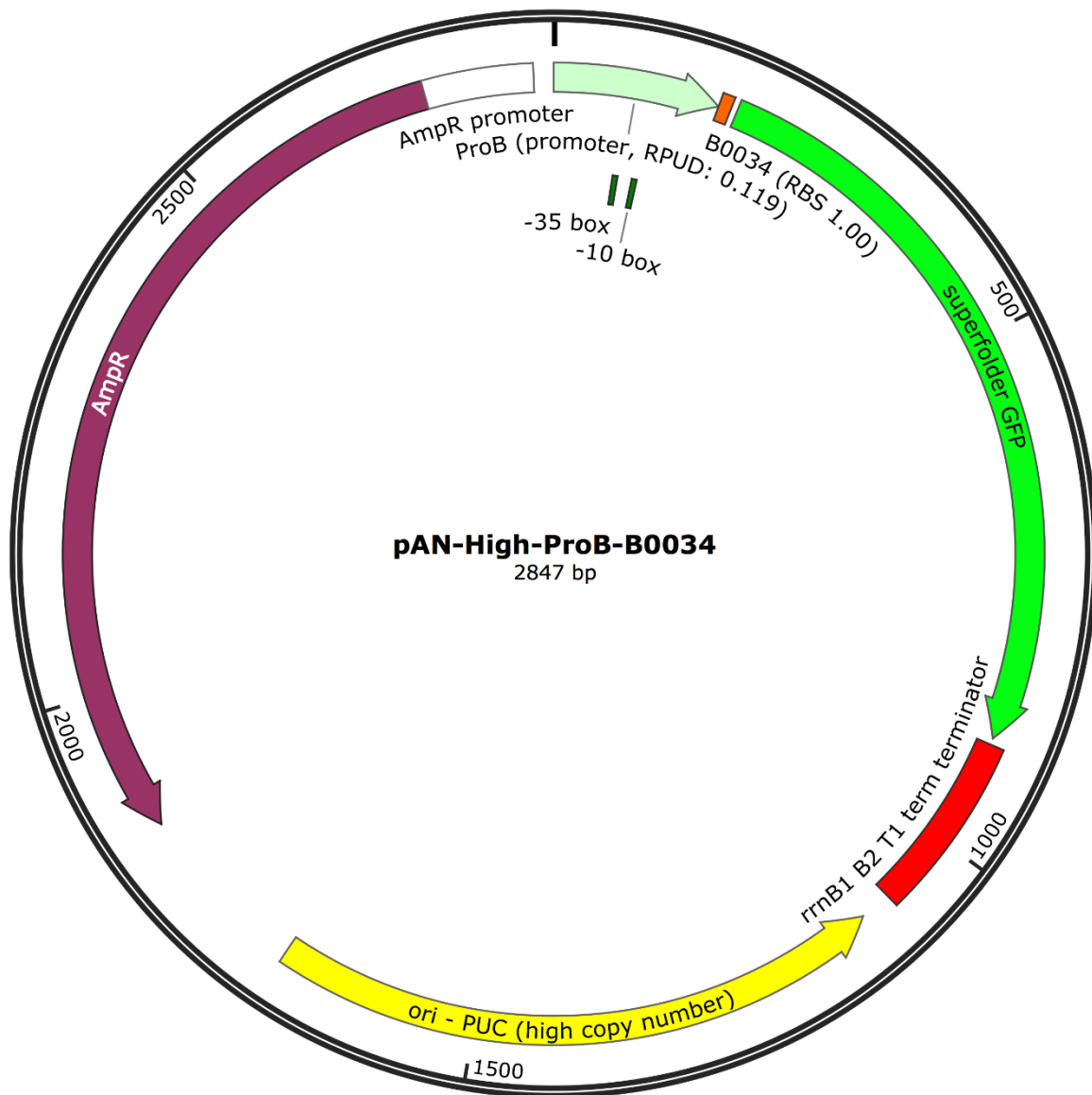
ttctagagCACAGCTAACACCACGTCGTCCTATCTGCTGCCCTAGGTCTATGAGT
GGTTGCTGGATAACTTTACGGGCATGCATAAGGCTCGTATAATATATTCAGGGA
GACCACAACGGTTTCCCTCTACAAATAATTTTGTTTAACTTTtactagagAAAGAGG
AGAAAtactagATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTT
GTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGGAGAGGGT
GAAGGTGATGCTACAAACGGAAACTCACCTTAAATTTATTTGCACTACTGGA
AACTACCTGTTCCATGGCCAACACTTGTCCTACTCTGACCTATGGTGTTCAAT
GCTTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTCAAGAGTGCCA
TGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGACCT
ACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCG
AGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAACTCGA
GTACAACCTTTAACTCACACAATGTATACATCACGGCAGACAAACAAAAGAATGG
AATCAAAGCTAACTTCAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTA

GCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA
GACAACCATTACCTGTCTGACACAATCTGTCCTTTTCGAAAGATCCCAACGAAAAG
CGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGC
ATGGATGAGCTCTACAAAtagtaaaattagccccggcgaaaggccagctcttcgactgagccttctgttt
atttgatgcctggcagttccctactctcgcatggggagtccccacactaccatcggcgctacggcgtttcacttctgagtt
cggcacatggggcaggtgggaccaccgcgctactgccgccaggcaaagaaccgtaaaaaggccgctgtctggcg
ttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaaccgcacag
gactataaagataaccaggcgttccccctggaagctccctcgctcgctctctgttccgaccctgccgcttaccggatac
ctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcg
ctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggttaactatcgctcttgagttc
aaccgcgtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggc
ggtgctacagagttctgaagtgggtgcctaactacggctacactagaaggacagtatttggtatctcgctctgctgaa
gccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaccaccgctggtagcggtggtttttgttg
caagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacggggtctgacgctcagtg
gaacgaaaactcacgttaagggttttggctatgagattatcaaaaaggatcttcacctagatccttttaataaaaaat
gaagtttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatct
agcgatctgtctatttcgttcacatagttgcctgactccccgctgctgtagataactacgatacgggaggggttaccatct
ggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaaccagccagcc
ggaagggccgagcgcagaagtggctcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctaga
gtaagtagttcgccagtttaatagtttgcgcaacgttggtgccattgctgcaggcatcgtggtgtcacgctcgtcgtttggtat
ggcttcattcagctccggttccaacgatcaaggcgagttacatgatcccccattgttgcaaaaaagcggttagctcc
ttcggctctccgatcgttgtcagaagtaagttggccgagtggtatcactcatggttatggcagcactgcataattctcttac
tgtcatgccatccgtaagatgctttctgtgactgggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgacc
gagttgctcttgccccggcgtaataacgggataataaccgcgccacatagcagaactttaaaagtgtcatcattggaaa
acgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgaaccactcgtgcacccaact
gatcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggga
ataagggcgacacggaaaatgttgaatactcatactcttcttttcaatattattgaagcatttatcagggttattgtctcatg
agcggatacatatttgaatgtatttagaaaaataaacaatataggggttccgcgcacatttccccgaaaagtg



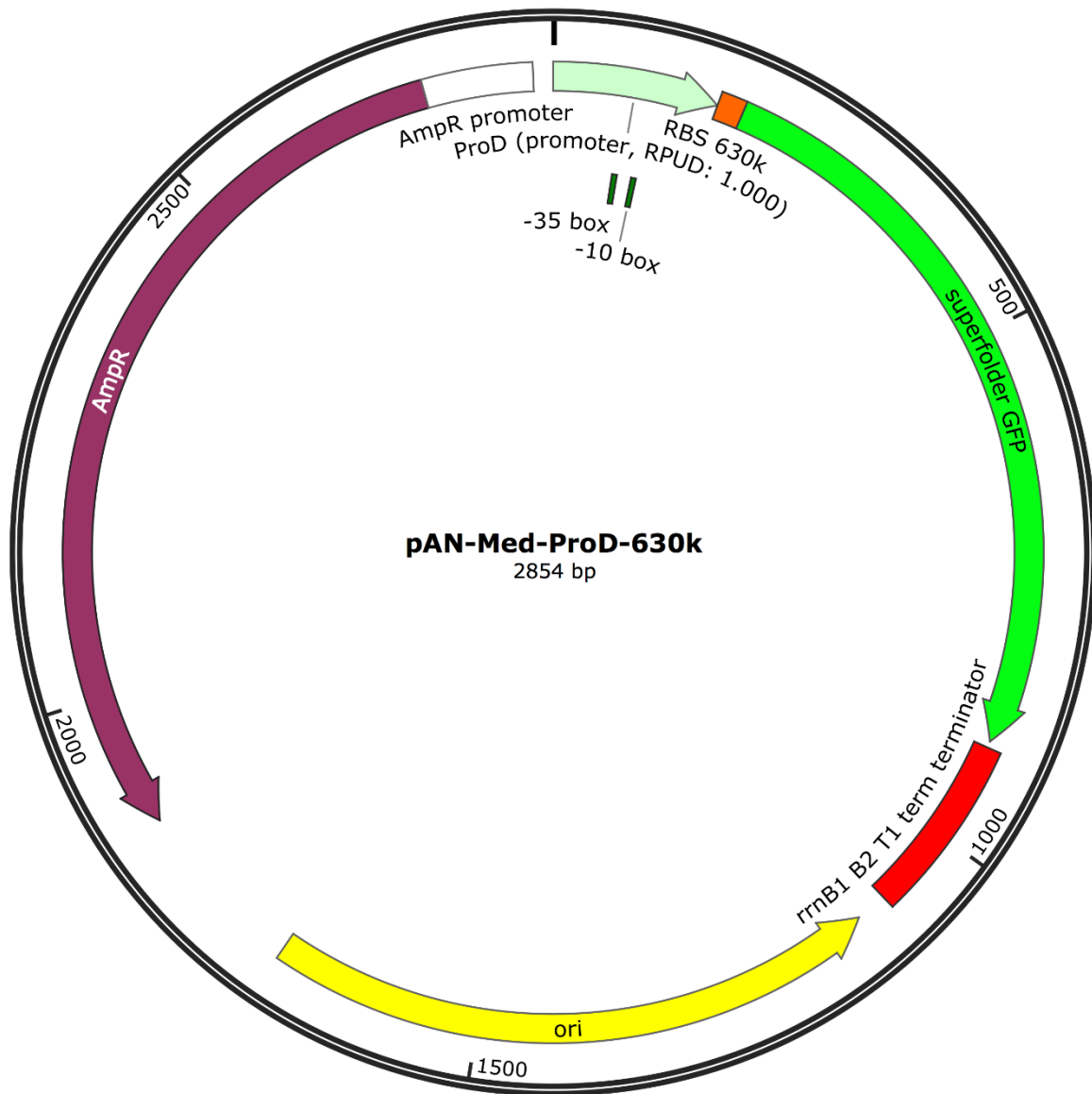
ttctagagCACAGCTAACACCACGTCGTCCTATCTGCTGCCCTAGGTCTATGAGT
 GGTGCTGGATAACTTTACGGGCATGCATAAGGCTCGTATAATATATTCAGGGA
 GACCACAACGGTTTCCCTCTACAAATAATTTTGTTTAACTTTtactagagATTAAAGA
 GGAGAAtactagATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCT
 TGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGGAGAGGG
 TGAAGGTGATGCTACAAACGGAAACTCACCCTTAAATTTATTTGCACTACTGGA
 AAATACTGTTCCATGGCCAACACTTGTCACTACTCTGACCTATGGTGTTCAAT
 GCTTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTCAAGAGTGCCA
 TGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGACCT
 ACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCG
 AGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAACTCGA
 GTACAACTTTAACTCACACAATGTATACATCACGGCAGACAAACAAAAGAATGG
 AATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTA
 GCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA

GACAACCATTACCTGTGCGACACAATCTGTCCTTTTCGAAAGATCCCAACGAAAAG
CGTGACCACATGGTCCTTCTTGAGTTTGTAAGTGTGCTGGGATTACACATGGC
ATGGATGAGCTCTACAAAtagtaaaattagcccgggcgaaaggccagctcttcgactgagccttcgtttt
atttgatgcctggcagttccctactctcgcatggggagtgcccacactaccatcggcgctacggcggttcacttctgagtt
cggcatggggtcaggtgggaccaccgcgctactgccgccaggcaaagaaccgtaaaaaggccggttgctggcg
ttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacag
gactataaagataccaggcggttccccctggaagctccctcgcgctctcctgttccgaccctgccgcttacgggatac
ctgtccgctttctcccttcgggaagcggtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcggtcg
ctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc
aacccggttaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggc
ggtgctacagagttctgaagtgggtgcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaa
gccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcgggtgtttttgtttg
caagcagcagattacgcgcagaaaaaaaggatctcaagaagatccttgatctttctacggggtctgacgctcagtg
gaacgaaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaataaaaaat
gaagtttaaatcaatctaaagtatatatgagtaaacttggctgacagttaccaatgcttaatcagtgaggcacctatctc
agcgatctgtctatttcgttcacatagttgcctgactccccgctgtagataactacgatacgggaggggcttaccatct
ggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaaccagccagcc
ggaagggccgagcgcagaagtggctcctgcaactttatccgcctccatccagttctattaattgttgccgggaagctaga
gtaagtagttcgccagttaatagtttgcgcaacgttggccattgctgcaggcatcgtggtgtcacgctcgtcgtttggtat
ggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccattgtgtgcaaaaaagcggttagctcc
ttcggctcctccgatcgttgtcagaagtaagtggccgcagtggtatcactcatggttatggcagcactgcataattctcttac
tgtcatgccatccgtaagatgcttttctgtgactggtagtactcaaccaagtcattctgagaatagtgtatgcggcgacc
gagttgctcttggccggcgtaatacgggataataccgcgccacatagcagaactttaaagtgtcatcattggaaa
acgttcttcggggcgaaaaactctcaaggatcttaccgctgttgagatccagttcgatgaacccactcgtgcacccaact
gatcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggga
ataagggcgacacggaaatgttgaatactcactcttcttttcaatattattgaagcatttatcagggttattgtctcatg
agcggatacatatttgaatgtatttagaaaaataacaaataggggttccgcgcacattccccgaaaagtg



ttctagagCACAGCTAACACCACGTCGTCCTATCTGCTGCCCTAGGTCTATGAGT
GGTTGCTGGATAACTTTACGGGCATGCATAAGGCTCGTAATATATATTCAGGGA
GACCACAACGGTTTCCCTCTACAAATAATTTTGTTTAACTTTtactagagAAAGAGG
AGAAATACTAGATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTC
TTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGGAGAGG
GTGAAGGTGATGCTACAAACGGAAAACCTCACCTTAAATTTATTTGCACTACTG
GAAAACCTGTTCCATGGCCAACACTTGTCCTACTCTGACCTATGGTGTTTC
AATGCTTTTCCCGTTATCCGGATCACATGAAACGGCATGACTTTTTCAAGAGTG
CCATGCCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGA
CCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTA

TCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAAACT
CGAGTACAACCTTTAACTCACACAATGTATACATCACGGCAGACAAACAAAAGAA
TGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCA
ACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTA
CCAGACAACCATTACCTGTCGACACAATCTGTCCTTTCGAAAGATCCCAACGAA
AAGCGTGACCACATGGTCCTTCTTGAGTTTGTAAGTGTGCTGGGATTACACAT
GGCATGGATGAGCTCTACAAAtagtaaaattagccccggcgaaaggcccagctttcgactgagcctt
cgttttattgatgcctggcagttccctactctcgcatggggagtccccacactaccatcggcgctacggcggttcacttct
gagttcggcatggggtcaggtgggaccaccgcgctactgccgccaggcaaaagaaccgtaaaaaggccgcggttgct
ggcgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagagggtggcgaaaccg
acaggactataaagataaccaggcggttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccg
gatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtc
gttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggttaactatcgtcttga
gtccaacccggttaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgagggtatgta
ggcggtgtacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatgttgatctgcgctctgct
gaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaacaaaccaccgctggttagcggtggtttttt
gtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatccttgatctttctacggggtctgacgctc
agtggaaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaa
aaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacct
atctcagcgatctgtctatttcgttcacatagttgcctgactccccgctcgtgtagataactacgatacgggaggggttac
catctggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaaccagcc
agccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagcttattaattgttgccgggaagc
tagagtaagtagttcgccagtttaatagtttgcgcaacgttggtgccattgctgcaggcatcggtgtcacgctcgtcgttt
gggtatggcttcattcagctccggttccaacgatcaaggcgagttacatgatccccatgttggtgcaaaaaagcgggtta
gctccttcgggtcctccgatcgttgtcagaagtaagttggccgcagtggttatcactcatggttatggcagcactgcataattc
tcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattctgagaatagtgtatgcgg
cgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgtcatcatt
ggaaaacgttcttcggggcgaaaactctcaaggatcttacggctgttgagatccagttcgatgtaacccactcgtgcac
ccaactgatcttcagcatcttttactttaccagcggttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaa
aggggaataagggcgacacggaaatgttgaatactcatactcttcttttcaatattattgaagcatttatcaggggtattgt
ctcatgagcgggatacatatttgatgtatttagaaaaataaacaataaggggttccgcgcacatttccccgaaaagtg



ttctagagCACAGCTAACACCACGTCGTCCTATCTGCTGCCCTAGGTCTATGAGT
GGTTGCTGGATAACTTTACGGGCATGCATAAGGCTCGTATAATATATTCAGGGA
GACCACAACGGTTTCCCTCTACAAATAATTTGTTTAACTTTtactagagAGACATAG
ATTAAGGAGGTATACCCATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCC
CAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGG
AGAGGGTGAAGGTGATGCTACAAACGGAAAACCTCACCCTTAAATTTATTTGCAC
TACTGGAAAACCTGTTCCATGGCCAACACTTGTCACTACTCTGACCTATGG
TGTTCAATGCTTTTCCCGTTATCCGGATCACATGAAACGGCATGACTTTTTCAAG
AGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGAC
GGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAAT

CGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACA
AACTCGAGTACAACCTTTAACTCACACAATGTATACATCACGGCAGACAAACAAAA
GAATGGAATCAAAGCTAACTTCAAATTCGCCACAACGTTGAAGATGGTTCCGT
TCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCT
TTTACCAGACAACCATTACCTGTCGACACAATCTGTCCTTTGAAAGATCCCAAC
GAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAAGTGTCTGCTGGGATTACA
CATGGCATGGATGAGCTCTACAAAtagtaaaattagccccggcgaaaggcccagctttcgactgag
cctttcgttttattgatgcctggcagttccctactctcgcatggggagtccccacactaccatcggcgctacggcggttca
cttctgagttcggcatggggtcaggtgggaccaccgcgctactgccgccaggcaaagaaccgtaaaaaggccgcg
ttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaac
ccgacaggactataaagataaccaggcggttccccctggaagctccctcgctgcgctctcctgttccgaccctgccgcta
ccggatacctgtccgcctttctcccttcgggaagcggtggcgctttctcatagctcacgctgtaggtatctcagttcgggtga
ggctggtcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgt
cttgagtccaacccggttaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggt
atgtaggcggtgctacagagttcttgaaagtggtaggcttaactacggctacactagaaggacagtatttggtatctgcgct
ctgctgaagccagttaccttcggaaaaagagttgtagctcttgatccggcaaacaaccaccgctggtagcggtggt
tttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacgggggtctgac
gctcagtggaaacgaaaactcacgtaagggttttggctcatgagattatcaaaaaggatcttcacctagatccttttaaat
taaaaatgaagttttaaatcaatctaaagtatatatgagtaaaacttggtctgacagttaccaatgcttaatcagtgaggca
cctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggct
taccatctggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaaccag
ccagccggaagggccgagcgcagaagtggctcctgcaactttatccgcctccatccagctattaattgttgccgggaa
gctagagtaagtagttcgccagtttaatagtttgcgcaacggtgttgccattgctgcaggcatcgtggtgtcacgctcgtcg
tttggtatggcttattcagctccggttccaacgatcaaggcgagttacatgatcccccattgtgtgcaaaaaagcggtt
agctccttcggtcctccgatcggtgtcagaagtaagtggccgcagtggtatcactcatgggtatggcagcactgcataatt
ctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcg
gcgaccgagttgctcttgccggcgctcaatacgggataataccgcgccacatagcagaactttaaaagtgtcatcat
tgaaaaacgttcttcggggcgaaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgca
cccaactgatcttcagcatctttactttcaccagcggtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaa
aagggaataagggcgacacggaaatgttgaatactcatactcttcttttcaatattattgaagcatttatcagggttatt
gtctcatgagcggatacatatttgaatgtatttagaaaaataaacaataaggggtccgcgcacatttccccgaaaagt