Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from non-model bacteria

Simon J. Moore[a,b,1], James T. MacDonald[a,b,1], Sarah Wienecke[c], Alka Ishwarbhai[d,e], Argyro Tsipa[d,e], Rochelle Aw[a,f], Nicolas Kylilis[a,b], David J. Bell[b,d], David W. McClymont[b,d], Kirsten Jensen[a,b,d], Karen M. Polizzi[a,f], Rebekka Biedendieck[c] and Paul S. Freemont[a,b,d,2]

[a]Centre for Synthetic Biology and Innovation, Imperial College London, London SW7 2AZ, United Kingdom;
[b]Section for Structural Biology, Department of Medicine, Imperial College London, London SW7 2AZ, United Kingdom;
[c]Institute of Microbiology and Braunschweig Integrated Centre of Systems Biology (BRICS), Technische Universität Braunschweig, Braunschweig, Germany
[d]London DNA Foundry, Imperial College London, London SW7 2AZ, United Kingdom;
[e]Department of Bioengineering, Imperial College London, London SW7 2AZ, United Kingdom;
[f]Department of Life Sciences, Imperial College London, London SW7 2AZ, United Kingdom;

[1]Joint First Authors

[2]To whom correspondence should be addressed:

Professor Paul Freemont
Tel.: +44 (0) 207 594 5327
Fax: +44 (0) 207 594 3057
E-mail address: p.freemont@imperial.ac.uk

www.pnas.org/cgi/doi/10.1073/pnas.1715806115

Table of Contents

Table of Contents - Supplementary Figures continued

Supplementary Text - Extended materials and methods

Assembly of plasmids

For the construction of plasmids pRBBm257 and pRBBm258 the promoter region of the operon *bmd_1326-1329* (encoding the four components of the pyruvate dehydrogenase of *B. megaterium*) was amplified via polymerase chain reaction (PCR) using the Phusion polymerase (New England Biolabs, Frankfurt, Germany) from genomic DNA of *B. megaterium* DSM319 with the primers RB216fw (AflII) and RB217rv (SpeI). AflII-/SpeI-digested PCR product was introduced into AflII/SpeI-digested p3STOP1623hp (1) and pSSBm85 (1), respectively, replacing the *xylR* gene and xylose-inducible promoter, resulting in pRBBm257 and pRBBm258, respectively (Table S1). For construction of pSWBm25, primer pair SW55fw (BglII) / SW56rev (EagI) was used to PCR-amplify *mCherry* from pKMMBm2 (2). The PCR product was cloned with BglII and EagI into p3STOP1623hp (1). For pRBBm266, the *mCherry* gene was PCR-amplified from pKMMBm2 (2) with primer pair RB233fw (SpeI) / RB234rv (SphI). SpeI/SphI-digested PCR product was cloned into pRBBm257 giving pRBBm266. pRBBm267 and pSWBm24 were constructed in a similar way. A *B. megaterium* codon optimised mCherry variant was synthesised by Life Technologies™ GeneArt™ (Regensburg, Germany) and cloned into BglII and SphI. This was adapted for Golden Gate cloning using PCR with the primers mCherry_Bm_NdeI_F and mCherry_Bm_BamHI_R and cloned into pBP-ORF-*lacZ*. For further details please see DNA sequences. pRBEc2 was constructed by PCR-amplifying $xylR_{DSM319}$ with primers RB128fw (BamHI) and RB129rv (BamHI) from pMM1525(3). BamHI-digested PCR product was introduced into BamHI-digested pET14b (Novagen, Germany) in such a way that the pET14b encoded His_6-tag was fused in frame to the 5' end of *xylR,* hereafter referred to as $XylR^{His}$. The malachite green aptamer (MGapt) was cloned with the Bba_B0015 terminator (4) through two rounds of PCR and introduced between the AgeI and SphI sites, downstream of the *gfp* stop codon in pKMMBm5(2) resulting in pKMMBm5-MGapt (pSJM673). For EcoFlex assembled plasmids, please refer to later in the text.

Plasmids were purified with the Qiagen Plasmid Mini and Midi Kit (Qiagen, Hilden, Germany) and dissolved in distilled water. dsDNA and ssRNA concentrations were measured using a nanodrop reader (Peqlab, Erlangen, Germany). All DNA sequences of the cloned genes were verified by DNA sequencing (GATC Biotech AG, Konstanz, Germany and Source Bioscience, Cambridge, UK, respectively). PCR oligonucleotides are listed in Table S2 and S4. Sequencing oligonucleotides are provided in Table S4.

B. megaterium cell-free protocol

Step 1 – Growth and S30A wash steps
For extract optimisation, samples were prepared as triplicate biological repeats. A 50 mL 2YT overnight pre-culture of *B. megaterium* DSM319 derived from a single colony was grown at 30°C, 100 rpm for 12 hrs. The pre-culture was diluted 50-fold into 0.6 L 2YT media and incubated at 37°C, 200 rpm until an OD_{578} of 2.0 was reached (early-logarithmic growth phase). Cells were centrifuged at 6000 × *g*, 4°C for 20 min. The supernatant was discarded and cells were washed with 200 mL of buffer S30A and centrifuged at 6000 × *g*, 4°C for 10 min. This was repeated again, before re-suspending the cells in 25 mL buffer S30A. Then to remove excess buffer, cells were centrifuged at 45,000 × *g*, 4°C for 10 min, followed by careful removal of the supernatant. The cell pellet was weighed (~7.8 gram per litre of culture) and re-suspended in 0.9 volumes of buffer S30A per gram of wet cells, before flash freezing in liquid nitrogen.

Step 2 – Sonication
Sonication was the first variable to be optimised and tested by varying energy input. Triplicate biological repeat pre-cultures were used to inoculate 0.5 L of 2YT medium, which were incubated at 30°C with fast shaking until an OD_{600} of 2.0 was reached. After washing and re-suspension (0.9 mL S30A buffer per mL of wet cells), 0.25 mL of wet-cell paste was sonicated for 10, 20, 40, 60 and 90 seconds in triplicate and average energy inputs were recorded as 85, 165, 274, 368, 538 Joules per mL, respectively (Fig. S5). In addition, we also tested the effect of sample pre-treatment with 250 µg mL^{-1} of lysozyme at 37°C for 10 min to increase cell-lysis. After sonication, all samples were incubated at 37°C for 1 hour, then clarified by centrifugation before freezing in liquid nitrogen. Lysozyme pre-treated samples decreased activity by approximately 35% in comparison to non-treated samples. Optimum activity was observed with an energy input of 165 Joules per mL in the absence of lysozyme (Fig. S5).

Step 3 - "Run-off" incubation
In cell-free extract preparation a heated incubation step post cell-lysis is unofficially referred to as a "run-off" step. Although not investigated at the biochemical level in detail, this particular stage of optimisation is believed to degrade fragmented genomic DNA through the release of endogenous nucleases, whilst allowing incomplete metabolic reactions to run to completion. For *B. megaterium* DSM319, triplicate biological repeats of supernatant aliquots (250 µL) were incubated in a run-off reaction, which was tested at 30°C and 37°C for a range of incubation of periods. Aliquots were then centrifuged at 18,000 × *g*, 4°C for 10 min to remove precipitate and then tested for activity. The optimum and minimal incubation temperature duration was 37°C for 1 hour for both growth temperature (Fig. S5). Additionally, we observed that if *B. megaterium* DSM319 was grown at 30°C, GFP synthesis was higher, whilst extract activity during a 30°C "run-off" peaked later and remained stable up to 6 hours (Fig. S5).

Step 4 – Dialysis: Dialysis is a common step for removing small molecules after the "run-off" step that potentially inhibit *in vitro* transcription-translation reactions. Biological repeats of *B. megaterium* DSM319 cell-extracts were also dialysed

with Buffer S30B for 4 hours at 4°C and then compared against non-dialysed extract batches. In summary, we didn't observe a significant and consistent response to dialysis treatment (Fig. S5) and subsequently, dialysis was not used for further extract preparations.

Activity comparison of the constitutive and xylose-inducible promoters in vivo

96-well plate reader characterisation
A single colony of *B. megaterium* DSM319 recombinant plasmid strains was inoculated and grown in 100 μL LB with 10 μg mL^{-1} of tetracycline (LB+Tet) in a 96-well plate at 600 rpm, 30°C for 6 hours using a ClarioStar (BMG LabTech, Germany) plate reader. For constitutive promoter plasmid strains, the initial OD_{600} was recorded and the cultures were diluted into fresh medium to an OD_{600} = 0.02. For the xylose-inducible promoter (pSSBm85) plasmid strain, the initial OD_{600} was set to 0.4. The experiment was performed as a quadruplicate biological repeat in either LB+Tet or LB+Tet with 0-1% (w/v) D-xylose to induce gene expression from the xylose inducible promoter plasmid strain (pSSBm85). The growth was continued for 16 hrs under the same conditions. The OD_{600} and GFP measurements were recorded every 10 minutes (Fig. S3-4).

48-well micro-bioreactor characterization
2.5 mL of LB+Tet pre-culture in a 15 mL falcon tube was inoculated from a single colony and incubated at 37°C, 220 rpm overnight. For 48 Well FlowerPlate® cultivation, 1 mL of LB+Tet medium was transferred to each well and inoculated with pre-culture volume adjusting to an $OD_{578\ nm}$ of 0.055. 48 Well FlowerPlate® was covered with gas-permeable membrane (m2p-labs, Baesweiler, Germany) and incubated in Biolector® at 30°C and 1400 rpm. The relative fluorescence units of GFP were directly measured by Biolector® at a wavelength of 488 nm. Biomass (scattered light units) was measured at a wavelength of 620 nm. A calibration of OD_{600} measurements was run in parallel to convert scattered light units into OD_{600}.

Shake-flask cultivations and SDS-PAGE
An overnight culture of 100 mL LB+Tet was diluted 1/100 into fresh medium and aerobically cultivated at 37°C and 250 rpm. Time point zero (t = 0) was set when the cultures reached an OD_{578} of 0.3. At this time-point, 0.5 % (w/v) D-xylose was added to cultures carrying a recombinant plasmid with the xylose inducible promoter. Samples for intracellular protein (Fig. S1) were taken after 6 hours of growth and analysed by SDS-PAGE as described previously (5).

GFP and mCherry calibration standards
GFP and mCherry were assembled by Golden Gate cloning method under the control of the T7 promoter with a N-terminal His_6-tag and the BBa_B0015 terminator using the EcoFlex assembly kit (4). *E. coli* Rosetta (DE3) pLysS strain (Novagen, Germany) was individually transformed with plasmids pSJM579 and pSJM687 encoding His_6-tagged GFP and mCherry, respectively, and grown for 24 hrs in 2 mL of 2YT medium (20 g tryptone, 10 g yeast extract and 5 g of NaCl per litre) at 30°C. His_6-tagged proteins were purified as described with $XylR^{His}$. Purified enzymes were dialysed against 2 litres of 20 mM of HEPES (pH 7.5) and 50 mM NaCl for 16 hrs at 4°C. mCherry protein

concentration was determined using the absorbance extinction coefficient of 72,000 M^{-1} cm^{-1} at 587 nm. For GFP, this was estimated by A_{280} and Bradford assay. Calibration standards were prepared in triplicate 10 µL cell-free reactions with 10 mg mL^{-1} of cell-extract and standard energy mix conditions, then equilibrated at 30°C for 10 min before measuring on a ClarioStar plate-reader. Fluorescence measurements were recorded using a filter for GFP (Ex 482/15, Em 530/40, 600 gain) and monochromator for mCherry (Ex 570/15, Em 620/20, 1600 gain). For quantifying GFP and mCherry concentrations in experimental data, the average background for the negative control (pRBBm99 – Empty vector) incubation at each time point was subtracted from the corresponding data point. The concentration of each data point was then estimated using the GFP and mCherry calibration standards that were fitted in GraphPad (Prism7) (Fig. S2).

Preparation of amino acids for NCF reactions

The *B. megaterium* DSM319 extracts were found to be sensitive to high salt conditions, so we prepared the amino acid mixture with minimal potassium hydroxide base to neutralise the pH to 7.5. The amino acid solution was prepared as follows. 40 mL of distilled water was stirred and heated to approximately 42°C. Each amino acid was then weighed (Table S6) and dissolved one at a time until the liquid was clear. This ranged from a few seconds up to 5 min for cysteine and some of the hydrophobic amino acids (leucine, tyrosine). The solution was adjusted to pH 7.5 using 25 □L aliquots of 4 M potassium hydroxide until a final concentration of 26 mM of KOH was reached. The volume was then adjusted to 50 mL and aliquots were stored at -80°C. Aliqouts were defrosted at 42°C for 5 min to re-dissolve any precipitate after freezing. The final 4 □ stock concentration was 5 mM for L-leucine and 6 mM for the remaining amino acids.

in vitro *RNA synthesis of the malachite green aptamer (MGapt) and cell-free measurements*

The *gfp-MGapt* fragment (Fig. S11) of pSJM673 was placed under the transcriptional control of a T7 promoter by sub-cloning the NcoI/PvuI-fragment into pTU1-A-T7-ET-GFP-Bba_B0015 resulting in pSJM734. Further, the region containing the T7 promoter, RBS, *gfp* gene, *MGapt* and the BBa_B0015 terminator of pSJM734 was amplified by *in vitro* transcription using the primers VF2 and BBa_B0015endREV. *in vitro* transcription was carried out with the MEGAscript® T7 Transcription Kit (Ambion®). MEGAclear™ Transcription Clean-Up Kit (Ambion®) was used to purify mRNA transcript and checked for purity on a 5% TBE-urea PAGE gel (Fig. S12) with the low range ssRNA ladder (New England Biolabs, Hitchin, UK). For preparing standards, purified *gfp-MGapt* mRNA was used in concentrations from 50 to 1000 nM and mixed with 10 mg mL^{-1} of cell-free extract, amino acids and energy buffer as described for real-time transcription-translation experiments. To reduce mRNA decay during preparation, plates were pre-incubated at -20°C, whilst mRNA standards were prepared on ice, transferred to plates as 10 µL aliquots (quadruplicate repeat) and then briefly centrifuged (~1,000 × *g* for 10 sec at 4°C). Fluorescence was optimised for malachite green concentration (Fig. S13). mRNA standards were recorded every 30 sec at 30°C for 120 min (Fig. S14). A 1 µM GFP-MGapt mRNA transcript was also spiked into the reaction and monitored for malachite

green and GFP fluorescence signals, in comparison to pRBBm258-MGapt plasmid control (Fig. S16).

qRT-PCR

A cell-free transcription-translation time-course reaction containing 10 nM of the PDH-RiboJ-GFP-MGapt-Bba_B0015 plasmid was incubated on a heat block at 30°C (Fig. S15). 150 μL samples were removed at set time points and frozen in liquid nitrogen. 150 μL of RNAprotect™ (Qiagen) was added to each sample, thawed at room temperature and mixed. For quantitative Reverse Transcription (qRT)-PCR, RNA was isolated using RNeasy® Mini Kit (Qiagen) according to manufacturer's instructions. cDNA was prepared using the High-Capacity cDNA Archive Kit (Applied Biosystems, Warrington, UK) according to the manufacturer's instructions. 1 μg RNA was used in a total reaction volume of 20 μL. qRT-PCR reactions were set up using the 2X SYBR® Green JumpStart Taq Ready Mix (Sigma-Aldrich, Dorset, UK). An Eppendorf Mastercycler® ep *realplex* quantitative cycler was used (Eppendorf UK Ltd, Histon, UK). Data was analysed using the Pfaffl method, based on ΔΔ-Ct (6, 7) and normalised to 16S RNA as the housekeeping gene. Primers for 16S were TTCGAAGCAACGCGAAGAAC and AACCCAACATCTCACGACAC. Primers for *gfp* were ATACTCCAATTGGCGATGGC and TCCCAGCAGCAGTTACAAAC. Cycling conditions were 95°C for 5 minutes followed by 40 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 30 seconds with a melting curve afterwards to ensure a single product was being measured. All samples were normalised to 0 minutes for analysis.

Recombinant production and purification of XylR^{His}

E. coli strain BL21 (DE3) Star™-pLysS (LifeTechnologies, Darmstadt, Germany) and auto-induction strain *E. coli* BL21 were individually transformed with plasmid pRBEc2. The recombinant BL21 strain was grown in 1 L of LB medium with 100 μg ml^{-1} of carbenicillin at 37°C for 6 hrs. After the addition of 0.4 mM of IPTG the culture was shifted to 19°C and was incubated for 16 hrs. For cell harvesting cultures were centrifuged (6000 × *g*, 15 min) and the cells were re-suspended in 20 mL of binding buffer containing 40 mM of HEPES (pH 7.5), 500 mM of NaCl and 10 mM of imidazole. Cells were lysed by sonication for 2 s on and 2 s off cycles for 5 min at 65% amplitude, with cooling on ice. After centrifugation at 12,000 x *g* for 15 min at 4°C, recombinant XylR^{His} was purified from the supernatant at 4°C with metal ion affinity chromatography (IMAC) using 2 mL nickel resin (GE Healthcare). After loading, the column material was washed with 10 column volumes (CV) of wash buffer I (50 mM imidazole), 5 CV of wash buffer II (100 mM imidazole) and 5 CV of elution buffer (400 mM imidazole). Soluble and insoluble crude extracts and elution fractions were collected and analysed for purity by 12 % SDS-PAGE (Fig. S17) using a protein molecular weight marker (Fisher). Purified XylR^{His} is relatively unstable and slowly precipitates after purification or freezing despite optimisation of buffer conditions, whilst glycerol cannot be added since it is detrimental to cell-free activity with 0.1% (v/v) and 0.5% (v/v) glycerol reducing activity by 47% and 80%, respectively (Fig. S28). To study XylR^{His}, the purified protein was freshly purified for each experiment. To do this, eluted protein was rapidly buffer-exchanged with a PD-10 column (GE Healthcare) into 20 mM HEPES (pH 7.5) and 400 mM NaCl (Buffer A) to remove imidazole, which we found to accelerate

precipitation of XylR^{His}. XylR^{His} was further dialysed in Buffer A for 2 hours at 4°C using a 10 kDa MWCO membrane (ThermoFisher). Precipitate was removed by centrifugation and then protein concentration was determined by Bradford assay. XylR^{His} was freshly purified for each experiment and rapidly dispensed with energy buffer, amino acids, extract, xylose and plasmid DNA using an Echo® 525 liquid handling system (LabCyte).

Mathematical modelling of cell-free transcription-translation
A general ordinary differential equations (ODE) model scheme was created to account for the expression of single or multiple genes in the same reaction. This model included accounting for the use of shared resources (NTPs, amino acids, RNA polymerase, and ribosomes). Transcription was modelled using three reactions: RNAP binding to the promoter, promoter escape, and transcription elongation.

$$b_i + p \rightleftarrows c_i$$
$$c_i \xrightarrow{NTP} b_i + c_i^{elong}$$
$$c_i^{elong} + L_{m_i}\, NTP \longrightarrow m_i + p$$

Where, p is an RNA polymerase, b_i is a free promoter site, and c_i is the promoter site with a bound RNAP, c_i^{elong} is the corresponding coding sequence being transcribed, m_i is a corresponding mRNA transcript, and L_{mi} is the length of the transcript m_i. Similarly, translation was modelled using three reactions: ribosome binding to the RBS, translation initiation and translation elongation.

$$m_i + r \rightleftarrows d_i$$
$$d_i \xrightarrow{NTP} m_i + d_i^{elong}$$
$$d_i^{elong} + 2 \times L_{x_i}\, NTP + L_{x_i}\, AA \longrightarrow x_i + 2 \times L_{x_i}\, NMP + r$$

Where, r is a ribosome, d_i is an mRNA with a ribosome bound to the RBS, d_i^{elong} is an mRNA with a ribosome in elongation phase, x_i is a translated protein, L_{xi} is the length of protein x_i. Four additional reactions were also used to model first order mRNA degradation into NMPs, first order degradation of NTP to NMP to model the effect of energy consumption by non-productive NTPase side-reactions, the regeneration of NTP from NMP using the secondary energy source, S, and a first order degradation reaction to model the inactivation of ribosomes (8).

$$m_i \longrightarrow L_{m_i}\, NMP$$
$$NTP \longrightarrow NMP$$
$$NMP + S \longrightarrow NTP$$
$$r \longrightarrow \varnothing$$

Transcription elongation, translation elongation, and NTP regeneration reactions were modelled assuming multi-substrate Michaelis-Menten kinetics with random order binding of reactants for simplicity.

Bayesian modeling of the cell-free system

Two different ordinary differential equation models of cell-free transcription/translation were created based on the reaction scheme outlined in the main paper. A single promoter model was created for the xylose inducible promoter, and a two promoter model was created for the resource competition experiments. Both of these models are available as SBML and C header files in the GitHub repository (https://github.com/jmacdona/ODE_MCMC_tools).

A Markov chain Monte Carlo (MCMC) algorithm with an adaptive proposal distribution was used to sample from the posterior distribution given the experimental data. C++ source code implementing the algorithm, together with instructions and an example data set are available to download from the same repository. The software can also search for local optima using quasi-Newton minimization. A system of ODEs is defined in a C header file using the same header format as exported by COPASI (9) and then compiled into a binary executable file for fast parameter inference. The Boost.odeint library was used to numerically integrate the system of ODEs and the Ceres Solver library was used for quasi-Newton optimization (10). Jacobian matrices and gradient vectors were determined numerically using a finite difference method. The software allows parameter inference from multiple experimental datasets under different experimental conditions and time resolutions.

The C++ MCMC software takes a number of command line arguments including the initial multivariate normal (MVN) proposal distribution, experimental data sets, and initial parameter values. Example input files and command line arguments are provided with the software distribution. Data sets are formatted as tab-separated files and contain the means and standard deviations of experimentally measured GFP or mRNA concentrations over time at different template DNA, XylR, and xylose concentrations. The MVN proposal distribution covariance matrix is also formatted as a tab separated text file.

The experimentally measured GFP and mRNA trajectories, and corresponding experimental conditions (i.e. defined as a subset of defined parameter values and/or initial values) are stored in a std::map data structure and used to evaluate the log likelihood function during MCMC.

The system of ODEs were numerically integrated using the Boost.Numeric.odeint C++ library for multiple experimental conditions with the same parameter vector and the overall log-likelihood function evaluated (equation S1). The Gaussian log-likelihood function can be replaced with a Student's t-distribution with a user-definable number of degrees of freedom, ν, for more robust parameter estimation (11). The Metropolis-Hastings algorithm with the standard acceptance criterion (equation S2) was used to generate the Markov chain using a MVN proposal distribution and a uniform or MVN prior.

Equation S1

$$\log \mathcal{L}(\boldsymbol{\theta}|\boldsymbol{\mu}, \boldsymbol{\sigma}) = \sum_{c \in \boldsymbol{C}} \sum_{s \in \boldsymbol{S}} \sum_{t=0}^{T} \log f(x_{c,s,t}(\boldsymbol{\theta})|\mu_{c,s,t}, \sigma_{c,s,t})$$

where $\boldsymbol{C}$ is the set of experimental condition (e.g. XylR, xylose and DNA concentrations), $\boldsymbol{S}$ is the set of chemical species (e.g. mRNA or $\mathrm{GFP_{mat}}$), and t is the time point. $f(x|\mu, \sigma)$ is the probability density function of the Gaussian distribution with mean, μ, and standard devation, σ. $x_{c,s,t}(\boldsymbol{\theta})$ is the simulated concentration of species, s, at time, t, given experimental condition, c, and parameter vector, $\boldsymbol{\theta}$. $\mu_{c,s,t}$ and $\sigma_{c,s,t}$ are the mean and standard deviation of the experimentally measured concentration of species, s, at time, t, and experimental condition, c.

Equation S2

$$A(\boldsymbol{\theta'}|\boldsymbol{\theta}) = \min\left(1, \left[\frac{\mathcal{L}(\boldsymbol{\theta'})\pi(\boldsymbol{\theta'})}{\mathcal{L}(\boldsymbol{\theta})\pi(\boldsymbol{\theta})}\right]^{\beta}\right)$$

where $\pi(\boldsymbol{\theta})$ is the prior distribution, $\mathcal{L}(\boldsymbol{\theta})$ is the likelihood function, $\boldsymbol{\theta}$ is the current parameter vector, $\boldsymbol{\theta'}$ is the proposed parameter vector and β is the thermodynamic beta. Normally $\beta = 1$ but is set to a range of different values during parallel tempering to improve sampling of parameter space.

The proposal distribution was adapted using the robust adaptive Metropolis (RAM) algorithm with a target acceptance rate of 0.234 (12).

In addition to standard MCMC, the software is able to carry out parallel tempering MCMC with multiple threads (using the Boost.thread C++ library to manage the threads). This permits the simultaneous exploration of parameter space at different "temperatures" (as controlled by the thermodynamic beta parameter in equation S2) with the probability of exchanges between the different Markov chains determined by equation S3. This was used to extensively explore parameter space to find regions of high probability before running standard MCMC production runs to obtain the final posterior distribution estimates.

Equation S3

$$p = \min\left(1, \frac{\left(\mathcal{L}(\boldsymbol{\theta_i})\pi(\boldsymbol{\theta_i})\right)^{\beta_j} \left(\mathcal{L}(\boldsymbol{\theta_j})\pi(\boldsymbol{\theta_j})\right)^{\beta_i}}{\left(\mathcal{L}(\boldsymbol{\theta_i})\pi(\boldsymbol{\theta_i})\right)^{\beta_i} \left(\mathcal{L}(\boldsymbol{\theta_j})\pi(\boldsymbol{\theta_j})\right)^{\beta_j}}\right)$$

where i and j are two independent MCMC chains with parameter vectors $\boldsymbol{\theta_i}$ and $\boldsymbol{\theta_j}$, and thermodynamic betas β_i and β_j respectively.

Multiple independent Markov chains were run from different initial parameter values and random seeds. Chains were checked for convergence using with the Gellman-Rubin diagnostic (13) as implemented in the R CODA package (14).

LC-MS/MS proteomics

A LC-MS/MS protocol was developed based on previous studies (15–17) to quantify proteins in the cell-extracts. To begin a preliminary tryptic digest of a *B. megaterium* cell-extract experiment was performed and analysed by LC-MS/MS. Two diagnostic tryptic peptides for each protein to be measured were selected for the targeted LC-MS/MS method. This was achieved by comparison of the peptides identified for each protein from auto-MS/MS analysis of digested samples with those predicted to be suitable for measurement by LC-MS/MS using Peptide Selector software (Agilent, Santa Clara, CA). By combining the recorded retention LC times and target precursor masses, a method to determine the concentration of proteins was developed. From this initial screen, we could not identify sufficient coverage of the Sigma A protein (RpoD), which is the major regulator of gene expression in *Bacillus* species. Based on this evidence, a *B. megaterium* DSM319 synthetic peptide (QconCAT_Bm_NCF) was designed to target two peptides for the S4, S7, L2, L4 ribosomal proteins, RNA polymerase β subunit and alanyl- and methionyl-tRNA synthetases. The QconCAT peptide was also fused with a sacrificial peptide and internal calibration standard (GluFib B). To positively identify the Sigma A factor (RpoD) from the cell-extracts, this was prepared as a full-length sequence, amplified by PCR from *B. megaterium* DSM319 genomic DNA. Both QconCAT_Bm_NCF and *rpoD* gene products were assembled with EcoFlex (47) with a T7 promoter, N-terminal His_6-tag and transformed into *E. coli* BL21 Gold (DE3), grown and induced (with 0.4 mM IPTG) in either LB or M9 minimal medium supplemented with 0.2% (w/v) glucose and $^{15}NH_3Cl$ for isotope labelling. From expression in LB medium, the 27.4 kDa ^{His}QconCAT protein was purified with IMAC under denaturing (8M urea) conditions. Inclusion bodies were unfolded prior to purification by incubating at 80°C for 30 min in unfolding buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 8M urea and 5 mM imidazole). RpoD (45.6 kDa) was purified under native conditions. Proteins were verified by SDS-PAGE (Fig. S23-24) and confirmed to be highly pure. For LC-MS/MS, 2 mL of concentrated protein was dialysed overnight in 2 L of ddH_2O at 4°C. Proteins were then lyophilised in a vacuum centrifuge and then dissolved in 0.1M ammonium bicarbonate, 10% (v/v) methanol (ABC buffer). Protein concentration was estimated by absorbance at 280 nM with the extinction coefficients of 11,460 and 25,900 M^{-1} cm^{-1} for QconCAT_Bm_NCF and RpoD, respectively.

For preparation of the cell-extracts for SRM LC-MS/MS, 100 μL of cell-extract (30 mg mL^{-1}) was dialysed in 2 L of 20 mM Tris-HCl at 4°C for 2 hours. A 1.5-fold dilution was recorded after dialysis. 100 μL of purified standards and cell-extracts (20 mg mL^{-1}) were reduced with 50 μL of 25 mM dithiothreitol at 80°C for 15 min. 50 μL of 25 mM iodoacetic acid was added, vortexed and left in the dark for 30 min. Samples were diluted with 250 μL of ABC buffer and vortexed briefly. 75 μL of samples containing cell-extract and internal ^{15}N calibration standards (0.5 μM) were digested with 25 μL of 0.12 mg mL^{-1} mass spectrometry grade trypsin (Promega, USA).

The targeted LC-MS/MS method was applied using an Agilent 1290 LC system coupled to an Agilent 6550 quadrupole – time-of-flight (Q-ToF) mass spectrometer with electrospray ionization (Agilent, Santa Clara, CA). The LC column used was an Agilent Zorbax Extend C-18, 2.1 x 50mm and 1.8um

particle size. The LC buffers were Solvent A - 0.1% formic acid in water and Solvent B - 0.1% formic acid in acetonitrile (v/v). The gradient elution method is detailed in Table S11. Quantitation was based on the LC retention times of standards (Fig. S26-27) and the area of accurately measured diagnostic precursor or fragment ion for each targeted peptide. The protonated molecules of each peptide, $[M+2H]^{2+}$, were targeted and subjected to collision induced dissociation, with product ions accumulated throughout the targeted period. Standard solutions of GluFib B were prepared using [Glu1]-Fibrinopeptide B Standard (Waters, Milford, MS) and analysed by targeted LC-MS/MS. The resulting calibration curve was then used to determine the concentration of the GluFib B released from a digested QconCAT stock solution. This value was then used for the QconCAT solutions to generate calibration curves for target peptides, which were in turn used to measure the concentration of peptides in the digested sample solutions. Concentrations of peptides were calculated using the integrated area of the peak corresponding to the elution of the peptide of interest at the retention of the standards. This was measured from either the response for the precursor ion or for a fragment ion from the product ion spectrum of each peptide. Calibration curves generated from the standards using the same data type were used to calculate concentrations, with normalisation using the respective internal standard peptide response. Additional quantitative measurements were made by using the internal standard as a single calibration point to calculate the concentration of target peptides in the samples.

Semi-automated cell-free screening of RBS library parts

For clarity, elements of this workflow used both manual and automated steps (i.e. semi-automated), which we feel is appropriate for any standard molecular biology laboratory to perform even without liquid handling capability.

Library construction

To ensure high-yields of DNA from 96-well purification, the backbone of the positive control plasmid (PDH-RiboJ-GFP-MGapt-Bba_B0015) was modified to include a pUC19 origin of replication, using Golden Gate assembly. The plasmid also included the *B. megaterium* shuttle vector parts RebB and TetA for *in vivo* experiments. Inverse PCR primers were designed to amplify the plasmid (6884 - 6889 bp) using Q5 polymerase (NEB). The spacer region (WWW) was varied from three to eight W nucleotides to provide the RBS-3, -4, -5, -6, -7 and -8 libraries.

RBS library primer(s)
5’ AGGTCTCAATCARRRRRRRWWWATGACTAGTTCGAAGATCTGGGCTAG 3’

Reverse primer
5’ AGGTCTCTTGATTTAAGTGAACAAGTTTAAACAAAATTATTTGTAGAGG 3’

Ligation

100 ng of purified PCR product was then digested and ligated in a one-pot reaction using a standard EcoFlex Level 1 15 µL reaction. 0.5 µL of DpnI was also added to remove any residual template DNA. 2 µL of EcoFlex mix was

transformed into 25 µL of competent *E. coli* cells, plated onto carbenicillin plates and grown at 37°C overnight.

96-well DNA purification

For the RBS libraries, single colonies were cultured in 1.2 mL of Terrific Broth (Merck) with 100 µg mL^{-1} carbenicillin in a 2 mL deep-well plate (Greiner) for 16 hours shaking at 37°C, 300 rpm with a Breathe-Easy® (Sigma) seal. For each library, 88 clones were picked, along with 4 positive and 4 negative control colonies to complete a 96-well plate. Plasmid DNA was purified with a PureLink® Pro Quick96 Plasmid Purification Kit (Invitrogen) as outlined by the manufacturer's guidelines. Modifications to the protocol were made as follows. After resuspension of cells in 250 µL of A1 buffer, to minimise time during the alkaline lysis step, addition of the alkaline lysis buffer (250 µL of A2 buffer) was carried out with a multi-channel dispenser. The plate was then immediately sealed with an aluminium plate seal (ThermoFisher), before inverting four times slowly to mix. The plate was then briefly centrifuged at 500 x *g* for 20 seconds to ensure that upon removal of the seal, no cross contamination between wells occurred. This procedure of plate sealing, plate inversion and brief centrifugation was also repeated for the following neutralisation step (350 µL of Buffer A3). Finally, after following the recommended wash steps, the plate was centrifuged at 4000 x *g* for 15 min at room temperature to ensure full removal of ethanol. DNA was eluted with 50 µL of elution buffer. Plate DNA concentrations were quantified with PicoGreen™ assay (ThermoFisher) (Fig. S38).

Low-volume cell-free screening reactions

Cell-free reactions were set-up as follows. 4 µL of 88 clones, 4 positive controls and 4 negative controls were dispensed into a low dead volume (2 µL) 384 well COC Source Microplate (Labcyte). 576 µL of cell-free reaction mix was warmed to room-temperature and transferred to a 6-well source plate (Labcyte), with 200 µL dead-volume. 0.5 µL of plasmid DNA was then randomly dispensed with a Echo® 550 liquid handling robot (Labcyte) as four replicates into a 384 well Small Volume™ LoBase Microplate (Greiner). Plate randomisation was prepared using the Echo® Cherry Pick software (Labcyte) to minimise time-differences between wells. Reactions were started with the addition of 1.5 µL of cell-free reagent into each well using the Echo® 525 liquid handling robot (Labcyte). The plate was then briefly centrifuged, sealed with an aluminium plate seal (ThermoFisher). Plates were then measured on a Synergy™ 2 (BioTek) plate reader and incubated for 5 hours (end-point) at 30°C. GFP fluorescence was measured every 15 mins with the following settings: excitation 488 nm, emission of 525 nm and a gain of 85.

Supplementary Text - Extended Results

Optimisation of a B. megaterium DSM319 cell-free extract
Following the protocol of Sun *et al* describing preparation of a *E. coli* cell-free extract (18), a modified protocol for cell-free protein synthesis using the wild-type *B. megaterium* DSM319 strain was developed. For initial extract optimisation steps, a 0.33 volume of extract (~10 mg mL^{-1}) was incubated with 5 nM of pRBBm258 plasmid, 1 × 3-PGA energy buffer, 1 × amino acid mix, 6 mM Mg-glutamate, 50 mM K-glutamate and 2% (w/v) PEG8K in a technical triplicate repeat.

Extracts were prepared through four stages of preparation – growth, cell-disruption, incubation at a controlled temperature ("run-off" period) and dialysis. Following the S30 method with minor modifications the cells were lysed with sonication followed by a 1.5 hour "run-off" period at 37°C and a 3 hour dialysis. From this, significant levels of GFP synthesis (1.4 □M) were detected with 10 mg mL^{-1} of protein extract incubated with 10 nM of plasmid DNA, 1.5 mM amino acids and a 3-PGA energy buffer system over 4 hours. To optimise the method further, the sonication, "run-off" and dialysis stages of extract preparation were individually tested. Maximum activity was observed with an energy input of 154 kJ mL^{-1} for sonication and a "run-off" period of 1 hour at 37°C. These two modifications increased GFP batch synthesis up to 2.8 □M (Fig. S5). Further details of optimisation are outlined in the text. As a third preparation step for cell-free extracts, dialysis is used to remove small molecules. However, it was found that dialysis was not always essential for the preparation of all *E. coli* cell-free extracts (19, 20). For *B. megaterium* DSM319, dialysis was not essential to improving cell-extract activity (Fig. S5). With further optimisation of Mg-glutamate (2 mM) and K-glutamate buffer (50 mM) conditions with 3-PGA as the energy source, maximum levels of 4.96 □M GFP was produced at 30°C from 10 nM pRBBm258 PDH constitutive plasmid, with a band corresponding to GFP observable on SDS-PAGE (Fig. S1). This represented a 3.5-fold improvement over the initial test condition.

Optimisation of B. megaterium DSM319 cell-free buffer conditions
Mg-glutamate and K-glutamate buffer optimisation: The Mg- and K-glutamate buffer was tested as a two-dimensional grid using the Echo liquid handling robot. Mg-glutamate showed the greatest effect on cell-free protein synthesis activity peaking at 6 mM. With no added Mg-glutamate, only trace levels of GFP were detected. K-glutamate up to 50 mM showed only a slight increase in GFP synthesis, whilst between 50 and 200 mM, a gradual decline in activity was observed (Fig. S5).

Molecular crowding agents: For molecular crowding, the PEG concentration was varied from 0-4% (w/v) and PEG sizes 2K, 4K, 6K, 8K and 10K were tested. Significant activity was detected in the absence of PEG in comparison to the control condition containing 2% (w/v) PEG 8K (Fig. S5).

Energy buffer
3-PGA energy buffer: The reaction buffer of Sun *et al* (18) was kept constant throughout the preliminary testing and for prototyping of gene expression tools.

The buffer consisted of a 14 × stock of energy solution containing 700 mM HEPES pH 8, 21 mM ATP, 21 mM GTP, 12.6 mM CTP, 12.6 mM UTP, 2.8 mg mL^{-1} *E. coli* tRNA, 3.64 mM coenzyme A, 4.62 mM NAD, 10.5 mM cyclic AMP, 0.95 mM folinic acid, 14 mM spermidine and 420 mM 3-PGA. Aliquots were stored at -80°C.

Testing of the xylose-repressor system in an extract with minimal native XylR interference

Expression of the *xyl* operon is repressed from the divergently transcribed repressor XylR. In the absence of xylose, XylR binds to the operator sequence to provide tight transcriptional gene expression control. Xylose binding to XylR releases the protein to allow transcription to start with gene expression induced up to 350-fold in *B. megaterium* DSM319, thus acting as a sensing mechanism for the cell to switch metabolism and utilise xylose as a carbon source (21). To study this system in cell-free reactions, we used a xylose-promoter reporter plasmid lacking the XylR regulator (pKMMBm5). Previously, this plasmid was shown to constitutively produce GFP at similar levels in wild-type *B. megaterium* DSM319 and WH325 Δ*xylR* mutant strains (2). This suggested that under standard growth conditions, the levels of XylR produced from the genomic copy of the gene were insufficient to repress P_{xyl} controlled GFP expression from the medium copy plasmid pKMMBm5 (2). In addition, with the pKMMB5 plasmid, DSM319 extracts did not show any changes in GFP expression with the addition xylose (Fig. 3, *SI Appendix*, Fig. S28). This suggests that the level of XylR in the extract is below the limit of detection. We were also unable to detect peptides derived from XylR in trypsin-digested cell-extracts, which we performed to inform the design of the synthetic QconCAT peptide. Therefore, we suggest our experiments were performed with minimal interference from endogenous XylR. In addition, prior to further testing, we also verified that the cell-free reaction is not inhibited by high D-xylose concentrations or buffer components (NaCl, glycerol) derived from XylR purification and storage (*SI Appendix,* Fig. S28).

Cell-free GFP and mCherry production at different temperatures

Plasmids with the PDH promoter controlling GFP (pRBBm258) and mCherry (pRBBm267) genes were tested at different temperature-controlled incubations of 25°C, 30°C and 37°C. The highest level of mCherry protein production was 0.17 □M at 25°C, which decreased further with increasing temperature, whereas GFP levels were found to be optimum at 30°C (Fig. S6).

Codon optimised mCherry synthesis

The genome of *B. megaterium* DSM319 has a low G+C content of 38%. This is also reflected by the presence of the rare codons for Arg (CGG), Ala (GCC) and Thr (ACC), which were identified by “tRNAscan-SE” (22). The presence of these rare codons in the iGEM derived mCherry gene is relatively high (62% G+C) with a low codon adaption index score of 0.15. To test this, a *B. megaterium* codon-optimised mCherry variant (31% G+C) was tested in comparison, along with an additional high G+C (%) variant codon optimised for *Streptomyces* (64% G+C) expression. These variants were assembled into pTU1-A with the *pdh* promoter, RiboJ-RBS and Bba_B0015 terminator. 10 nM of plasmid was added in a standard NCF reaction and incubated at 30°C for 6 hours. The *B. megaterium* optimised mCherry produced the highest levels at 0.37 μM, whilst

the iGEM (0.019 μM) and *Streptomyces* (0.009 μM) produced only trace levels of mCherry (Fig. S7). In the absence of *E. coli* tRNA an 8-24% decrease in mCherry synthesis was observed between the three variants (Fig. S7). This suggests that genes codon optimised towards low G+C (%) content are more favourably synthesised in *B. megaterium*, and that the cell-extracts contain sufficient levels of native tRNA.

Bayesian modelling of energy consumption and efficiency in the NCF system

The total theoretical energy potentially available to the system based on the initial NTP and the 3-PGA secondary energy source concentrations (and ignoring metabolism of the amino acids) was estimated by assuming that 1 mol 3-PGA molecule yields 13.5 mols high energy phosphate bond equivalents through complete oxidative phosphorylation (23). The final reaction mix contains 4.8 mM of NTPs and 30 mM 3-PGA, giving a maximum of 414.6 mM of high-energy phosphate bond equivalents. Rather than directly using these values in the model we have inferred the initial concentrations of NTPs and of the 3-PGA secondary energy source (in units of NTP → NMP equivalents). This is because it is highly likely that the 3-PGA is not fully metabolised in the cell-free reaction and we wish to estimate the proportion of the theoretical energy that is accessible to the system in practice. The initial NTP concentration for these reactions was inferred to be 4.3-5.8 mM (95% confidence interval, median 5.1 mM), the initial 3-PGA concentration (in terms of NTP equivalents, i.e. two high energy phosphate bond equivalents) was inferred to be 7.9-9.7 mM (*SI Appendix*, Table S9-10), and the summed total accessible concentration of high energy phosphate bonds available to the reaction from both of these sources was 24.8–30.4 mM. This suggests that around 6-7% of the total theoretical energy input was accessible to the system in practice, and around 4-5% of the possible theoretical energy from complete oxidation of 3-PGA was used to recycle NTPs. The fraction of the total accessible energy used for protein synthesis depends on the promoter concentration and strength, and the RBS strength. Lower concentrations of promoter result in lower final protein concentrations as well as a lower protein accumulation rate because a larger fraction of the total energy is diverted to competing non-productive NTPase side-reactions. Taking the most productive condition in this experiment (40 nM *pdh*-RiboJ-GFP-MGapt and 0 nM *pdh*-RiboJ-mCherry plasmid concentrations), the model predicts that 12-15% of the total available energy was used for transcription, 9-11% was used for translation and 74-79% was wasted in non-productive NTPase side reactions. Therefore, of the maximum total theoretical energy (414.6 mM) available, only 0.64-0.66% of this was directed to protein translation. The model suggests that improving the metabolism of the secondary energy source would substantially increase the overall protein yield but increasing the concentration of NTPs would not have much effect (*SI Appendix,* Fig. S36). This is because NTPs were inferred to have a half-life of around 8.7-9.7 minutes (before being consumed in non-productive NTPase side-reactions), so increasing the NTP concentration just increases the rate at which the secondary energy source is used to regenerate NTPs that are then wasted in non-productive side-reactions.

Inactivity of the pyk and tpi constitutive promoters

The *pyk* promoter controls the pyruvate kinase (BMD_1526) gene, whist *tpi* drives a three-gene operon encoding the glycolysis enzymes triosephosphate isomerase, phosphoglycerate mutase and enolase. In a separate preliminary data collection (described under LC-MS/MS analysis section), LC-MS/MS analysis of trypsin-digested cell-extracts identified the presence of eight distinct peptides for the triosephosphate isomerase (BMD_5036), but not the pyruvate kinase (BMD_1526). Three distinct peptides of an alternative genomic copy of the pyruvate kinase (BMD_4760) was detected by LC-MS/MS, suggesting that the *pyk* promoter we isolated upstream of *bmd_1526* is redundant under the conditions tested. In addition, for the *tpi* promoter, considering that the growth conditions used (rich media) for both promoter testing and cell-extract preparation is identical, the presence of the triosephosphate isomerase protein in the extracts suggests that the upstream promoter is active. Considering we isolated and characterise the entire intergenic region for all of the promoters, we can only speculate that under the test conditions, the *tpi* promoter is controlled by an unknown regulatory loop, or the removal of the native RBS has negatively affected promoter activity.

Effect of NaCl, glycerol and xylose on cell-free activity

To maintain stability of XylR, the purification of His_6-tagged *B. megaterium* XylR requires a high NaCl concentration (400 mM) and 15% (v/v) glycerol to maintain solubility and storage at -80°C, respectively. For testing the xylose promoter activity in cell-free, the additives NaCl, glycerol and xylose were tested for their effect on cell-free activity, using 2.5 nM of pRBBm258 under standard conditions. This testing showed D-xylose did not have any effect on cell-free protein synthesis activity above 1 mM, which is the maximum used for the xylose promoter activity in cell-free. As glycerol did inhibit cell-free protein synthesis activity, we chose to freshly purify XylR for each experiment without glycerol. XylR could not be stabilised by alternative buffering components, so a high-salt concentration of 400 mM NaCl was essential for purification and stability. However, we concentrated XylR up to 60 µM and diluted to a maximum concentration of 1 µM XylR in cell-free reactions. The maximum diluted NaCl concentration is 6.67 mM, which is sufficient to avoid inhibition effects by high-salt concentrations. A summary of these findings is in Fig. S28.

DNA parts used in this study and assembled with EcoFlex

All DNA parts were cloned as described by EcoFlex assembly method (4). The *bmd_4746* (*gapdh*) promoter, RiboJ insulator (24), QconCAT_Bm_NCF sequences were synthesised by Integrated DNA Technologies (IDT), USA. The *B. megaterium* codon optimised mCherry variant was synthesised by Life Technologies™ Geneart™ (Regensburg, Germany).

The EcoFlex Level 1 assembly format was modified to standardise and characterise the *B. megaterium* promoter parts. Golden Gate BsaI overhang order of assembly is shown in highlight below. Parts were assembled with EcoFlex (4) into the destination vector pTU1-A-*lacZ*.

CTAT-GTAC	Promoter
GTAC-CATA	RiboJ-RBS - modified from Lou *et al* (24) to include RBS
CATA-TCGA	GFP-MGapt-Bba_B0015
TCGA-TGTT	TetA-RebB shuttle vector part

***B. megaterium* DSM319 putative Sigma A promoters**

The entire intergenic sequence from the upstream gene or operon was amplified. The native RBS site was removed, by eliminating the putative RBS site along with any preceding G nucleotides. The removed sequence of the putative RBS is listed beneath each gene, along with the start codon (coloured red). Flanking BsaI sites and overhangs are highlighted. Partial sequence alignment to a consensus Sigma A box (^{35}TTGACA and $^{-10}$TATAAT) is highlighted in green. We list these as putative Sigma A promoters.

Hexokinase (BMD_0536)

GGTCTCTCTATAAAATGAGGCTGAGACAAAAGTATTTTAGCCTAAGTGAAAAACGAGCCATT
GATCAAAATGTTTGATTCATGGTTCGTTTTTTTATTATGGTGAACGTAGGTTTCATCCGTTTT
GTTCCTTCTAGCAGCTGCATGGAAATCAACAGCGGGGTCACAAGCTATTCATACGAGCTAA
TGCATCCATTTGTTCGTCTTTCGACTGTATTTATTTCTTTATGTCTCAACCTCATTTTTGAAT
TTTTAATAATGTTTATTATAATAAAAAACAGGGAAACTAATTGCCTAATGTACAGAGACC

Native RBS - GAAGAGAGAGAAAGGGTTGTTTTAACGTATG

Phosphoglucose isomerase (BMD_4923)

GGTCTCTCTATACGCTGGAATTTCCTATACAGAAGCGGGACAAAAAAGGCTAATAAAAAAG
GCGATAAAATCACCAAGTGCAACCACGACAAAAGACAAACCTCCTTGTGATATTCAATTCTG
TAACCGATTATATAATAATTTAATTATGCATGCATCCATGCTAATTCTTAGTATCAGCAGAATT
TTCAAGGCTTATTAAATAAATAGAAACTCGTCACAAAATCCATAAAAAAAACATAAAAAAATTT
TTTTCGACGTATAAAAATCATGCGAAATTGTCACTCTAAGATTGAGGTGATGGAAATGTATA
GGAAACAAATAGCCATCAGTTTTCTCATTCTTACCGCACTATGTATGAGTGGATTTGTGCTG
AATAATCACTTGGAAATGAAAGAAGAAAACAGGATGCTGCTTACTTCTAAAAACGAAGCTGA
AAATTCCATATCAATAGACTCTTCCCTGTTAAAAGCCTCTTATAAACCGAGACAATATGTAAA
GATAGGTGAAATGAACTATATTGCACCATCCGGTAAGCAAATAAAAAACGATTGAAGACCAC
TCTTTAGTCGTTTTTTATTTGCTTATTATTTGTGTAATCTTGCATGAATTACGTTCTATAATCG
AAGAACATTAGTTGACTTCATAACTGAATTTATCGTTAAATAAAAGGGTACATACAATGATTG
CAGTTTCAACAAACTATAATGATTGTGTCAAACTGTAAGTGGATTCACTTGAATTCAGTTAAC
AGTTTCGATAAAGTGAAAAATAAAGATATTTTAATGTACAGAGACC

Native RBS – GGAGGAACAGACTATG

Phosphofructokinase (BMD_3400)

GGTCTCTCTATGTTCACAAATCACTACATTCGTACGAACACTCTGCAATAATTCACTTTTAGG
TTGGAGTGTAGAGATAAGAAAAGCAGCGTGATGTTTTACATCCGCTGCTTTTTTATATCGTA
AAATTTATTTAGTACAAGGCGCCCGTTGTATAGCTAAATTGCTTAGGGGCGGGTTTTAAAAC
CGGATATACTGAAATAGTGAAACGGTCAGAATTAGGGTGGATAAGGTATCAGCAGACATTA
ATGTGATGATTTTCCCTCGATTTTTCTCTTCATCTAACAAATCCAACAGAAAAAATCTCTGAA
CATAAAAGAACAGAGATTCAACAGAATGATGATTCTTATATAATACAAAGGATGACTTTTCCT
GTCTCAGAAAGGTCATTTTTATTGTAACACCAGCAAAGTCTTAACGCTGTGAGATATCTCAC
AGCTAGGGAAGAAAAAACCTGCTATTATATAGAGGTCGGAGTAAACTTTCTTAATTATGTGC
GCTTCTTTCATATGTATCGAACCTGTTGACTGAGATCAGCAGGTTTTTTTATGTTACCAGCTA
GCTGTCATCTTTAAAGGGGCAAGATCGTTTAATCTTTATAAACAAATAGACTTTATACCATTT
TACCTATACGTAAAATAGGTTAAAATAAAGGGGTGAGGCGTGGAGAGAAAAGAGTTAAATAA
GAGGTGCGTTACATTATGATCACAATTTATCTTACAAGGTTATAACGGTGATTATACTAAATG
GTGAACGTTGAAAGTACAGAGACC

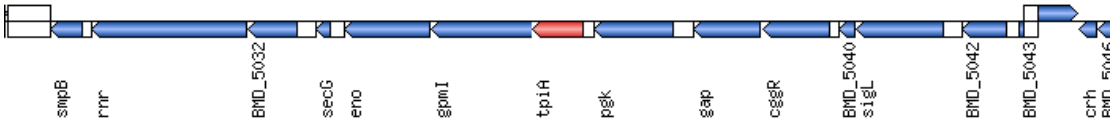
Native RBS – GGGGATACATAGTTATG

Fructose biphosphate aldolase (BMD_5161)

GGTCTCTCTATGGTTTTGGTAGGGTTTTCATCCTAACAAATTTTGAATATCAGCCTAAAAATG
TTGCTGTTTTCATTGTTAGTTGGTATCCTAATTACTGTATTCATAATTCGACAATCTATTATGT
AAAGATGCGCGTTGTTTCTACATGAAACACTCAATTCAAGAGCATACTATTTGAGTGTGCCT
TTTGGTTAGCATCTAATCTAAGGTTGATTTTAAGTACAGAGACC

Native RBS – GGAGGATTTCAATATG

Triosephosphate isomerase (BMD_5036) – shared operon with phosphoglycerate mutase (BMD_5034) and enolase (BMD_5035)



Graphic from http://megabac.tu-bs.de/index.php

GGTCTCTCTATGTTTATGTTTCAGTAAGACGGAGGACCGGCCTGACGACTGATAATATAAGA
AACTATCAAGCAGCGAAGAAGGGTGACCAAGAGATAAAAGGGACCGAATCACTTTTAACAG
CACCCTCTTGGACATCTAACGCGTGAAGGGAAAACCCACAATAAACAAAGTACAGAGACC

Native RBS – GGACGGTGTTATACATGCGTATG

Glyceraldehyde phosphate dehydrogenase (BMD_4746)

GGTCTCTCTATAGAGTGAATGATTGCATTCACTCTTTTTTTATTGGGAATATTTAAGGGCTTC
TAAAAAAGAATGTGTTAAATATCTGTTCTCGCTATATACATATAGTAAAACTTAATTAAAATAA

TTTATTAAAAAAAATTAATATTATTCACTTTTTAAATATTTATGTTATACTATTCAATATAAAAA
GTTAAATAGTATAACACATATAGTACAGAGACC

Native RBS – GAGAGGGGTATAGCCATG

Phosphoglycerate kinase (BMD_5037)
GGTCTCTCTATGAATAACTTAAAATTCGCTTCTATTTTATTCAATAAGACAGAAATCGACTAG
ACTTTCCTTTTTTATTGGATATCTGAAAATAGAACGACTATAATAAGACTTGCTAAGGAAGAG
AGGGGCATACCCCCTCTTCCTTGATTTATGTATTGGCTTCTACATGAAAAAGCGTCATAAGC
AGCTAGCTGGAGTTTCACTTTATATAGAAAGGCGCCAATATATAAAATGAAATCTACGGTGG
TTTGTGGATTGATGATTACATAGAAGCCTACTTGCTTAGTACCATAAATACTTAAGTACAGAG
ACC

Native RBS – GGAGGCCTTTTAACGATG

Pyruvate kinase (BMD_1526)
GGTCTCTCTATTAAAAGGGCTGTCTCTTAGGCAAGTTATTGACTAAGAGGCATCTTTTTGTG
TTGGTATACCTATTTTTGTATAAAGCTTATCTCTTTTATGAAAATGGGCATCCTACATTAAAAG
TACAGAGACC

Native RBS – GGATGGTGAAAGGCAATG

Pyruvate dehydrogenase (BMD_1326)
GGTCTCTCTATGAAGGCCTGTGAAGTTTACAGGTCTTTTTTTTGGCTGAAAAGGAAGGTTTT
CTGAAATAAAGAAAGAGTTTCGACATGATTCACGTAAAAATTACAAAAGGTTATTGTTTTAAA
AAATCGTTTATTTGTGGCAAGATAGGAGATGTGCAGACGTAAAAAAAGAAAAGGTTATTATA
AAACATATAATACAGTTTGCGTCTGTGTAATAGTACAGCATGTCTTTTTTATGTGTTTTATTTC
ACAAACTAAAAAAGCGTAATTTCTTGATGTAAATGCATTGACGAACTGTTTTAGTGTATTGTA
AACTAGGTAGCGTAGTAAGTGTTGTATTAATCATTTTTGCAAAATAATTAATACAGTTTAACG
AAACGTTCGTTTCGCTTAAACTTTCCTTTTACGTAAATGTTTGCACAGAATTCGCTGGGGAAT
ACATAGAAATGTTAAGGATTAGCGAATTCAAAAAGCTACTCGACAAAATATGACTTTTACTTG
TCCATTTGGGACATTCTAAAAGCGTGTATGTTTGTTTGATTAGTAAAATGAATTTATACGGCT
TAAATGAAAGTACAGAGACC

Native RBS – GGAAGAGGTGACTTAGATG

Promoter linker (PL) – used for negative control experiments for comparing activity of Sigma A constitutive promoter plasmids
GGTCTCTCTATTATATAGTACAGAGACC

Other DNA parts

Non-coding parts were cloned with either NdeI and SphI, or BsaI into pBP-*lacZ*. ORFs were cloned with BsaI or NdeI and BamHI into pBP-ORF-*lacZ*. Restriction sites are highlighted. Start and stop codons are coloured red. The MGapt sequence is highlighted in dark green.

pBP-ORF-lacZ – Modified from pBP-ORF (4) for blue/white screening

GGTCTCACATATGTTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGGTA
CAAAGAGGAGAAAGGACATGGGTGACCCGAAACTGCTGCTGGAAGTTCCGCACGCTATCT
CTTCTTCTCCGGGTAACTCTCTGGCTGTTGTTCTGCAACGTCGTGACTGGGAAAACCCGGG
TGTTACCCAGCTGAACCGTCTGGCTGCTCACCCGCCGTTCGCTTCTTGGCGTAACTCTGAA
GAAGCTCGTACCGACCGTCCGTCCCAGCAGCTGCGTTCTCTGATCCGTCTGCTGACCAAA
CCGGAACGTAAACTGTCTTGGCTGCTGCCGCCGCTGTCTAACAACTAACGCAAAAAACCCC
GCCCCTGACAGGGCGGGGTTTTTTCGCACTAGTGGATCCTCGAAGAGACC

RiboJ-RBS - derived from Lou *et al* (24)

GGTCTCTGTACAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC
AGCCTCTACAAATAATTTTGTTTAAACTTGTTCACTTAAATCAAGGAGGTGAATGTCATAAGA
GACC

GFP-MGapt-Bba_B0015 – (malachite green aptamer highlighted in dark green)

GGTCTCACATATGACTAGTTCGAAGATCTGGGCTAGCAAAGGAGAAGAACTTTTCACTGGA
GTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGG
AGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAATTTATTTGCACTACTGGAA
AACTACCTGTTCCATGGCCAACACTTGTCACTACTTTGACCTATGGTGTTCAATGCTTTTCC
CGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGT
ACAGGAACGCACTATATCTTTCAAAGATGACGGGAACTACAAGACGCGTGCTGAAGTCAAG
TTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGA
AACATTCTCGGACACAAACTCGAGTACAACTATAACTCACACAATGTATACATCACGGCAGA
CAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACATTGAAGATGGATCCG
TTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA
GACAACCATTACCTGTCGACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGCGTGACC
ACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAGCCCTA
CAAATAATGAATTCGCGGCCGCATGCTCGATATAGCAATACGGATCCCGACTGGCGAGAGC
CAGGTAACGAATGGATCCCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGG
CCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCT
TCGGGTGGGCCTTTCTGCGTTTATATGTTACCGGTCCAAGTCGAAGAGACC

B. megaterium DSM319 Sigma A factor – *rpoD*

GGTCTCACATATGGCTGAAAAATCAGCTCAATCAAAACAACTTGATTCAGATTTAACAATTGA
ACAAGTAAAAGATCAATTGACCGAAATCGGAAAAAAACGTGGTGTGCTCACATACGAAGAA
ATTGCAGAGCGTATGGCAAACTTCGAAATTGAATCCGATCAAATGGATGAATATTATGAATA
CCTTGGAGAACAAGGTGTTGAAGTAATGAGTGAATCTGAAGCAAATGATGACGATGCCGAT
CCCGATATTCAAGAACTGTCAAAAGAAGAAGAGTTTGATTTAAATGATTTAAGTGTACCTCCT
GGTGTGAAAATTAATGATCCAGTTCGTATGTACTTAAAAGAAATCGGACGCGTTGACCTATT
GTCGGCAGAGGAAGAAATCAATTTAGCAAAACGTATTGAAGAAGGCGATGAAGAAGCGAAA
CGTCGTCTTGCTGAAGCAAACTTACGTCTTGTTGTAAGTATCGCTAAACGCTATGTCGGGC
GCGGTATGCTGTTTCTTGATTTAATTCAAGAAGGAAACATGGGTCTTATTAAAGCGGTAGAA
AAATTTGATTATCGCAAAGGTTATAAATTCAGTACGTATGCGACATGGTGGATTCGTCAAGC
TATTACTCGTGCAATTGCTGACCAAGCAAGAACGATTCGTATTCCCGTTCATATGGTTGAAA

CGATTAACAAGCTAATTCGTGTTCAACGTCAGCTTCTACAAGATTTAGGTCGTGAACCATCT
CCTGAAGAAATTGCAGAAGATATGGACTTAACACCAGAAAAAGTTCGTGAAATCTTAAAAAT
TGCTCAAGAACCGGTTTCCCTTGAAACACCAATTGGTGAGGAAGATGATTCGCATTTAGGT
GACTTCATTGAAGACCAGGATGCAACATCACCTTCTGAACATGCTGCATATGAATTATTAAA
AGAGCAGTTAGAAGACGTCCTTGATACGTTAACAGATCGCGAAGAAAATGTTCTTAGACTCC
GCTTTGGTCTGGACGATGGACGTACGCGTACGCTTGAAGAAGTAGGAAAAGTATTTGGCGT
AACGCGTGAGCGTATTCGTCAAATTGAAGCTAAAGCGCTTCGTAAGCTTCGTCACCCTAGC
CGAAGCAAACGTTTAAAAGATTTCTTAGAATAATCGAAGAGACC

QconCAT_Bm_NCF

CACCATATGAAAGAAACCGCAGCAGCCAAATTTGAACGTCAGCACATGGATAGTCCGGATC
TGGGCACCCTGGTTCCGCGTGGTAGCATGGCAATTAGCGATCGTGAAGGTGTTAATGATAA
CGAAGAAGGCTTTTTTAGCGCACGCATTAGCGAATATGGTCTGCAACTGCAAGAAAAAGAA
GCACTGGAAGTTAATAACTTTGTGCCGGATTATCTGACCGTTGATGCAGAAAAAGCATTTGA
TCTGGTTCAAGAACGTTGGCTGGTTAATTATGCACGTAGCGCAGGCACCAGCGCACAGGTT
CTGGGTAAAGTTGCAACCATTGAATATGATCCGAATCGTGCACAGGATAATAGCATTGTTGT
TCTGGAAGCACTGAGCTTTGATGCACCGAAAGGTGGTGGTGTTGTTTTTGGTCCGGTTCCT
CGTCTGGCAGAAACCCTGGTTGATCCGGAAACCGGTGAAATTATTGCCGAAAAAGATATTC
CGAATGTGGGTGAAGATGCACTGCGTGTTGAAGGTCTGCTGGAAGAAATTCGTATTGTTAG
CGAAAGCGGTATTGGTGCAGGCACCCGTGTTGCACAGGTTACCCAGGTTGAACCGGTTAA
AGAAGTTCCGTTTGGTTCAGATGGTGTTTTTACACCGGAAGCATTTGTGGAACGTCATCATC
ACCACCATCACCATCATCATCACTAAGGATCCCAC

mCherry – codon optimised for *B. megaterium* (32% G+C)

CACCATATGGTATCTAAAGGTGAAGAAGATAATATGGCAATTATTAAAGAATTTATGCGTTTT
AAAGTACACATGGAAGGTTCTGTAAATGGTCATGAATTTGAAATTGAAGGTGAAGGTGAAG
GTCGTCCTTATGAAGGTACACAAACAGCAAAATTAAAAGTAACAAAAGGTGGTCCATTACCA
TTTGCATGGGATATTTTATCTCCACAATTCATGTATGGTTCTAAAGCGTATGTAAAACATCCA
GCAGATATTCCAGATTATTTAAAATTATCTTTTCCTGAAGGTTTTAAATGGGAACGTGTAATG
AATTTTGAAGATGGTGGTGTAGTAACAGTAACACAAGATTCTAGTTTACAAGATGGTGAATT
TATTTATAAAGTAAAATTACGTGGTACAAATTTTCCATCTGATGGTCCAGTAATGCAAAAAAA
AACAATGGGTTGGGAAGCATCTTCTGAACGTATGTATCCTGAAGATGGTGCATTAAAAGGT
GAAATTAAACAACGTTTAAAATTAAAAGATGGTGGTCATTATGATGCAGAAGTAAAAACAACA
TATAAAGCAAAAAAACCAGTACAATTACCAGGTGCATATAATGTAAATATTAAATTAGATATT
ACATCTCATAATGAAGATTATACAATTGTAGAACAATATGAACGTGCAGAAGGTCGTCATTC
TACAGGTGGTATGGATGAATTATATAAATAAGGATCCCAC

mCherry – codon optimised for *Streptomyces* (64% G+C)

CACCATATGGTGTCCAAGGGCGAGGAGGACAACATGGCCATCATCAAGGAGTTCATGCGC
TTCAAGGTCCACATGGAGGGCTCCGTCAACGGCCACGAGTTCGAGATCGAGGGCGAGGG
CGAGGGCCGCCCGTACGAGGGCACCCAGACCGCCAAGCTGAAGGTCACCAAGGGCGGCC
CGCTGCCGTTCGCCTGGGACATCCTGTCCCCGCAGTTCATGTACGGCTCCAAGGCCTACG
TCAAGCACCCGGCCGACATCCCGGACTACCTGAAGCTGTCCTTCCCCGAGGGCTTCAAGT
GGGAGCGCGTCATGAACTTCGAGGACGGCGGCGTCGTCACCGTCACCCAGGACTCCTCC
CTGCAGGACGGCGAGTTCATCTACAAGGTCAAGCTGCGCGGCACCAACTTCCCGTCCGAC
GGCCCGGTCATGCAGAAGAAGACCATGGGCTGGGAGGCCTCGTCCGAGCGCATGTACCC
CGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGCGCCTGAAGCTGAAGGACGGCGGCC
ACTACGACGCCGAGGTCAAGACCACCTACAAGGCCAAGAAGCCGGTCCAGCTGCCGGGC
GCCTACAACGTCAACATCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTCG
AGCAGTACGAGCGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAG
TGAGGATCCCAC

RepB-TetA shuttle backbone – originally derived from pWH1520

A BsaI site located in the *repB* origin of replication was removed by silent mutagenesis by the method of Sawano *et al* (25). Mutagenesis oligonucleotide is listed in Table S3.

GGTCTCATCGACCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATA
GGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCCTGTTATAAAAAAAGGATCAATTTTGA
ACTCTCTCCCAAAGTTGATCCCTTAACGATTTAGAAATCCCTTTGAGAATGTTTATATACATT
CAAGGTAACCAGCCAACTAATGACAATGATTCCTGAAAAAAGTAATAACAAATTACTATACA
GATAAGTTGACTGATCAACTTCCATAGGTAACAACCTTTGATCAAGTAAGGGTATGGATAAT
AAACCACCTACAATTGCAATACCTGTTCCCTCTGATAAAAAGCTGGTAAAGTTAAGCAAACT
CATTCCAGCACCAGCTTCCTGCTGTTTCAAGCTACTTGAAACAATTGTTGATATAACTGTTTT
GGTGAACGAAAGCCCACCTAAAACAAATACGATTATAATTGTCATGAACCATGATGTTGTTT
CTAAAAGAAAGGAAGCAGTTAAAAAGCTAACAGAAAGAAATGTAACTCCGATGTTTAACACG
TATAAAGGACCTCTTCTATCAACAAGTATCCCACCAATGTAGCCGAAAATAATGACACTCAT
TGTTCCAGGGAAAATAATTACACTTCCGATTTCGGCAGTACTTAGCTGGTGAACATCTTTCA
TCATATAAGGAACCATAGAGACAAACCCTGCTACTGTTCCAAATATAATTCCCCCACAAAGA
ACTCCAATCATAAAAGGTATATTTTTCCCTAATCCGGGATCAACAAAAGGATCTGTTACTTTC
CTGATATGTTTTACAAATATCAGGAATGACAGCACGCTAACGATAAGAAAAGAAATGCTATA
TGATGTTGTAAACAACATAAAAAATACAATGCCTACAGACATTAGTATAATTCCTTTGATATC
AAAATGACCTTTTATCCTTACTTCTTTCTTTAATAATTTCATAAGAAACGGAACAGTGATAATT
GTTATCATAGGAATGAGTAGAAGATAGGACCAATGAATATAATGGGCTATCATTCCACCAAT
CGCTGGACCGACTCCTTCTCCCATGGCTACTATCGATCCAATAAGACCAAATGCTTTACCC
CTATTTTCCTTTGGAATATAGCGCGCAACTACAACCATTACGAGTGCTGGAAATGCAGCTGC
ACCAGCCCCTTGAATAAAACGAGCCATAATAAGTAAGGAAAAGAAAGAATGGCCAACAAAC
CCAATTACCGACCCGAAACAATTTATTATAATTCCAAATAGGAGTAACCTTTTGATGCCTAAT
TGATCAGATAGCTTTCCATATACAGCTGTTCCAATGGAAAAGGTTAACATAAAGGCTGTGTT
CACCCAGTTTGTACTCGCAGGTGGTTTATTAAAATCATTTGCAATATCAGGTAATGAGACGT
TCAAAACCATTTCATTTAATACGCTAAAAAAAGATAAAATGCAAAGCCAAATTAAAATTTGGT
TGTGTCGTAAATTCGATTGTGAATAGGATGTATTCACATTTCACCCTCCAATAATGAGGGCA
GACGTAGTTTATAGGGTTAATGATACGCTTCCCTCTTTTAATTGAACCCTGTTACATTCATTA
CACTTCATAATTAATTCCTCCTAAACTTGATTAAAACATTTTACCACATATAAACTAAGTTTTA
AATTCAGTATTTCATCACTTATACAACAATATGGCCCGTTTGTTGAACTACTCTTTAATAAAAT
AATTTTTCCGTTCCCAATTCCACATTGCAATAATAGAAAATCCATCTTCATCGGCTTTTTCGT
CATCATCTGTATGAATCAAATCGCCTTCTTCTGTGTCATCAAGGTTTAATTTTTTATGTATTTC
TTTTAACAAACCACCATAGGAGATTAACCTTTTACGGTGTAAACCTTCCTCCAAATCAGACAA
ACGTTTCAAATTCTTTTCTTCATCATCGGTCATAAAATCCGTATCCTTTACAGGATATTTTGC
AGTTTCGTCAATTGCCGATTGTATATCCGATTTATATTTATTTTTCGGTCGAATCATTTGAACT
TTTACATTTGGATCATAGTCTAATTTCATTGCCTTTTTCCAAAATTGAATCCATTGTTTTTGAT
TCACGTAGTTTTCTGTATTCTTAAAATAAGTTGGTTCCACACATACCAATACATGCATGTGCT
GATTATAAGAATTATCTTTATTATTTATTGTCACTTCCGTTGCACGCATAAAACCAACAAGATT
TTTATTAATTTTTTTATATTGCATCATTCGGCGAAATCCTTGAGCCATATCTGACAAACTCTTA
TTTAATTCTTCGCCATCATAAACATTTTTAACTGTTAATGTGAGAAACAACCAACGAACTGTT
GGCTTTTGTTTAATAACTTCAGCAACAACCTTTTGTGACTGAATGCCATGTTTCATTGCTCTC
CTCCAGTTGCACATTGGACAAAGCCTGGATTTACAAAACCACACTCGATACAACTTTCTTTC
GCCTGTTTCACGATTTTGTTTATACTCTAATATTTCAGCACAATCTTTTACTCTTTCAGCCTTT
TTAAATTCAAGAATATGCAGAAGTTCAAAGTAATCAACATTAGCGATTTTCTTTTCTCTCCAT
GGTCTGACTTTTCCACTTTTTGTCTTGTCCACTAAAACCCTTGATTTTTCATCTGAATAAATG
CTACTATTAGGACACATAATATTAAAAGAAACCCCCATCTATTTAGTTATTTGTTTGGTCACT

TATAACTTTAACAGATGGGGTTTTTCTGTGCAACCAATTTTAAGGGTTTTCAATACTTTAAAA
CACATACATACCAACACTTCAACGCACCTTTCAGCAACTAAAATAAAAATGACGTTATTTCTA
TATGTATCAAGATAAGAAAGAACAAGTTCAAAACCATCAAAAAAAGACACCTTTTCAGGTGC
TTTTTTTATTTTATAAACTCATTCCCTGATCTCGACTTCGTTCTTTTTTTACCTCTCGGTTATG
AGTTAGTTCAAATTCGTTCTTTTTAGGTTCTAAATCGTGTTTTTCTTGGAATTGTGCTGTTTTA
TCCTTTACCTTGTCTACAAACCCCTTAAAAACGTTTTTAAAGGCTTTTAAGCCGTCTGTACGT
TCCTTAAGGAAGGTGTTAGAGACC

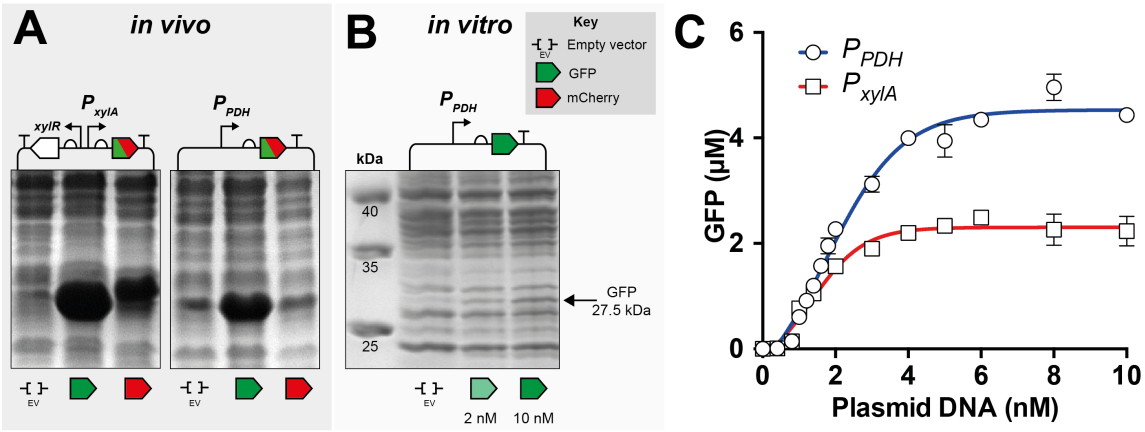


Fig. S1. Recombinant protein synthesis in *B. megaterium*. (*A*) SDS-PAGE of GFP (green arrow) and mCherry (red arrow) protein synthesised under the control of the xylose-inducible promoter or PDH constitutive promoter. An empty vector control (EV) – pMM1520, was used for comparison. (*B*) SDS-PAGE of NCF extracts of *B. megaterium* with GFP synthesised from the PDH promoter. (*C*) Time-course reaction of GFP synthesis from the xylose-inducible promoter and PDH constitutive promoter in NCF reactions with 10 nM of plasmid DNA. Error bars are representative of three technical repeats.

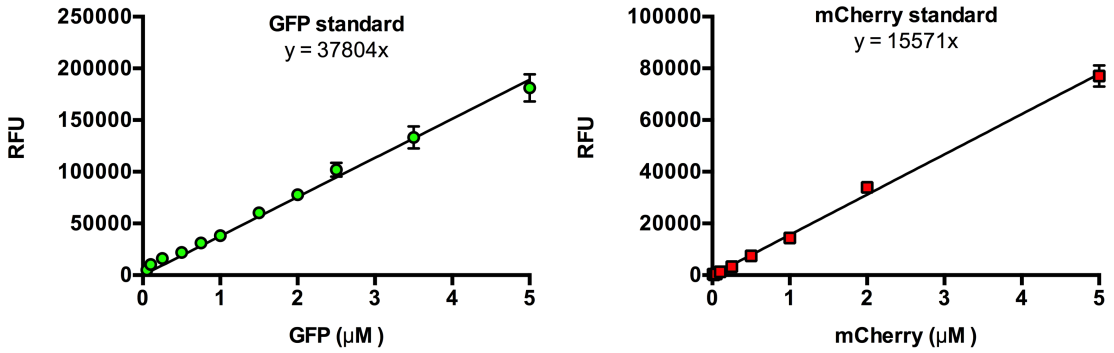


Fig. S2. Fluorescence calibration standards of purified His_6-tagged GFP and mCherry. Purified proteins were mixed with 10 mg mL^{-1} *B. megaterium* DSM319 extract, 3-PGA energy mix, amino acids, 6 mM Mg-glutamate, 50 mM K-glutamate and 2% PEG8K and pre-incubated at 30°C for 10 min.

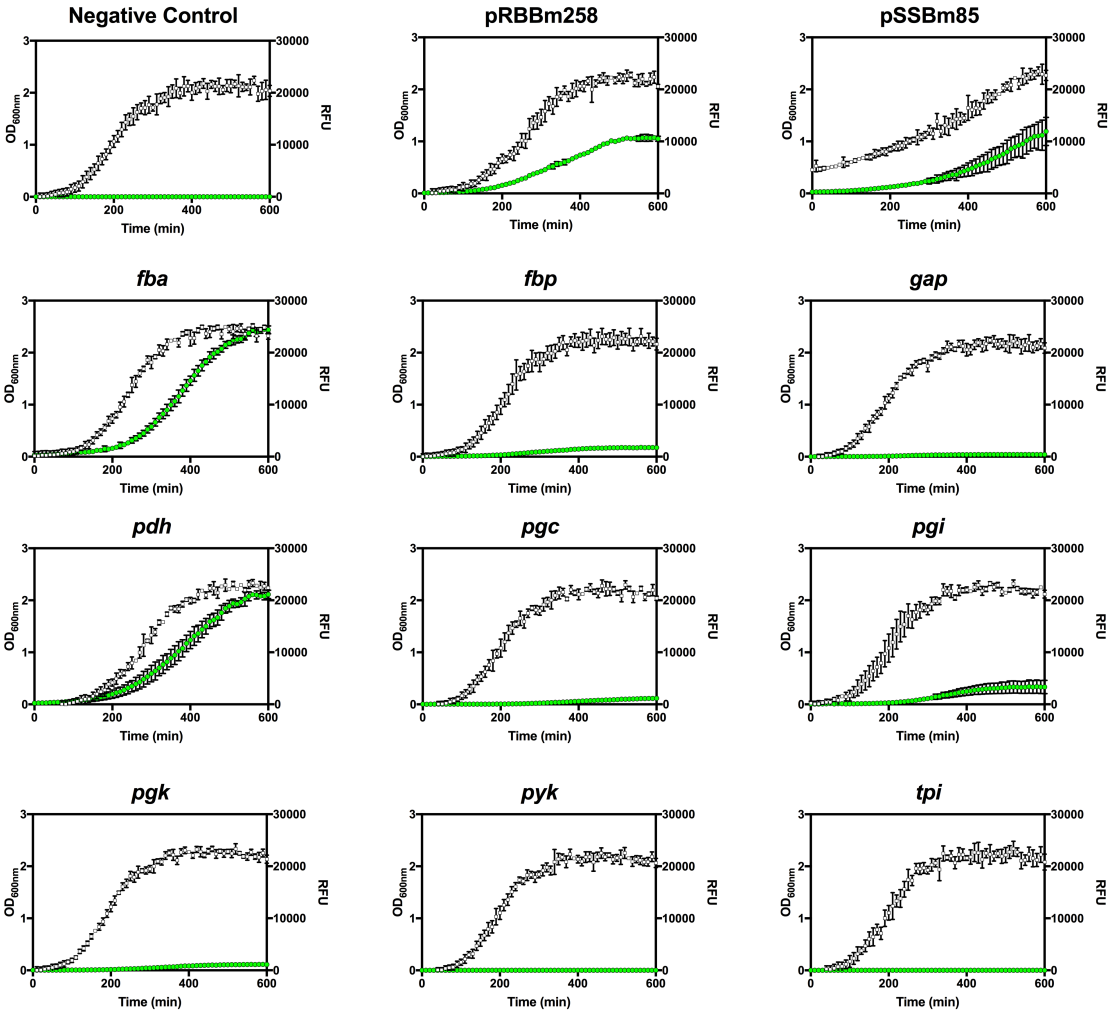


Fig. S3. Growth and GFP synthesis of promoter plasmid variants of *B. megaterium* DSM319 in 96-well plate cultures. Plasmid strains (controls, *fba*, *fbp*, *gap*, *pdh*, *pgc*, *pgi*, *pgk*, *pyk* and *tpi*) were grown as described in supporting materials and methods, then sub-cultured into fresh medium to an OD_{600} of 0.02 and grown for 10 hours at 30°C and 600 rpm shaking. The pRBm258 (P_{pdh}-GFP) and pSSBm85 (P_{xylA}-GFP) provide relative positive controls. The pSSBm85 plasmid strain was sub-cultured to an OD_{600} of 0.4 and induced with 0.25% (w/v) D-xylose. White squares represent OD_{600} and are plotted on the left y-axis. Green circles represent GFP synthesis in relative fluorescence units (RFU) and are plotted on the right y-axis. Error bars are derived from four biological repeats, from single colony growth and sub-culturing.

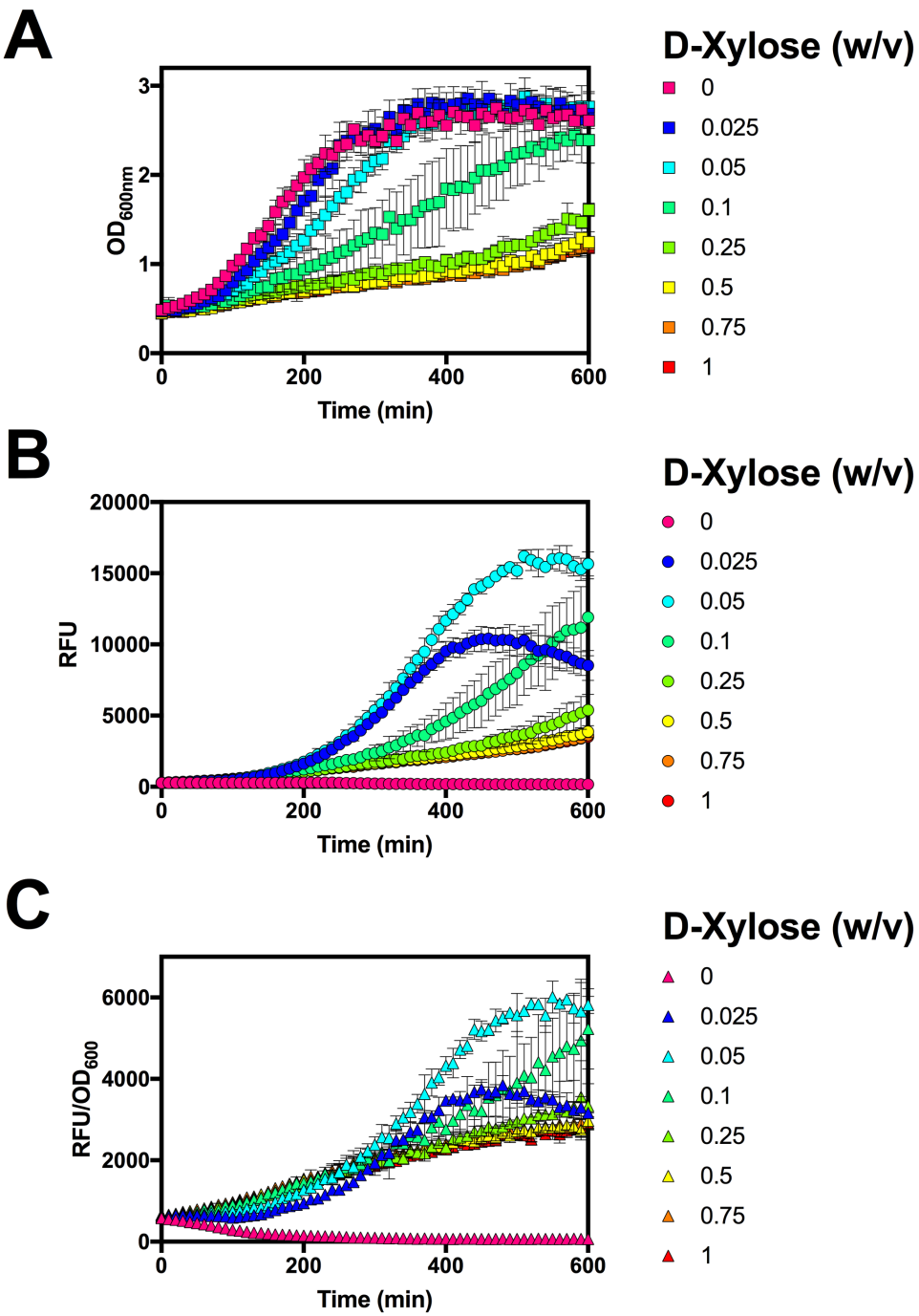


Fig. S4. Growth and GFP synthesis of the *B. megaterium* DSM319 pSSBm85 xylose-inducible promoter strain in a 96-well plate. pSSBm85 is a derivative of pKMMBm5 and contains the XylR repressor. Cultures were grown as described in supporting materials and methods, then sub-cultured into fresh medium to an OD_{600} of 0.4 with varying concentrations of xylose and grown for 10 hours at 30°C and 600 rpm shaking. (a) Squares represent OD_{600} growth. (b) Circles represent GFP synthesis in RFU. (c) Triangles represent RFU/OD_{600}. Error bars are derived from four biological repeats, from single colony growth and sub-culturing.

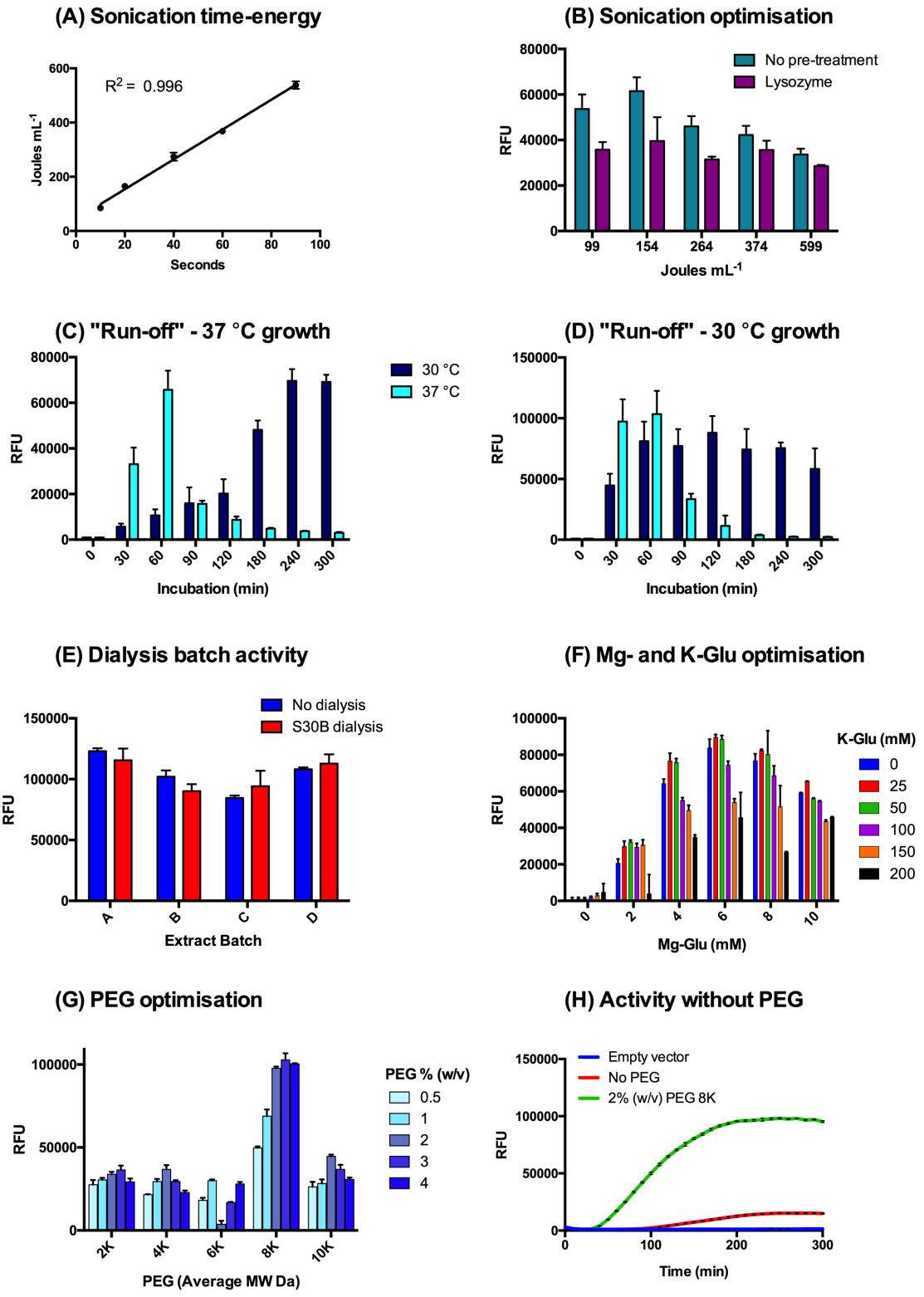


Fig. S5. Optimization of the *B. megaterium* DSM319 cell-free extract. See SI text for full details of extract preparation and testing.

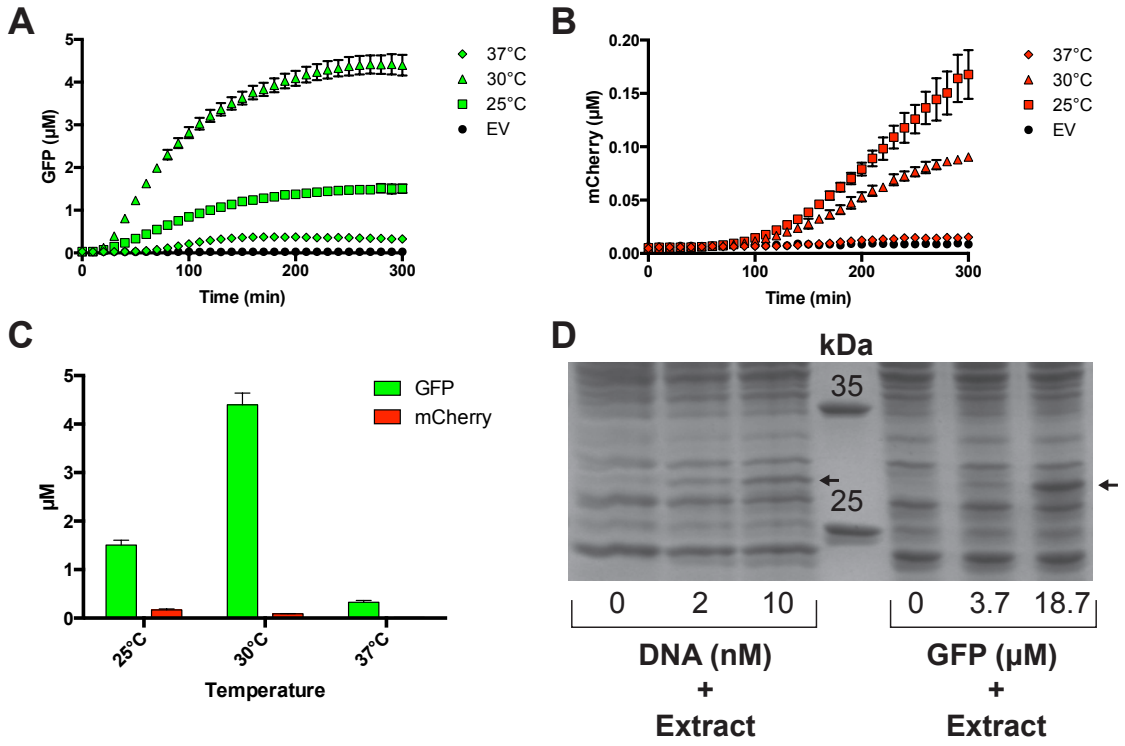


Fig. S6. GFP and mCherry production from the pRBBm258 and pRBBm267 plasmids under standard conditions with 3-PGA energy buffer. (a) 10 nM pRBBm258 incubated at 25, 30 and 37°C. (b) 10 nM pRBBm267 incubated at 25, 30 and 37°C. (c) End-point comparison of GFP and mCherry production at different temperatures. (d) SDS-PAGE analysis of pRBBm258 incubations in comparison to purified His_6-tagged GFP spiked into the extract. mCherry production was not significant and could not be observed. Empty vector (pMM1520) was used as a negative control for incubations. See supplementary text for further details. We noted a decreased yield of protein (Maximum 3.4 μM) where the *pdh* promoter (from pRBBm258) was cloned by EcoFlex with RiboJ, a standard RBS and GFP. The slight decrease in yield, in comparison to pRBBm258 may be due to differences in RBS and spacer sequences, with AGGAAGA-N_{10}-ATG (pRBBm258) and AGGAGGT-N_9-ATG (*pdh*-RiboJ-RBS-GFP-MGapt-Bba_B0015) used, respectively.

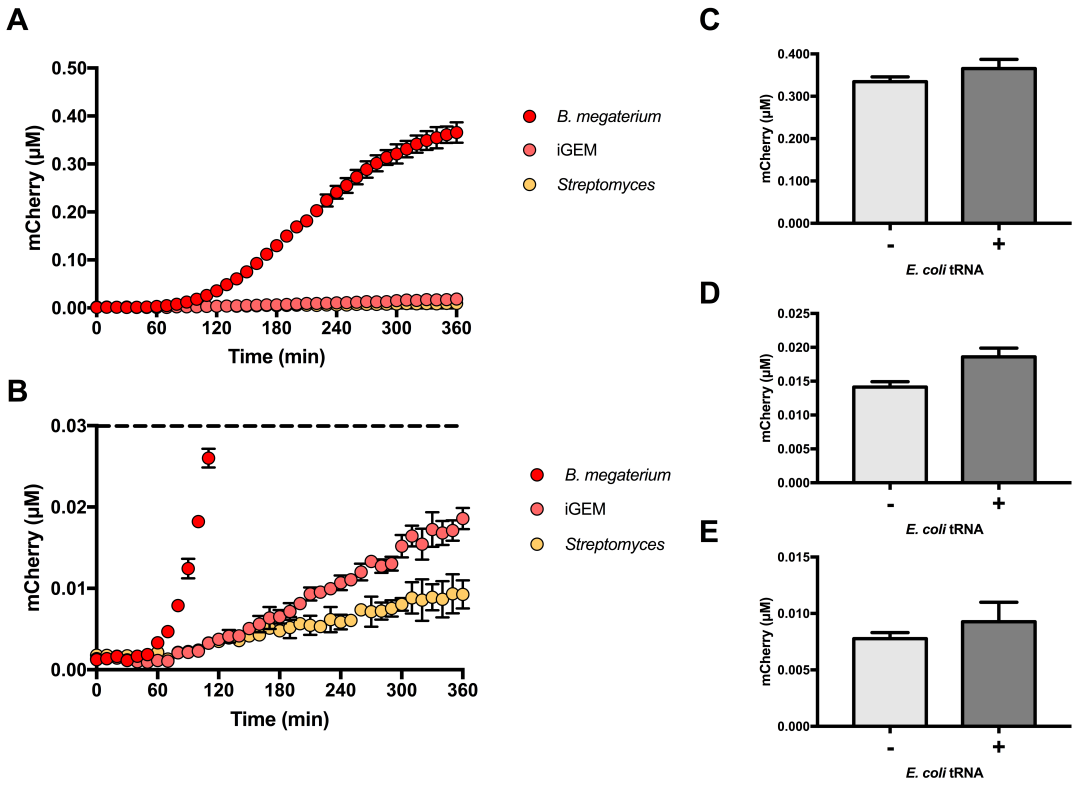


Fig. S7. mCherry production with different G+C (%) variants in the presence or absence of *E. coli* tRNA. (a) Time-course NCF reaction with 10 nM of pTU1-A-*pdh*-RiboJ-mCherry-Bba_B0015 variants. A *B. megaterium* codon-optimised (32% G+C) variant, iGEM (62% G+C from EcoFlex kit) and a *Streptomyces* (64% G+C) optimised variant were used for comparison. (b) Zoomed in view of iGEM and *Streptomyces* variants. (c) *B. megaterium*, (d) iGEM and (e) *Streptomyces* mCherry variants synthesised in the presence or absence of *E. coli* tRNA added into the 3-PGA energy buffer. Data is presented from three technical measurements.

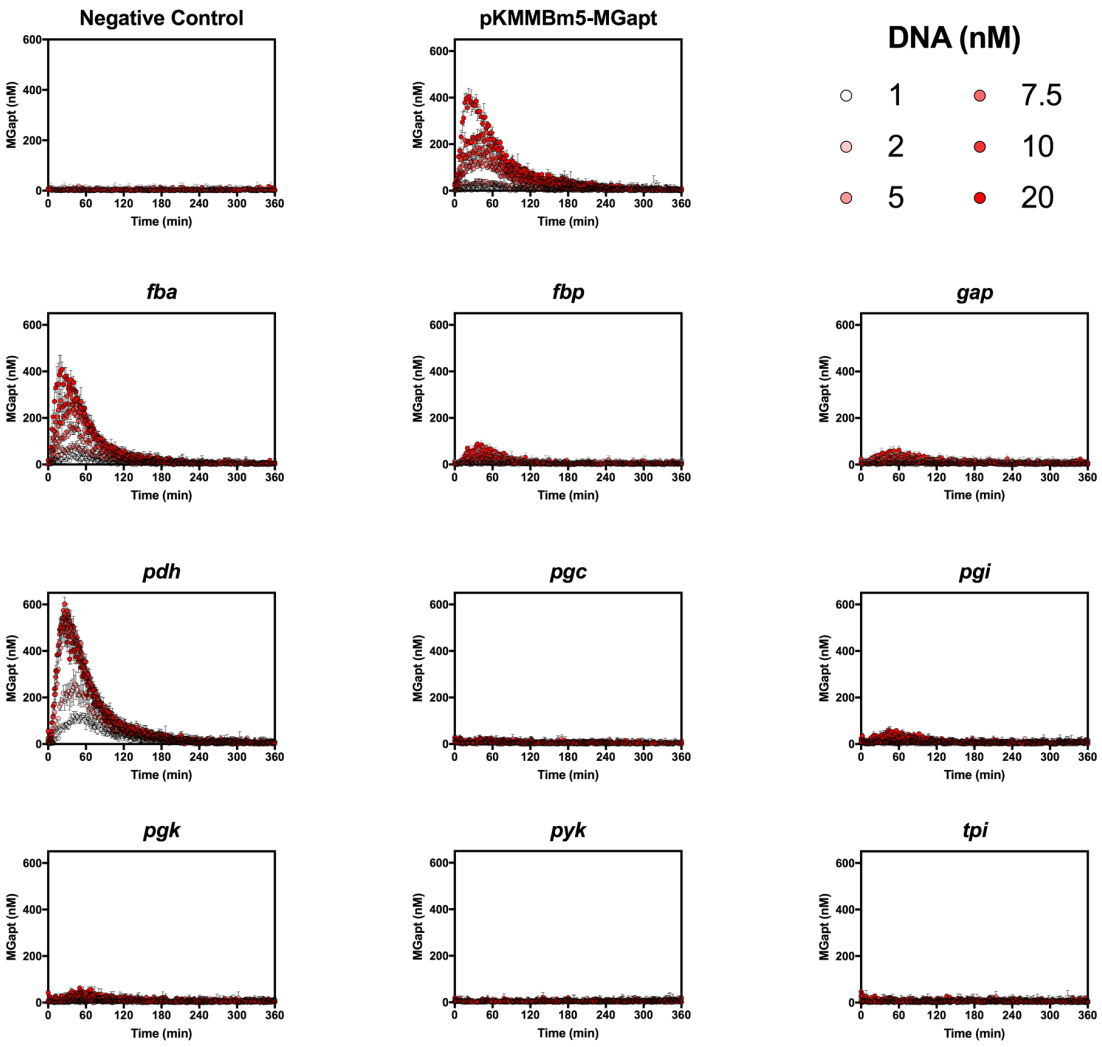


Fig. S8. RNA (nM) synthesis by the glycolysis and gluconeogenesis promoters with varying plasmid DNA (nM) concentration. pKMMBm5-MGapt was is provided as a relative comparison. In the negative control plasmid (P_L-RiboJ-GFP-MGapt-Bba_B0015), the promoter is replaced with TATATA promoter linker (P_L). Reactions were performed under standard conditions as three technical measurements.

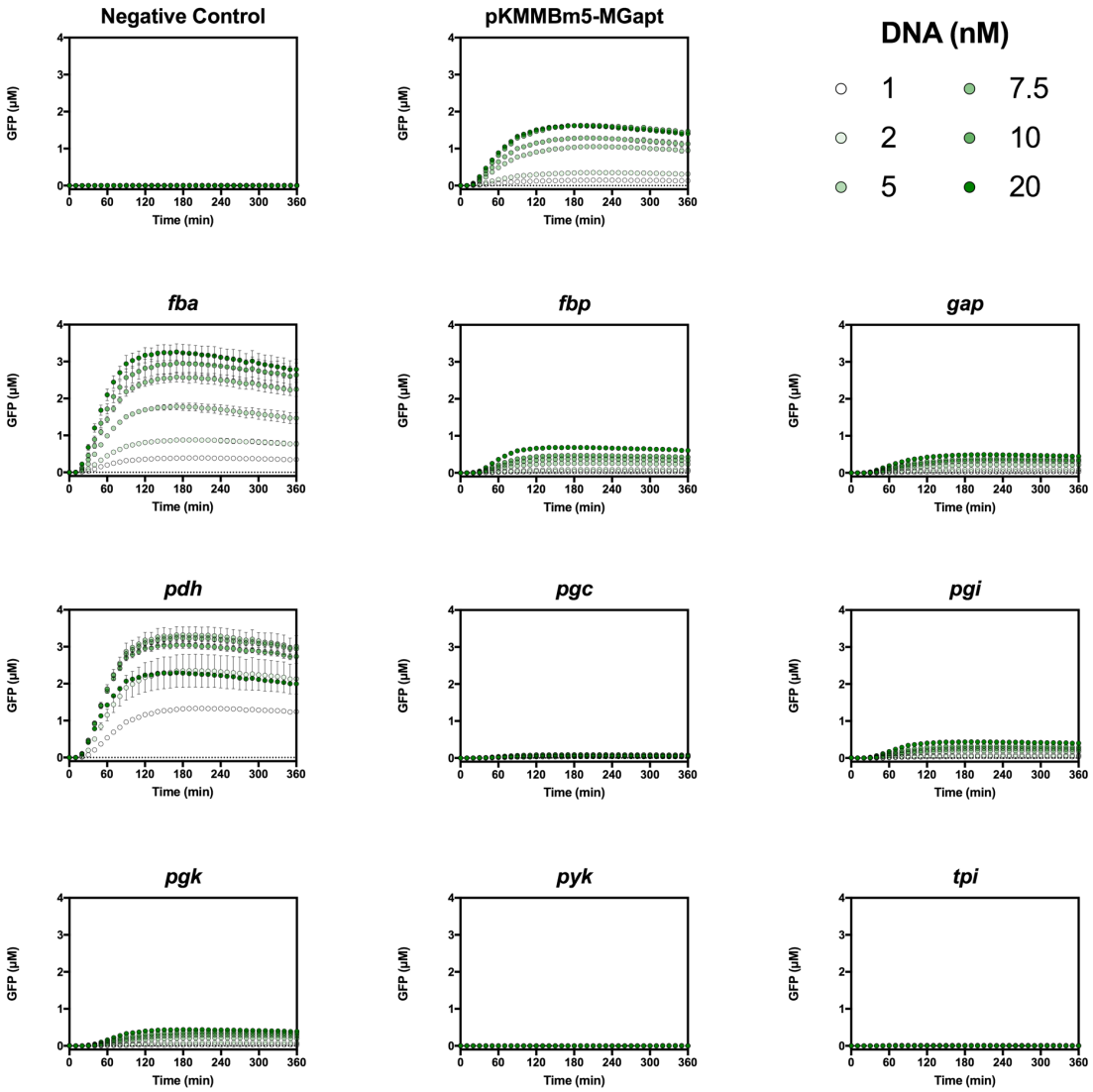


Fig. S9. GFP (μM) synthesis by the glycolysis and gluconeogenesis promoters with varying plasmid DNA (nM) concentration. pKMMBm5-MGapt was is provided as a relative comparison. In the negative control plasmid (P_L-RiboJ-GFP-MGapt-Bba_B0015), the promoter is replaced with TATATA promoter linker (P_L). Reactions were performed under standard conditions as three technical measurements.

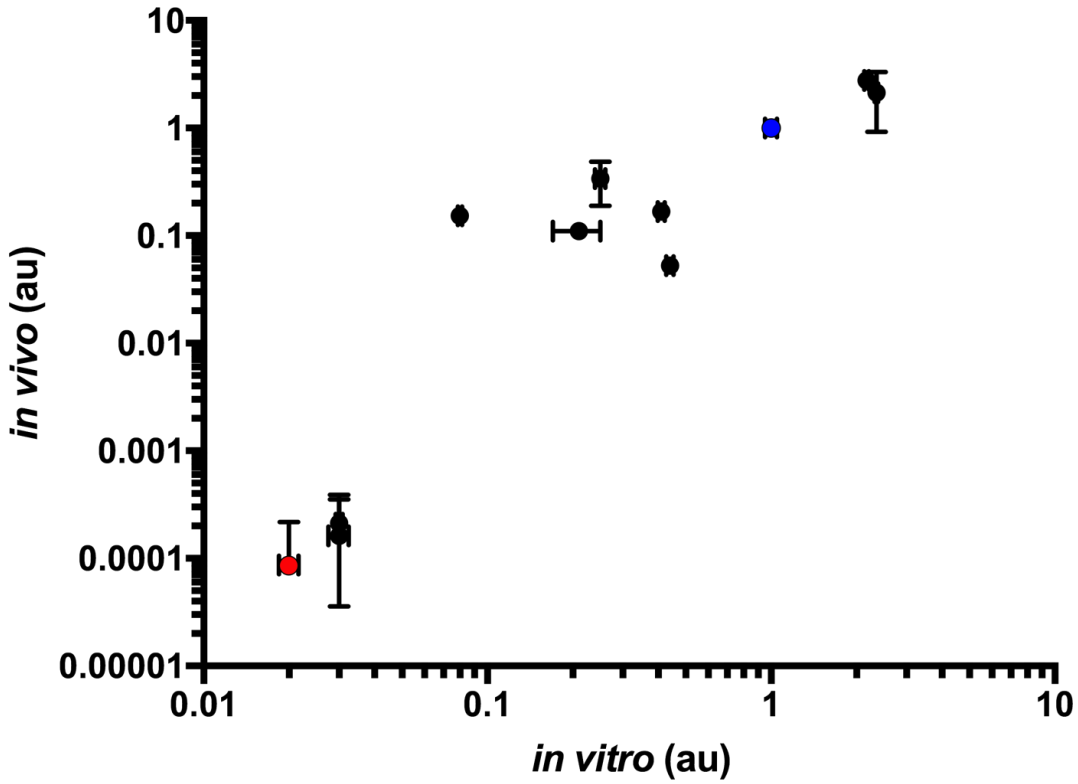


Fig. S10. Normalised *in vivo* and cell-free constitutive promoter activity relative to the strongest promoter (*pdh*). Correlation coefficient is 0.937 and P value = 0.0001. Negative control (red circle) and positive control (blue circle) data points are highlighted for comparison. Data is represented as normalised steady-state value relative to the positive control plasmid (au - arbitrary units).

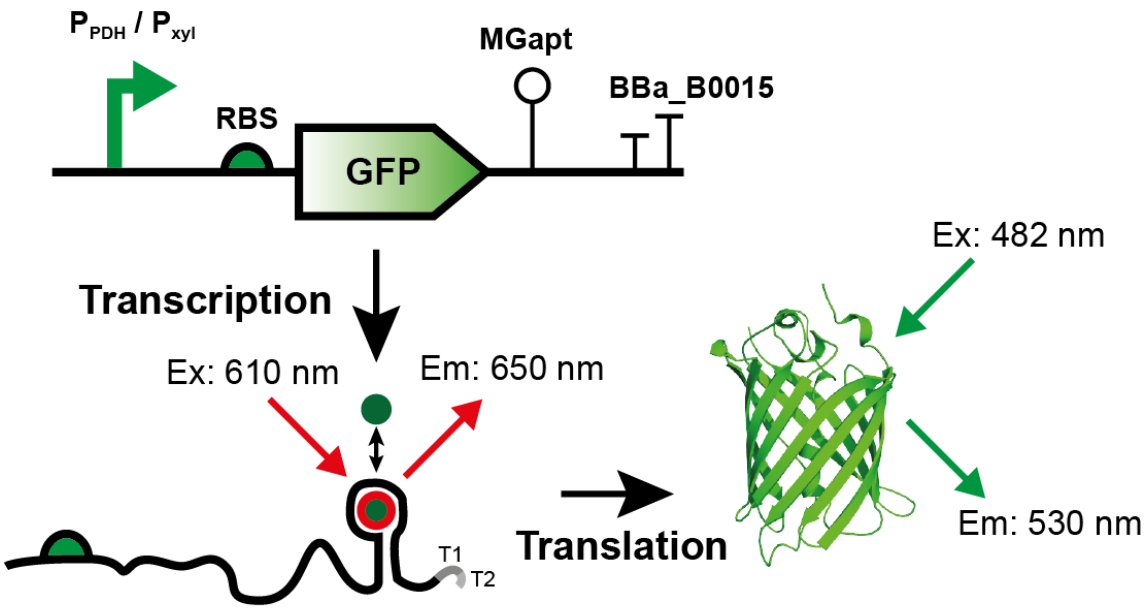


Fig. S11. Illustration of the GFP and malachite aptamer plasmids for cell-free transcription-translation in *B. megaterium* DSM319. Excitation and emission are abbreviated as Ex and Em, respectively. The transcript was detected as two bands of similar size (~920 and 950 bp, respectively) encoding the RBS, GFP and malachite green aptamer, representing two alternative transcription termination sites (T1 and T2) due to the presence of the Rho-independent double terminator Bba_B0015, which is composed of the Bba_B0010 and Bba_B0012 terminators.

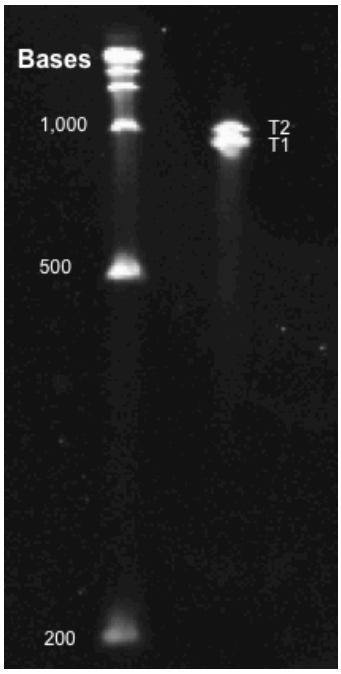


Fig. S12. Denaturing 5% TBE-urea PAGE gel of the GFP-MGapt transcript. Two bands of similar size (~920 bp) encoding the RBS, GFP and malachite green aptamer were detected (T1 and T2), representing the two alternative termination sites after the malachite green aptamer from the double terminator Bba_B0015, composed of joined Bba_B0010 and Bba_B0012 terminators.

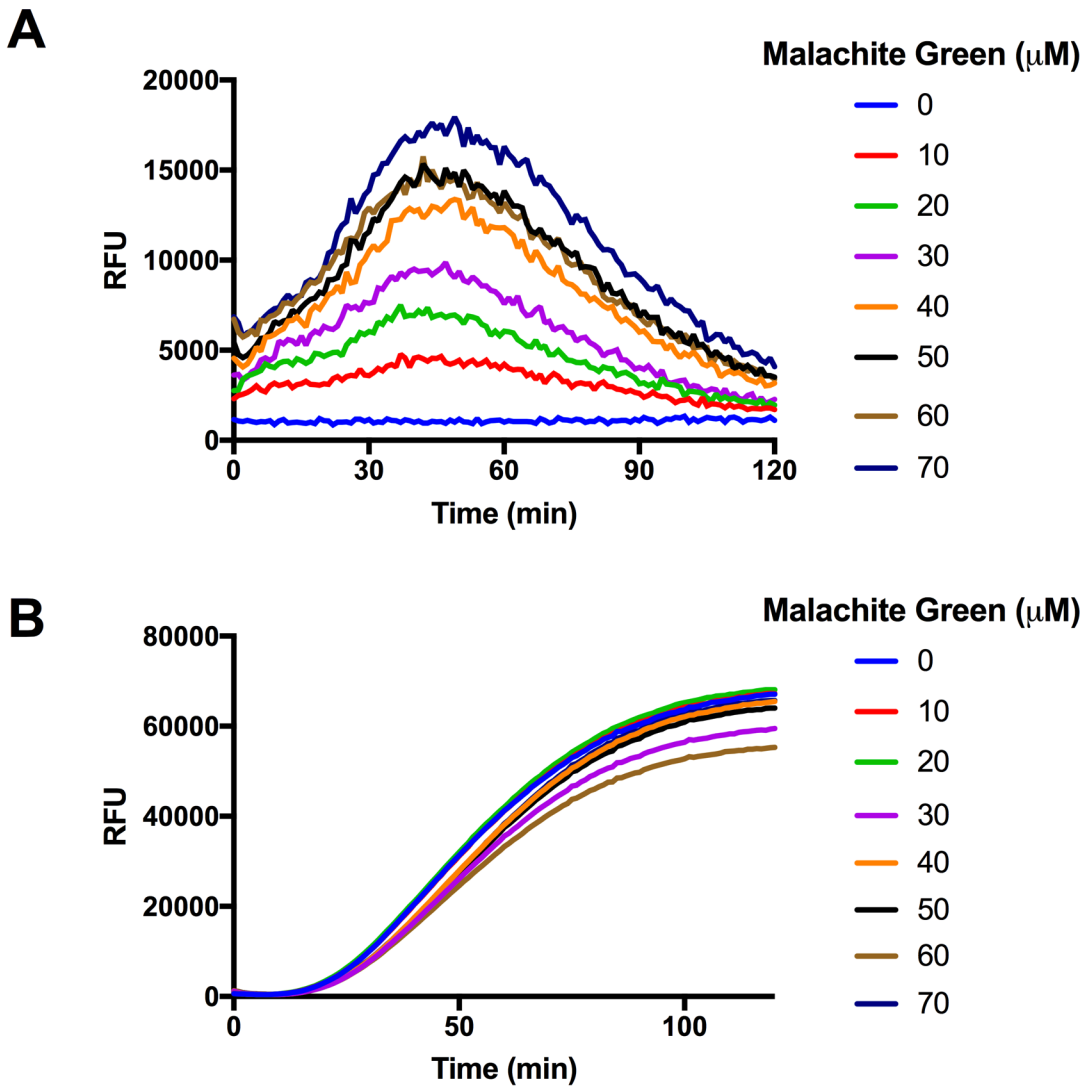


Fig. S13. Optimising the malachite green concentration for real-time (A) mRNA and (B) GFP synthesis in *B. megaterium* NCF. 10 nM of pKMMBm5-MGapt plasmid was used under standard conditions. Data is presented from three technical measurements, although error bars are removed for clarity between concentrations. Malachite green did not inhibit protein synthesis up to 70 μM.

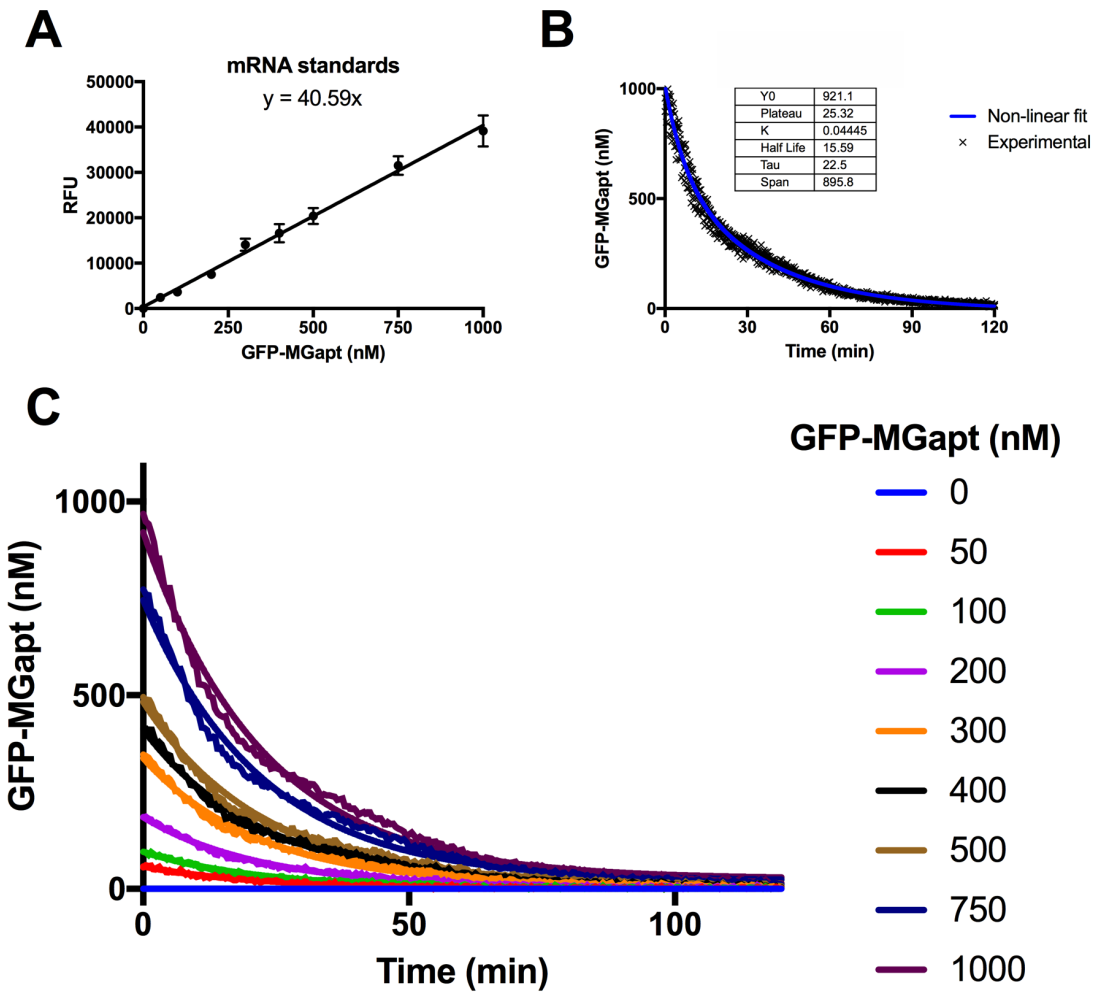


Fig. S14. mRNA standards and decay for the malachite green aptamer. (a) Purified mRNA standards were prepared in 10 mg mL^{-1} of extract with 3-PGA energy buffer, amino acids, 2% PEG8K, 6 mM Mg-glutamate, 50 mM K-glutamate and 25 μM malachite green. Fluorescent measurements (Ex: 610-10, Em: 650-10) were recorded every 30 s for 2 hrs. Standard curve of time zero fluorescence measurements of GFP-MGapt concentrations. (b) Decay rate in cell-free extract. Experimental data-points (black crosses) and a non-linear fit (blue line) of half-life was estimated based on a single-phase exponential decay function of $Y = (Y_0 - P)e^{-k\tau} + P$, where Y is mRNA concentration, Y_0 is the mRNA concentration at time zero, P is mRNA concentration at infinity, κ is the rate constant and τ is the time constant. (c) Decay rate of a range of GFP-MGapt concentrations from 50-1000 nM. Data is presented from three technical measurements

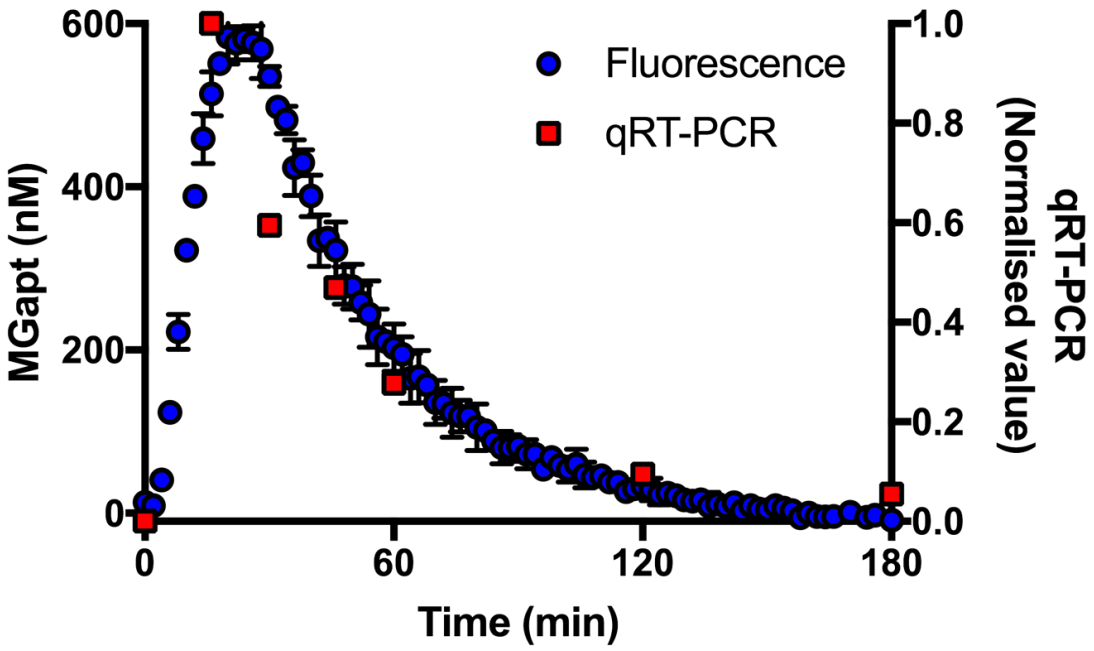


Fig. S15. qRT-PCR time-course measurement of the GFP-MGapt transcript in comparison to real-time MGapt fluorescence. Data was analysed as described in the supporting methods.

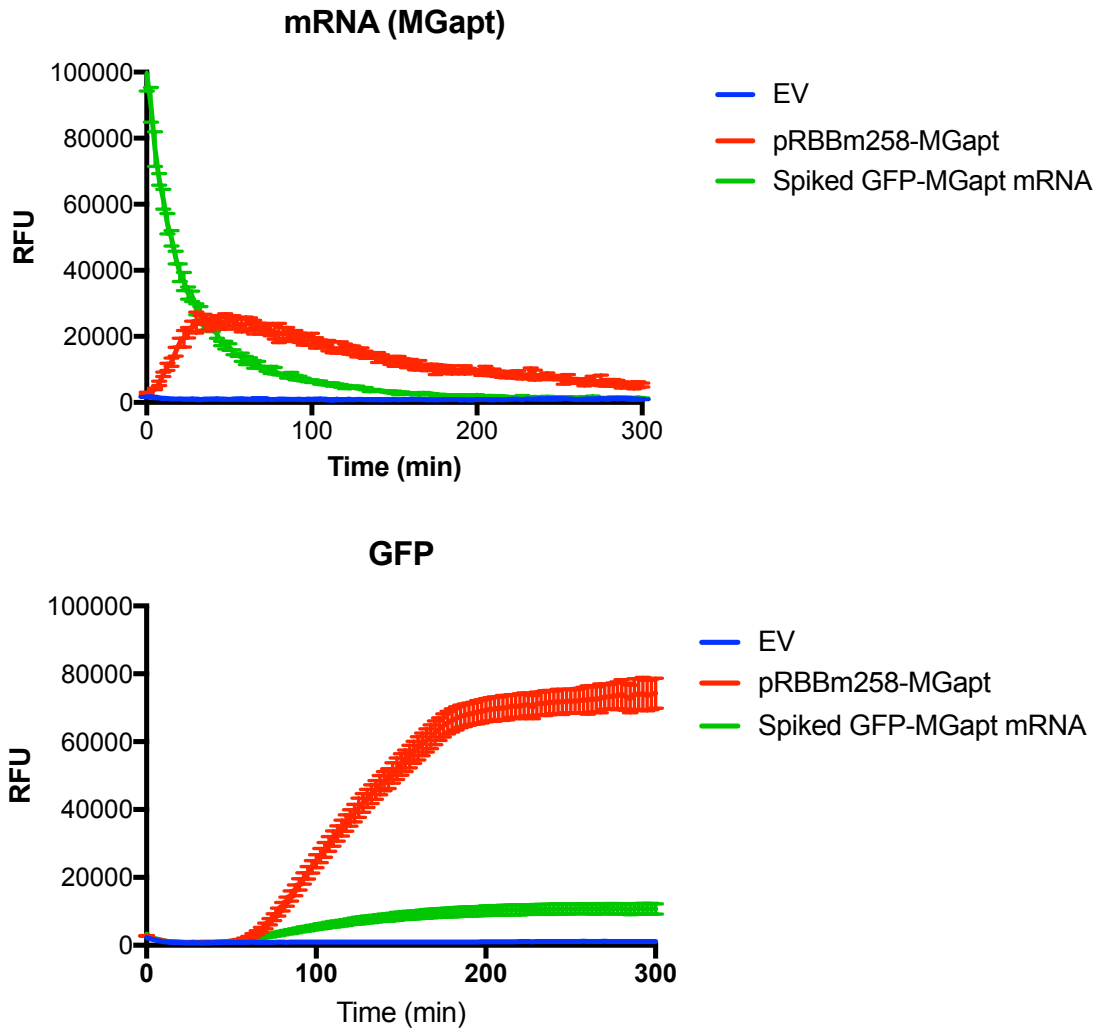


Fig. S16. Incubation of pure mRNA GFP-MGapt transcript into the cell-free reaction. Pre-spiked mRNA (green line) with 10 mg mL^{-1} *B. megaterium* DSM319 extract, 3-PGA energy mix, amino acids, 6 mM Mg-glutamate, 50 mM K-glutamate and 2% PEG8K. Empty vector (EV – blue line) and 10 nM pRBBm258-MGapt (red line) were run as controls for comparison. Reactions were performed under standard conditions. Data is presented from three technical measurements.

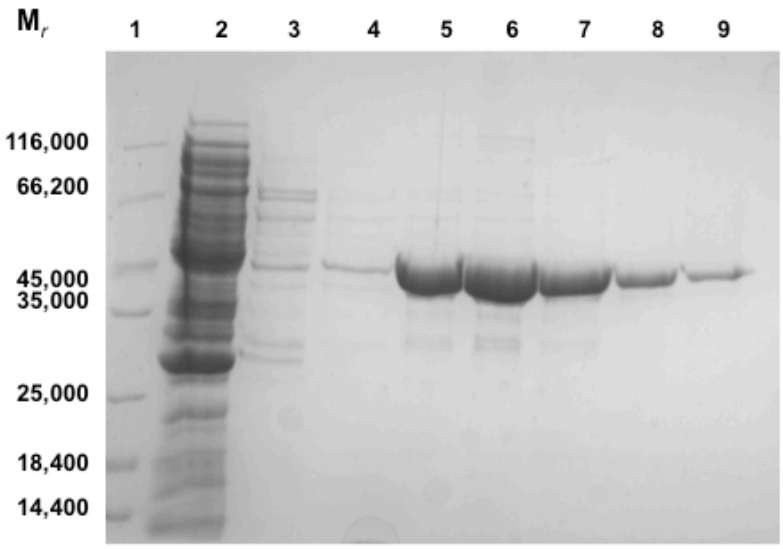


Fig. S17. Purification of XylR^{His} recombinantly produced in *E. coli* BL21 (DE3) pLysS. Purification fractions were analyzed with a 12 % SDS-PAGE gel. The gel was stained with Coomassie Brilliant Blue G250. Lane 1: Unstained Protein Molecular Weight Marker (Fermentas), Lane 2: soluble extract, Lane 3: 50 mM imidazole wash, Lane 4: 100 mM imidazole wash and Lanes 5-9: 400 mM imidazole elution fractions.

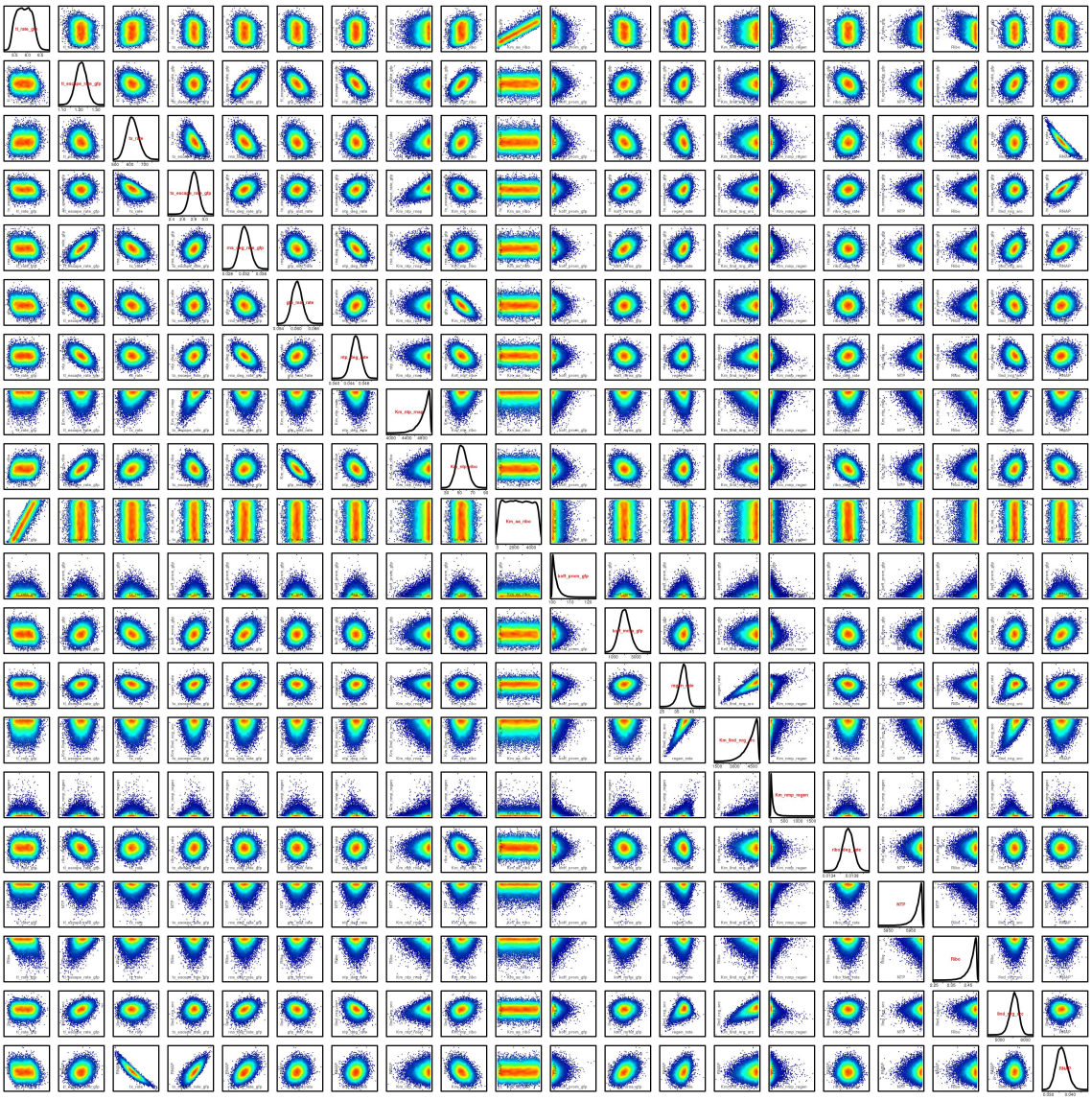


Fig. S18. Posterior distribution represented as univariate and bivariate marginal distributions after MCMC Bayesian model parameter inference with data from the pKMMBm5 plasmid titration experiment (Main text Fig. 2; see Table S9 for details on the model parameters). The posterior distribution is a complex multidimensional probability distribution of the model parameters inferred from the information contained in the experimental data (Main text Fig. 2). The diagonal histograms show the univariate marginal posterior distribution for each of the model parameters. The off-diagonal plots show the bivariate marginal posterior distributions, where each point is a sample from posterior distribution and the color indicates the density of points. These off-diagonal plots can indicate correlations between different parameters (e.g. tx_rate is negatively correlated with RNAP concentration as would be expected, see plot on bottom row, column three). The axes of the off-diagonal plots are not shown for reasons of space but have the same scales as the corresponding diagonal plots. Parameter units are the same as those in Table S9.

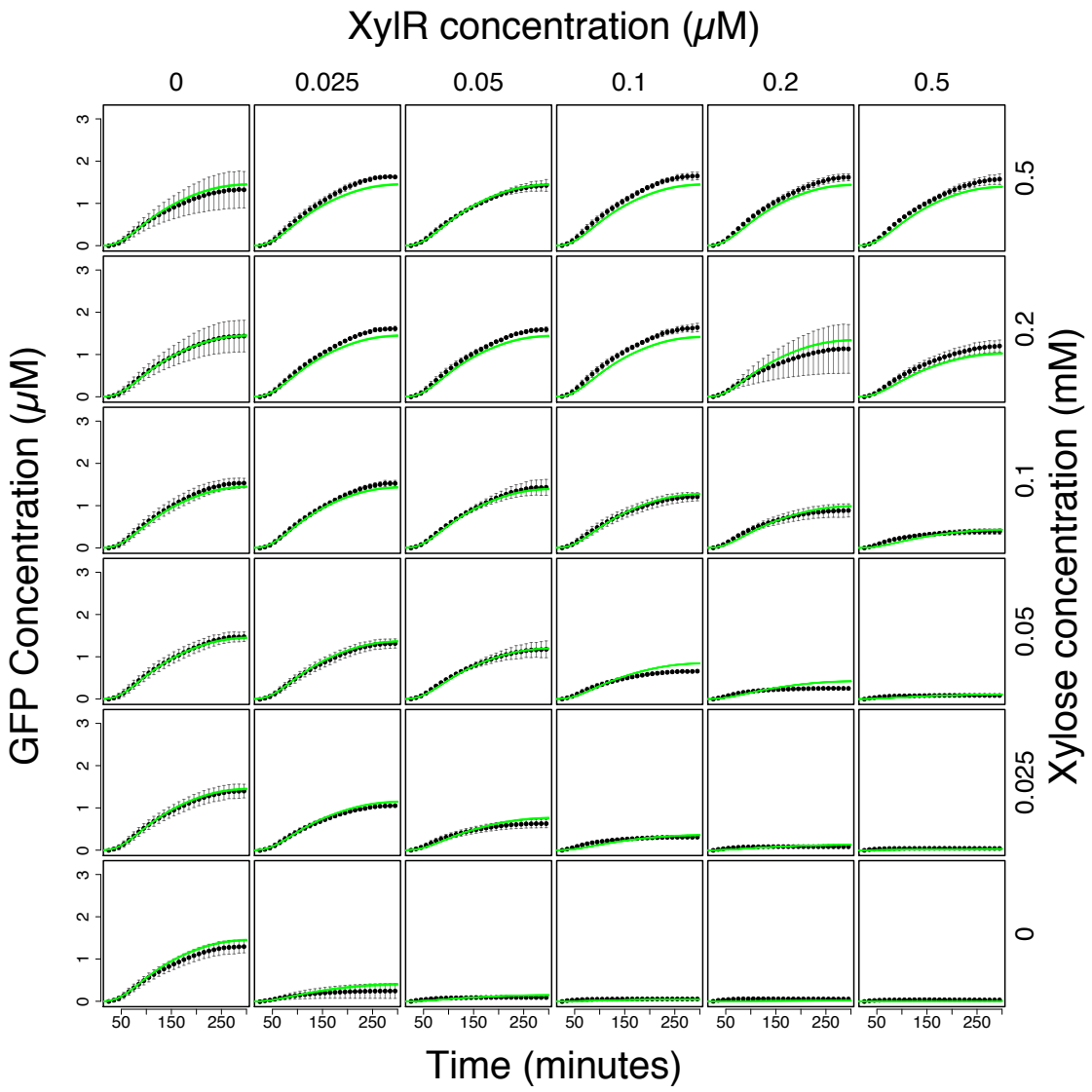


Fig. S19. Time-course experimental and simulated data of the xylose-XylR experiment (Main text Fig. 3) with 1 nM pKMMBm5 plasmid. The experimental trajectories are shown in black (where the error bars indicate standard error) and 100 different simulated trajectories are shown in green with parameter sets randomly selected from the posterior distribution (Fig. S22).

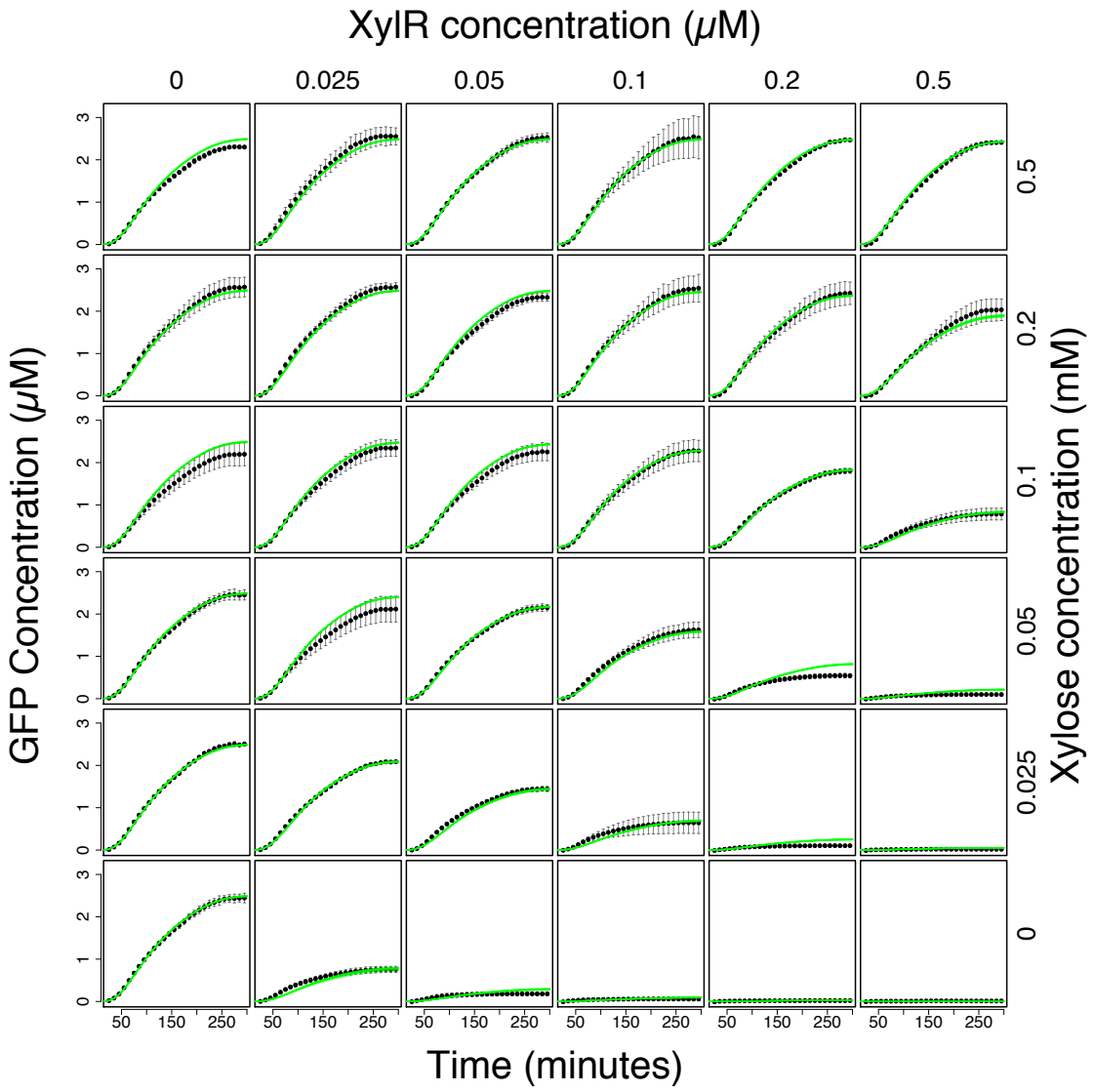


Fig. S20. Time-course experimental and simulated data of the xylose-XylR experiment (Main text Fig. 3) with 2 nM pKMMBm5 plasmid. The experimental trajectories are shown in black (where the error bars indicate standard error) and 100 different simulated trajectories are shown in green with parameters randomly selected from the posterior distribution (Fig. S22).

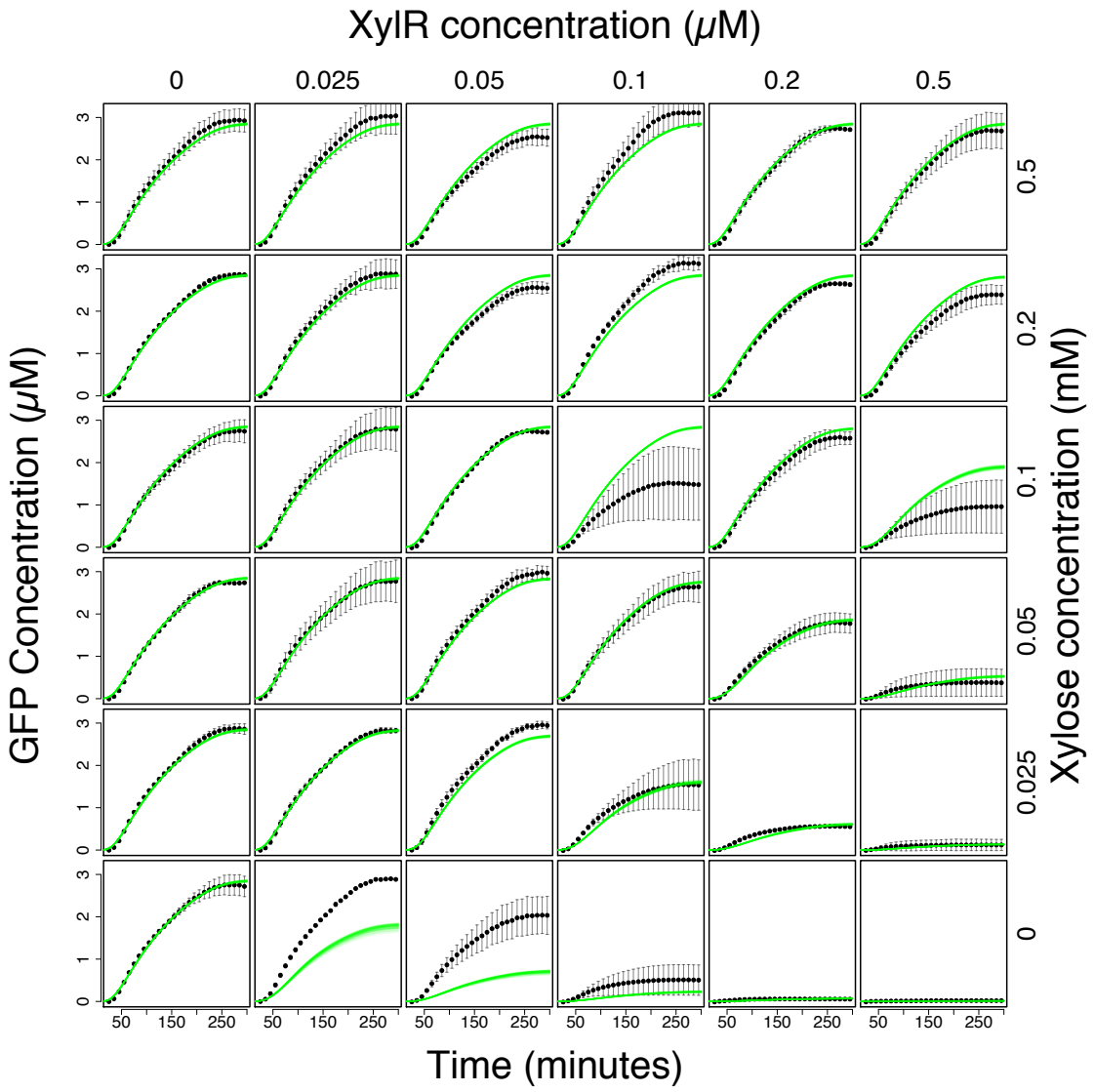


Fig. S21. Time-course experimental and simulated data of the xylose-XylR experiment (Main text Fig. 3) with 5 nM pKMMBm5 plasmid. The experimental trajectories are shown in black (where the error bars indicate standard error) and 100 different simulated trajectories are shown in green with parameters randomly selected from the posterior distribution (Fig. S22).

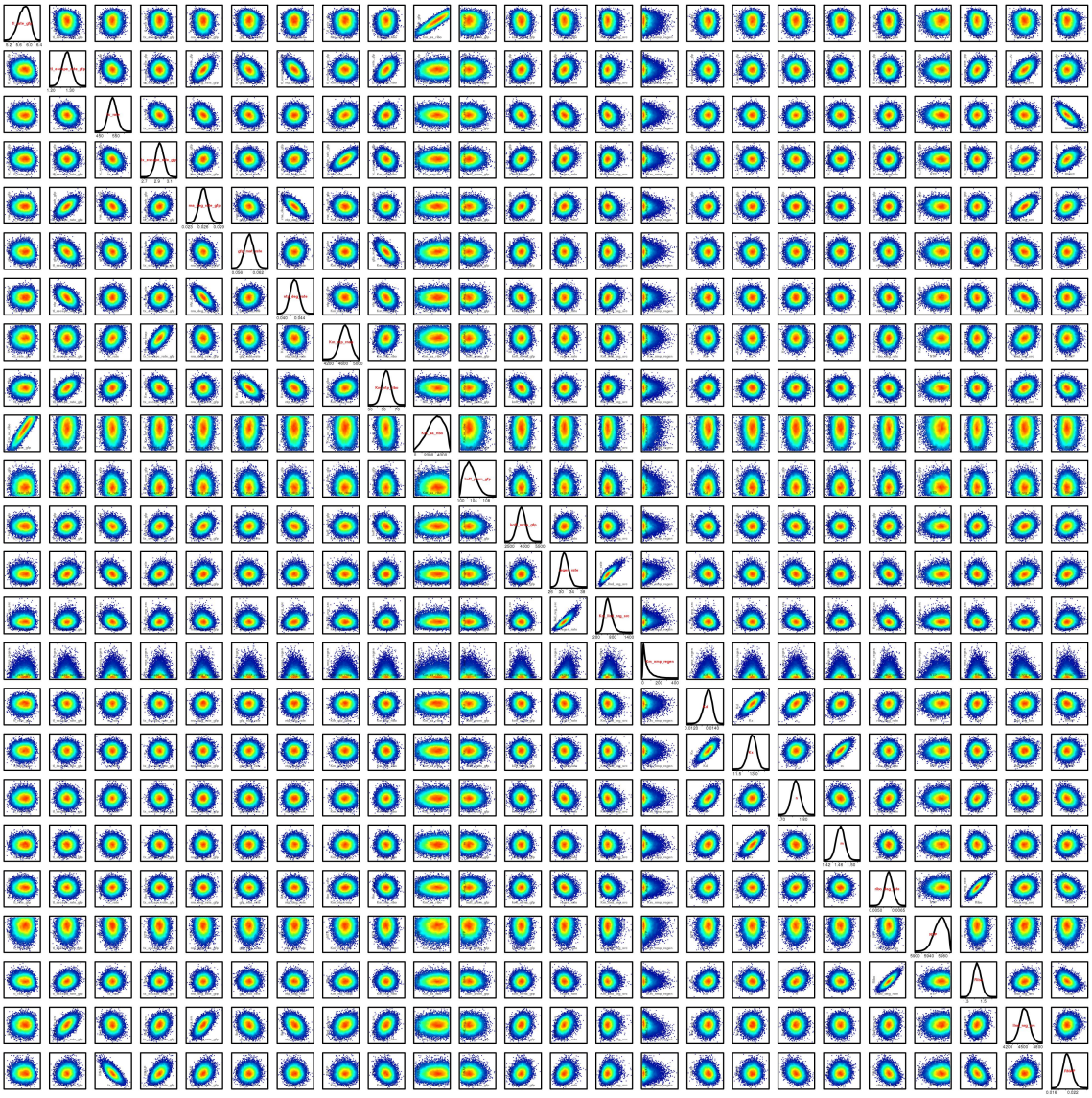


Fig. S22. Posterior distribution represented as univariate and bivariate marginal distributions after MCMC parameter inference with experimental data from the xylose-XylR experiment (Figs S19-S21 and Main text Fig. 3; see Table S9 for details on the model parameters). The posterior distribution is a complex multidimensional probability distribution over the model parameters inferred from the information contained in the experimental data (Figs S19-S21). The diagonal histograms show the univariate marginal posterior distribution for each of the model parameters. The off-diagonal plots show the bivariate marginal distributions, where each point is a sample from posterior distribution and the color indicates the density of points. These off-diagonal plots can indicate correlations between different parameters (e.g. tx_rate is negatively correlated with RNAP concentration as could be expected, see plot on bottom row, column three). The axes of the off-diagonal plots are not shown for reasons of space but have the same scales as the corresponding diagonal plots. Parameter units are the same as those in Table S9.

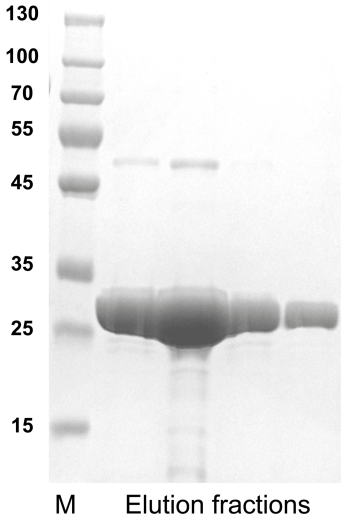


Fig. S23. Purified QconCAT_Bm_NCF^{His}. Purified fractions (in 8M urea) were analyzed with a 4-20 % gradient SDS-PAGE gel. The gel was stained with InstantBlue Protein Stain (Expedeon, USA). M - PageRuler^{TM} prestained protein ladder (ThermoFisher).

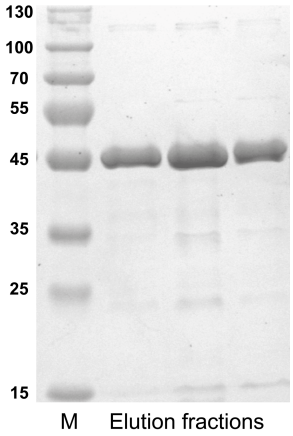


Fig. S24. Purified *B. megaterium* DSM319 RpoD^{His}. Purified fractions were analyzed with a 4-20 % gradient SDS-PAGE gel. The gel was stained with InstantBlue Protein Stain (Expedeon, USA). M - PageRuler^{TM} prestained protein ladder (ThermoFisher).

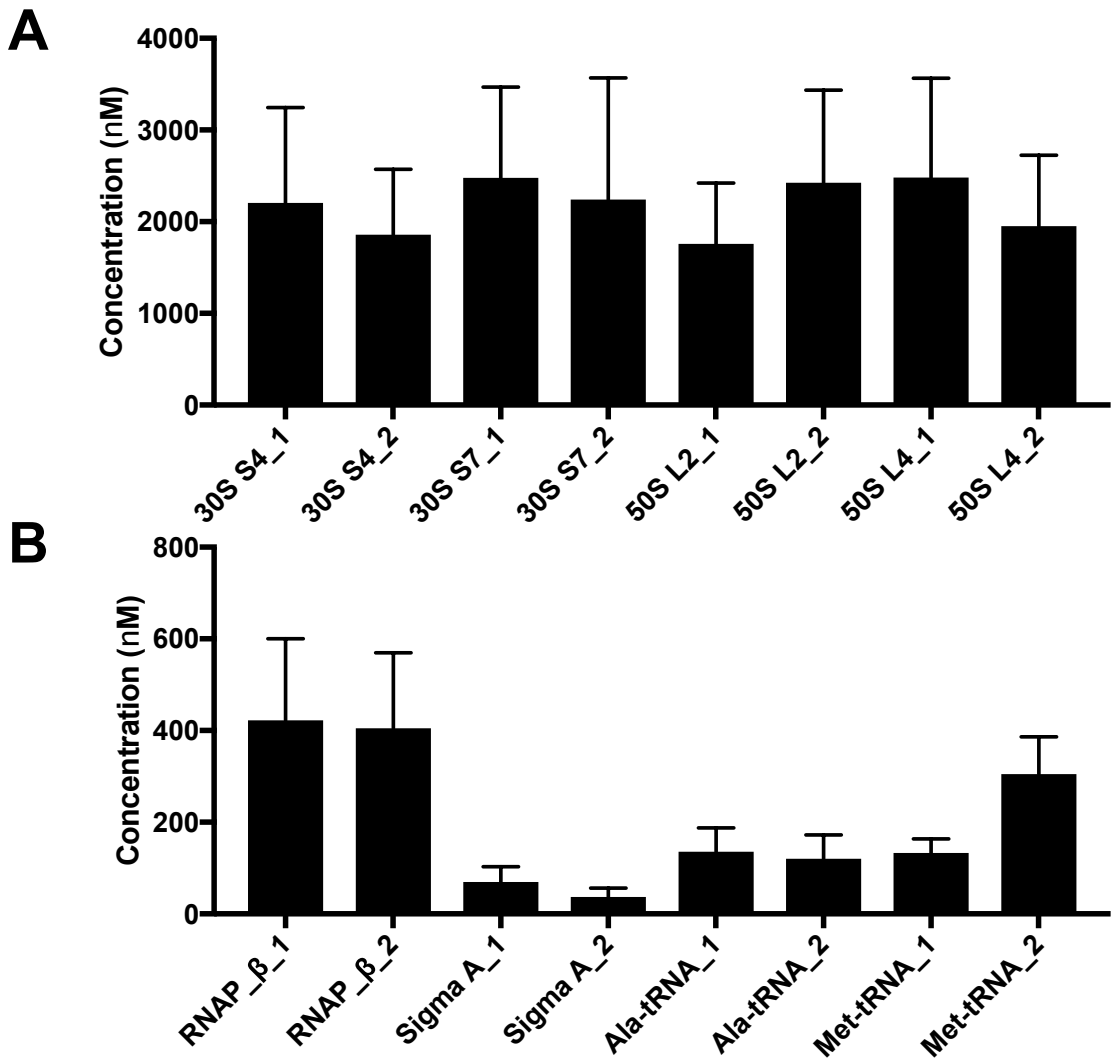


Fig. S25. Targeted LC-MS/MS values of selected transcription-translation components. (*A*) Concentration of ribosomal 30S S4, S7 and 50S L2, L4 ribosomal proteins. (*B*) Concentration of RNAP β subunit, Sigma A, alanyl-tRNA synthetase and methionyl-tRNA synthetase. Cell-extracts were digested with trypsin and analysed by LC-MS/MS as described in the supporting text. Data is representative of three biological repeats and quantified from a standard 10 mg mL^{-1} cell-extract reaction. Further detail on select peptides (two per protein, 1 and 2) is provided in *SI Appendix*, Fig. S25-S26 and Table S7-S8. Alanyl- and methionyl-tRNA synthetases are abbreviated to Ala-tRNA and Met-tRNA, respectively.

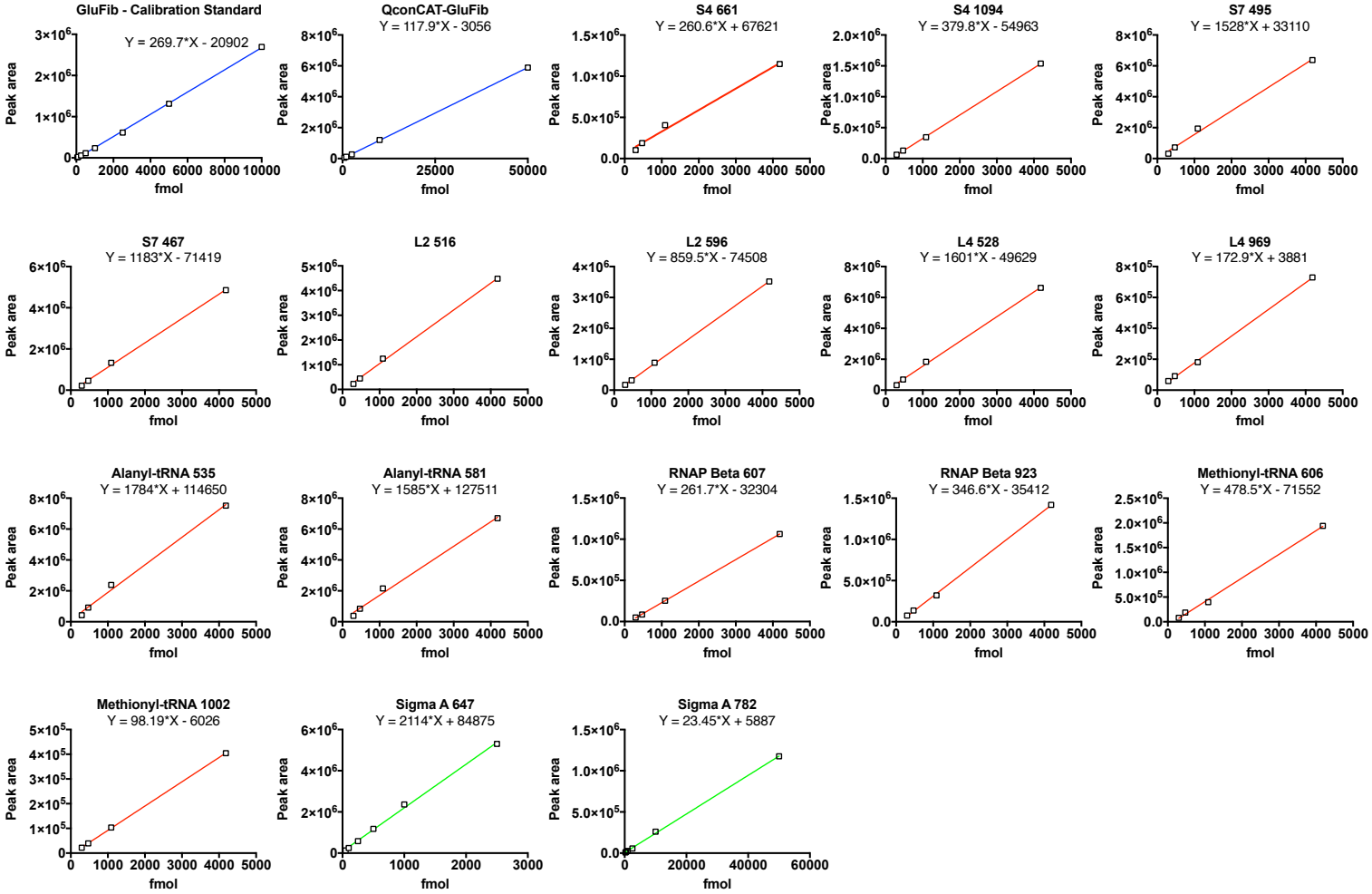


Fig. S26. LC-MS/MS ^{15}N labelled calibration standards for protein quantification

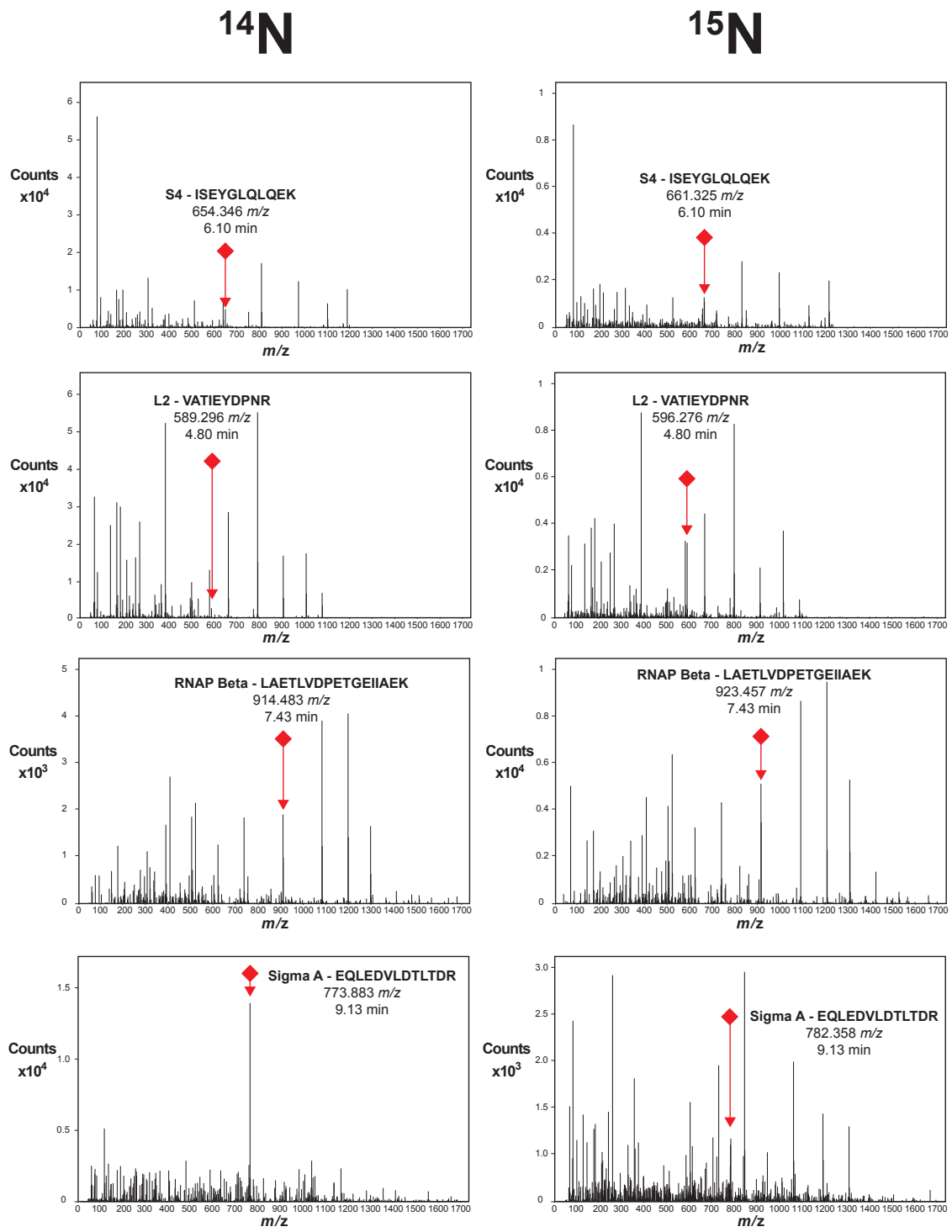


Fig. S27. ^{14}N and ^{15}N MS spectra of selected peptides for the S4, L2, RNAP β subunits and Sigma A. Please refer to Table S7-8 for further information

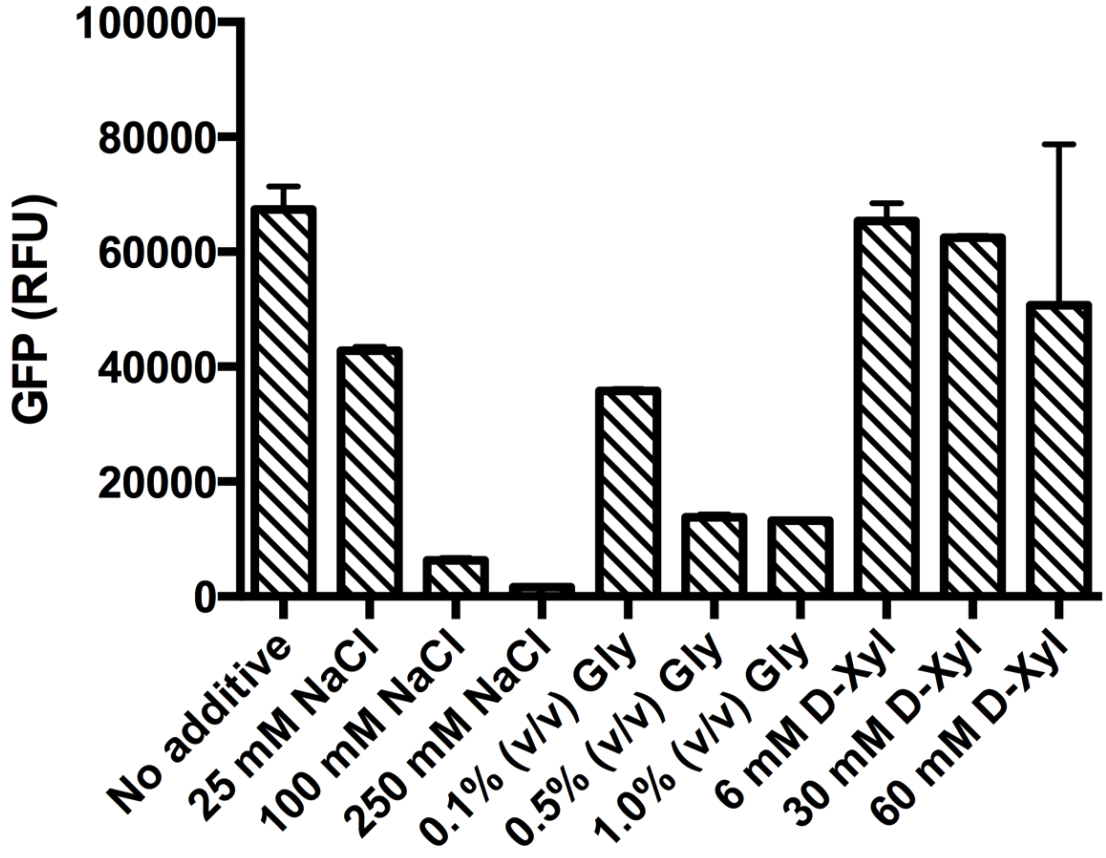


Fig. S28. Effect of additives on cell-free activity for the XylR-Xylose experiments. 10 nM of pRBBm258 was incubated in 10 mg mL^{-1} extract and 3-PGA energy buffer. Data is presented as an end-point reading of GFP fluorescence from three technical measurements.

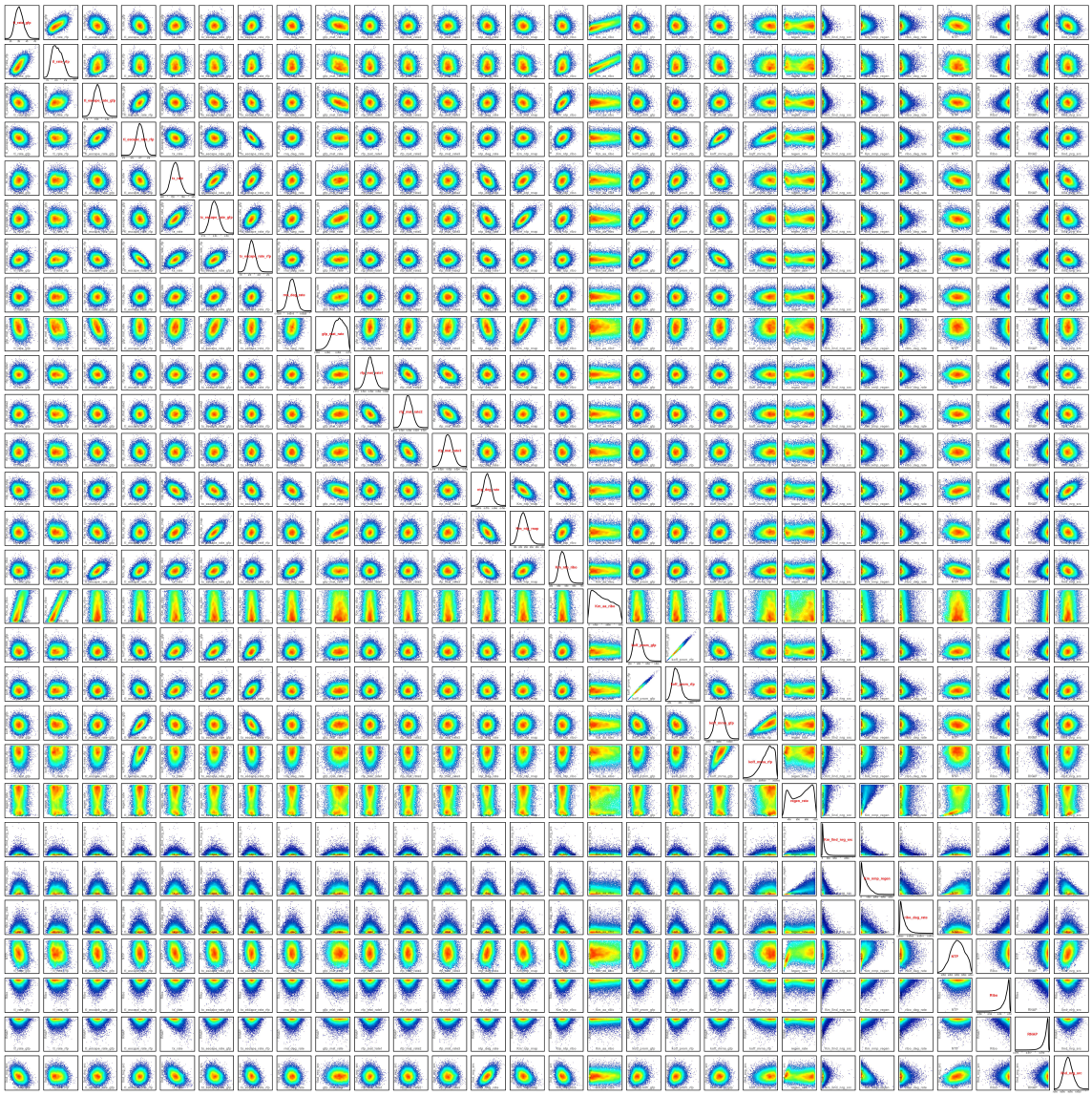


Fig. S29. Posterior distribution represented as univariate and bivariate marginal distributions after MCMC Bayesian parameter inference with data from the promoter competition experiment (Main text Fig. 4 and Fig. S29; see Table S10 for details on the model parameters). The posterior distribution is a complex multidimensional probability distribution over the model parameters inferred from the information contained in the experimental data (Main text Fig. 4 and Fig S29). The diagonal histograms show the univariate marginal posterior distribution for each of the model parameters. The off-diagonal plots show the bivariate marginal distributions, where each point is a sample from posterior distribution and the color indicates the density of points. These off-diagonal plots can indicate correlations between different parameters. The axes of the off-diagonal plots are not shown for reasons of space but have the same scales as the corresponding diagonal plots. Parameter units are the same as those in Table S10.

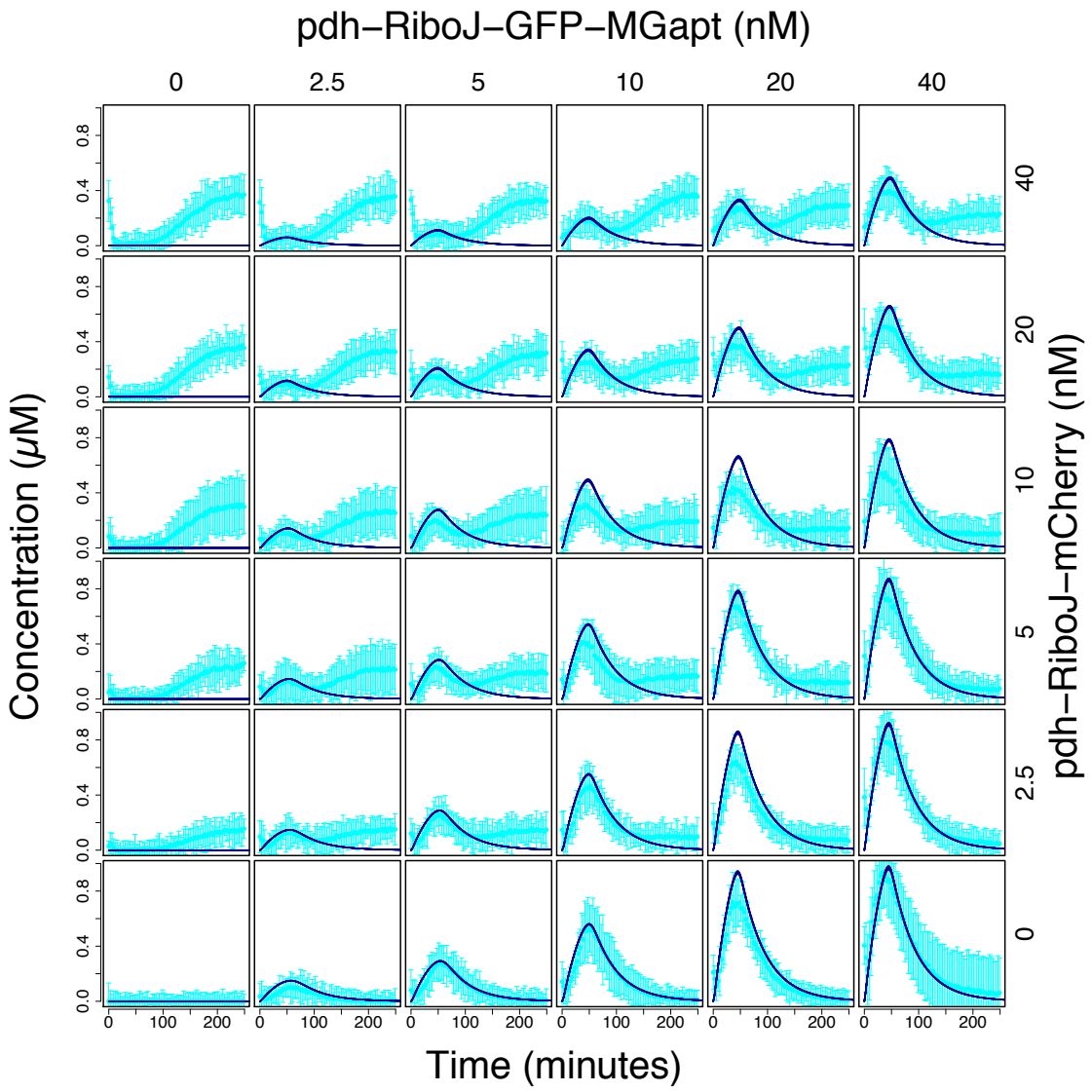


Fig. S30. Corresponding experimental and simulated GFP mRNA data corresponding from Fig. 4 in the main text. The GFP expressing plasmid (*pdh*-RiboJ-RBS-GFP-MGapt-Bba_B0015) also contained the MGapt for monitoring mRNA levels. This had crosstalk with the mCherry signal so only the initial ~90 minutes were usable. The light blue points represent the experimentally measured data (where the error bars indicate standard error), while the dark blue lines represent simulated trajectories for 500 parameter sets randomly picked from the posterior distribution (Fig. S29).

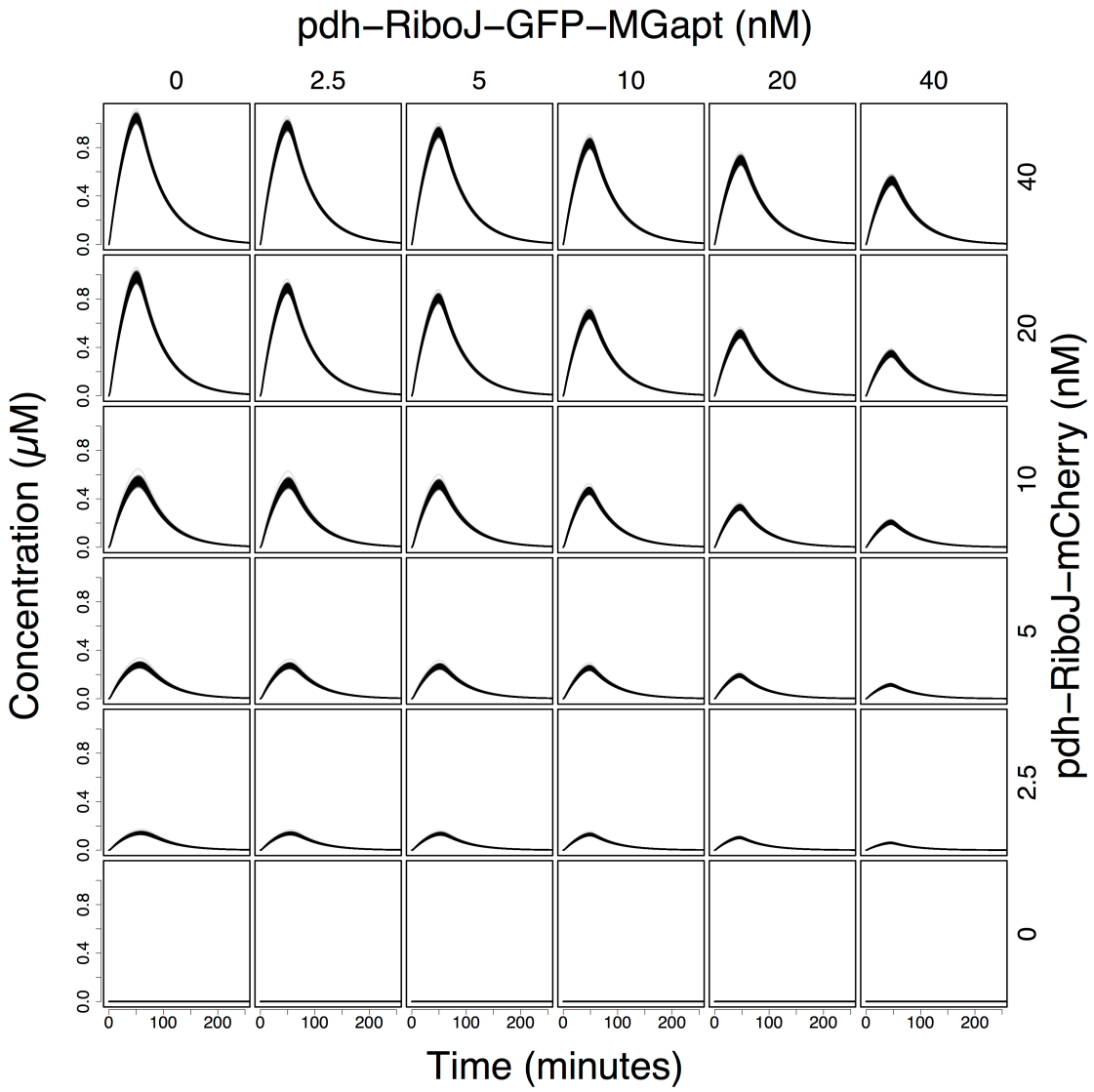


Fig. S31. Inferred mCherry mRNA data corresponding from Fig. 4 in the main text. The mCherry expressing plasmid (*pdh*-RiboJ-RBS-RFP-Bba_B0015) did not contain an aptamer for monitoring of mRNA levels. This plot shows inferred concentration levels for the mCherry mRNA for 500 parameter sets randomly picked from the posterior distribution (Fig. S28).

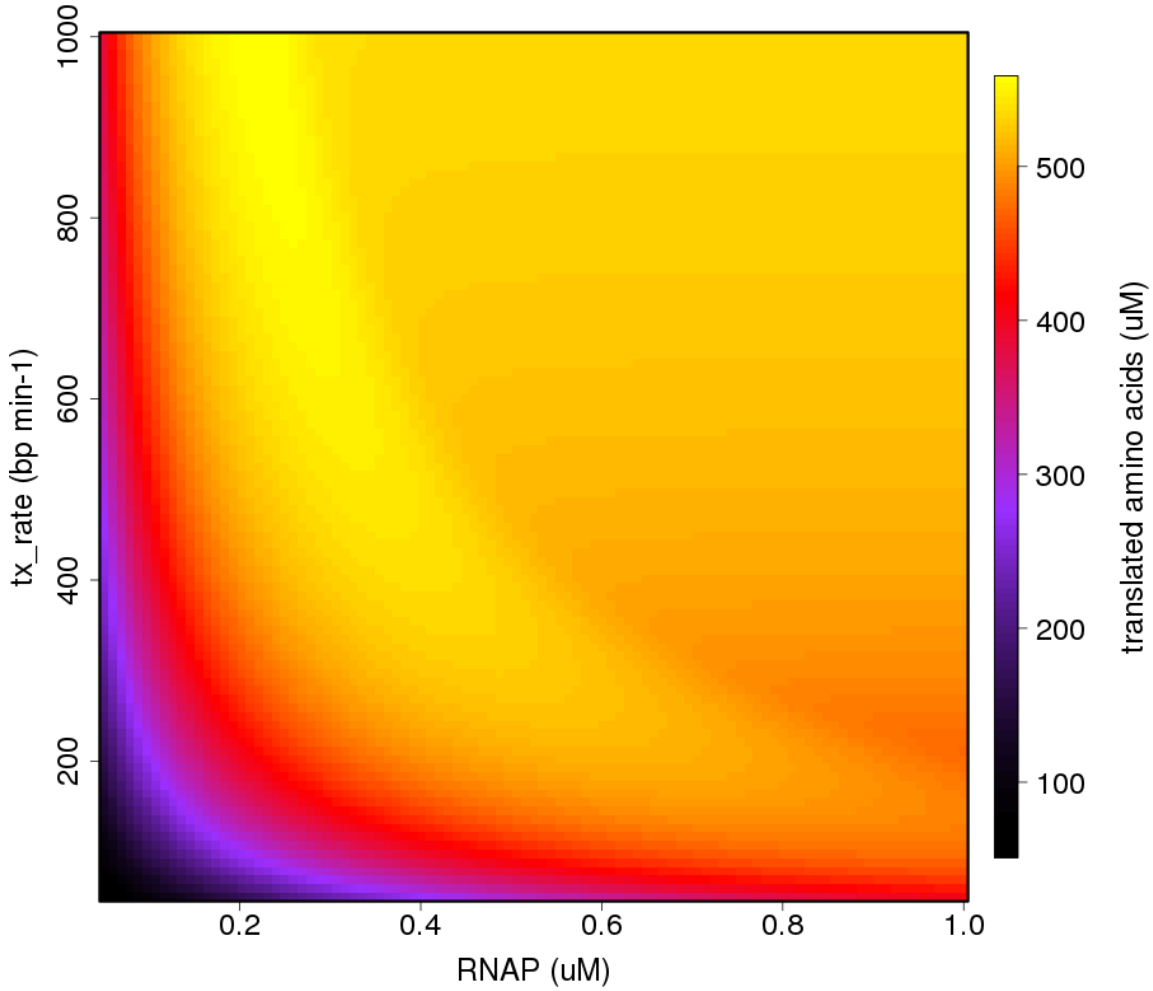


Fig. S32. Modeled total endpoint translated amino acid concentration (both mCherry and GFP) as a function of transcription elongation rate and RNAP concentration at steady state with 40 nM pdh-RiboJ-GFP-MGapt and 40 nM pdh-RiboJ-mCherry plasmid concentrations (Fig. 4 in the main text). For reference, the transcription elongation rate was inferred to be between 705-759 bp min^{-1}, while RNAP was inferred to be between 0.079 to 0.08 μM (Table S10). For simplicity, the maximum likelihood parameter set was used to simulate the data.

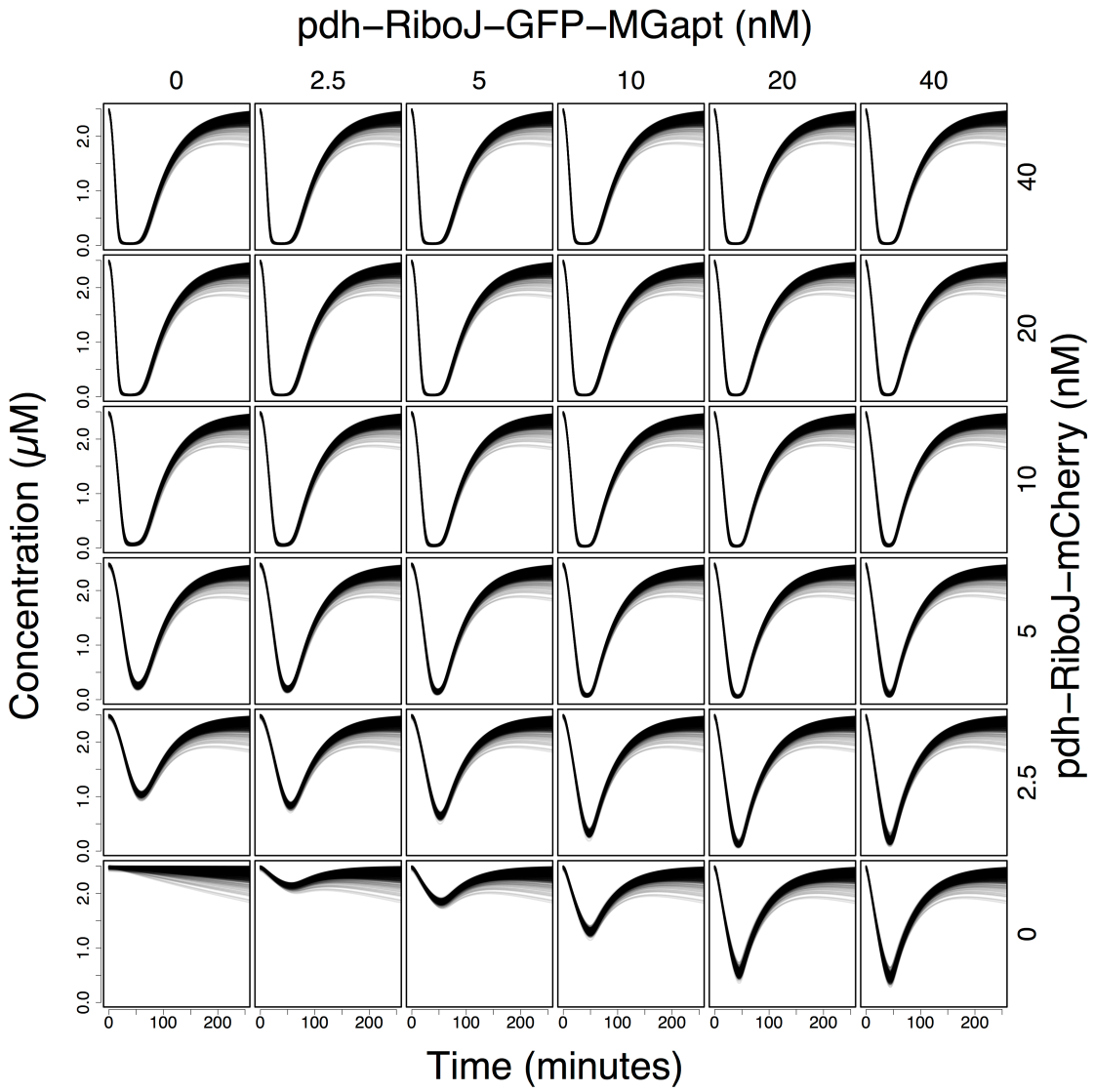


Fig. S33. Predicted free ribosome concentrations from the promoter competition experiment (Fig. 4 in the main text). 500 parameter sets were randomly picked from the posterior distribution (Fig. S28) and simulated trajectories generated for each set.

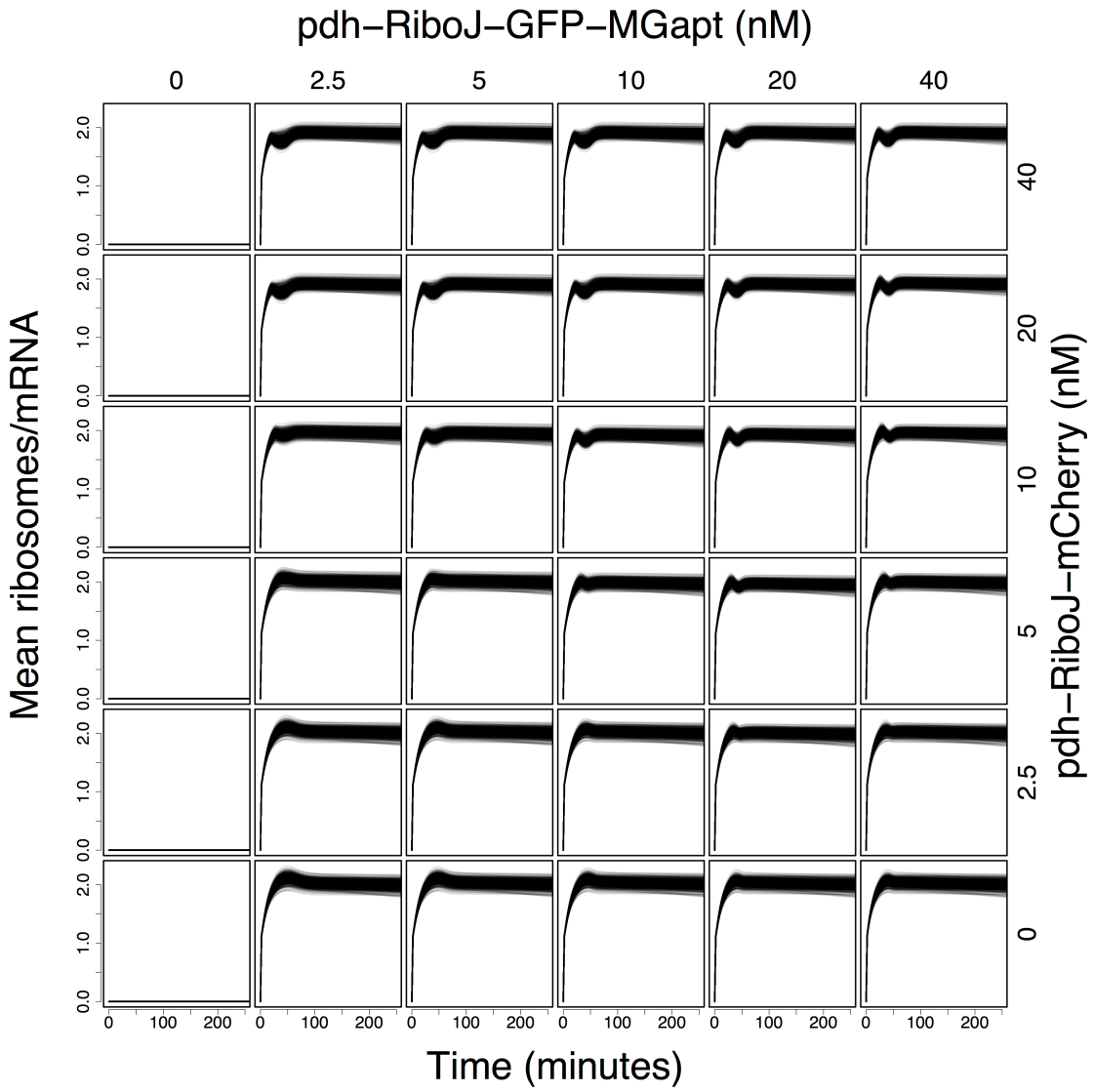


Fig. S34. Predicted mean number of ribosomes per GFP mRNA polysome from the promoter competition experiment (Fig. 4 in the main text). 500 parameter sets were randomly picked from the posterior distribution (Fig. S28) and simulated trajectories generated for each set.

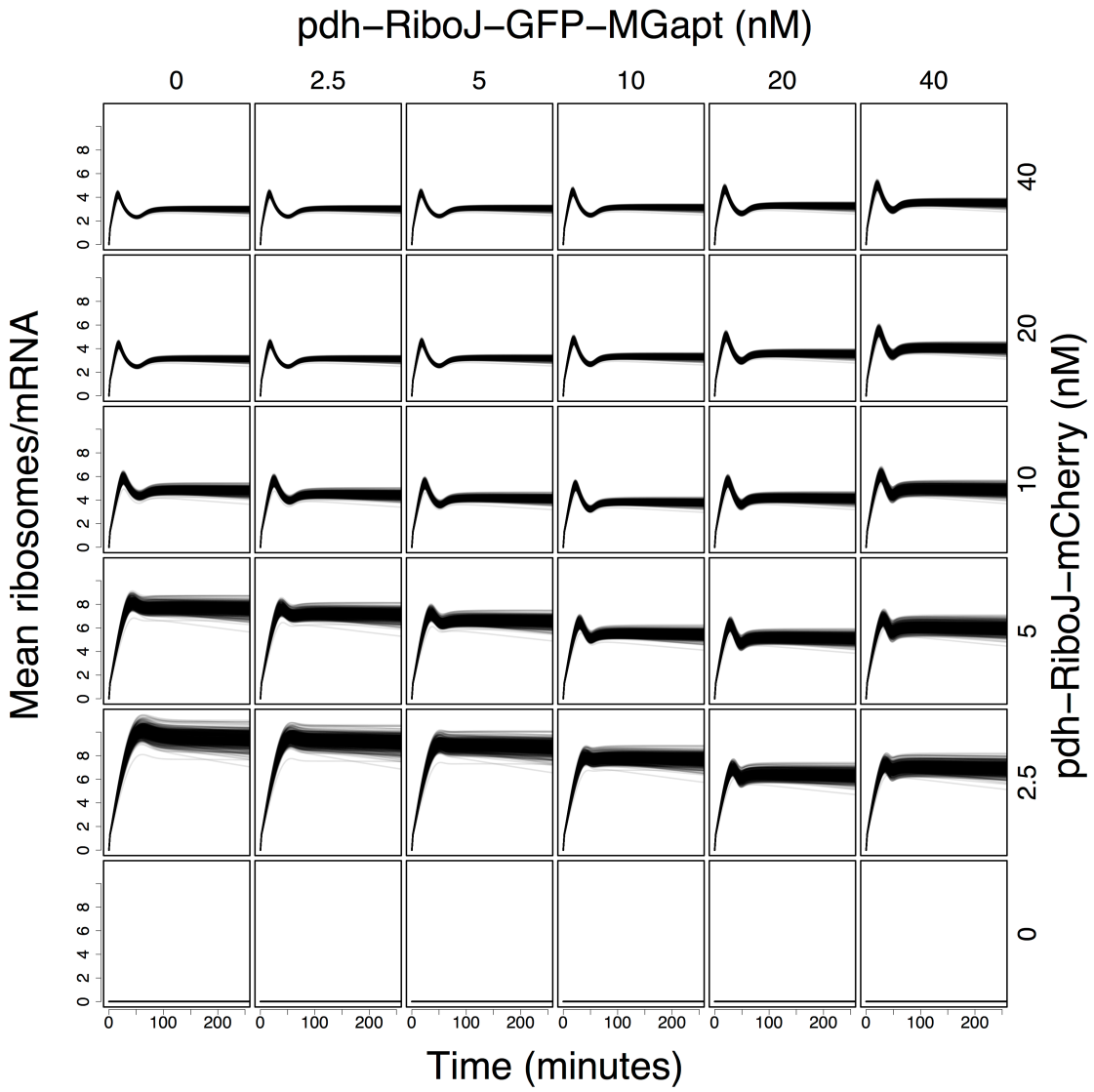


Fig. S35. Predicted mean number of ribosomes per mCherry mRNA polysome from the promoter competition experiment (Fig. 4 in the main text). 500 parameter sets were randomly picked from the posterior distribution (Fig. S28) and simulated trajectories generated for each set.

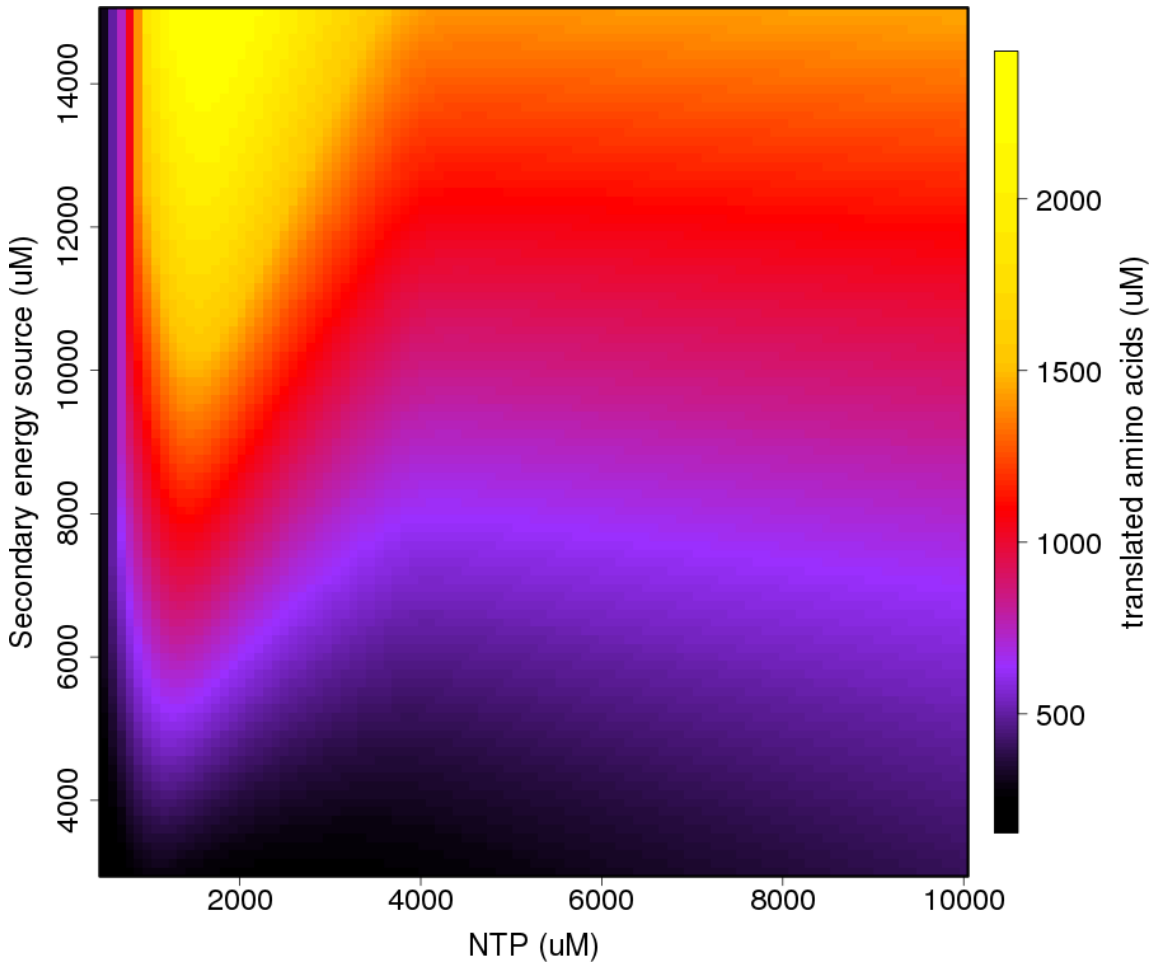


Fig. S36. Modelled total endpoint translated amino acid concentration (both mCherry and GFP) as a function of initial NTP and secondary energy source concentrations at steady state with 40 nM pdh-RiboJ-GFP-MGapt and 0 nM pdh-RiboJ-mCherry plasmid concentrations (Fig. 4 in the main text). For reference, the initial NTP concentration was inferred to be between 4330-5840 μM, while initial concentration of the secondary energy source was inferred to be between 7870-9690 μM (Table S10). For simplicity, the maximum likelihood parameter set was used to simulate the data.

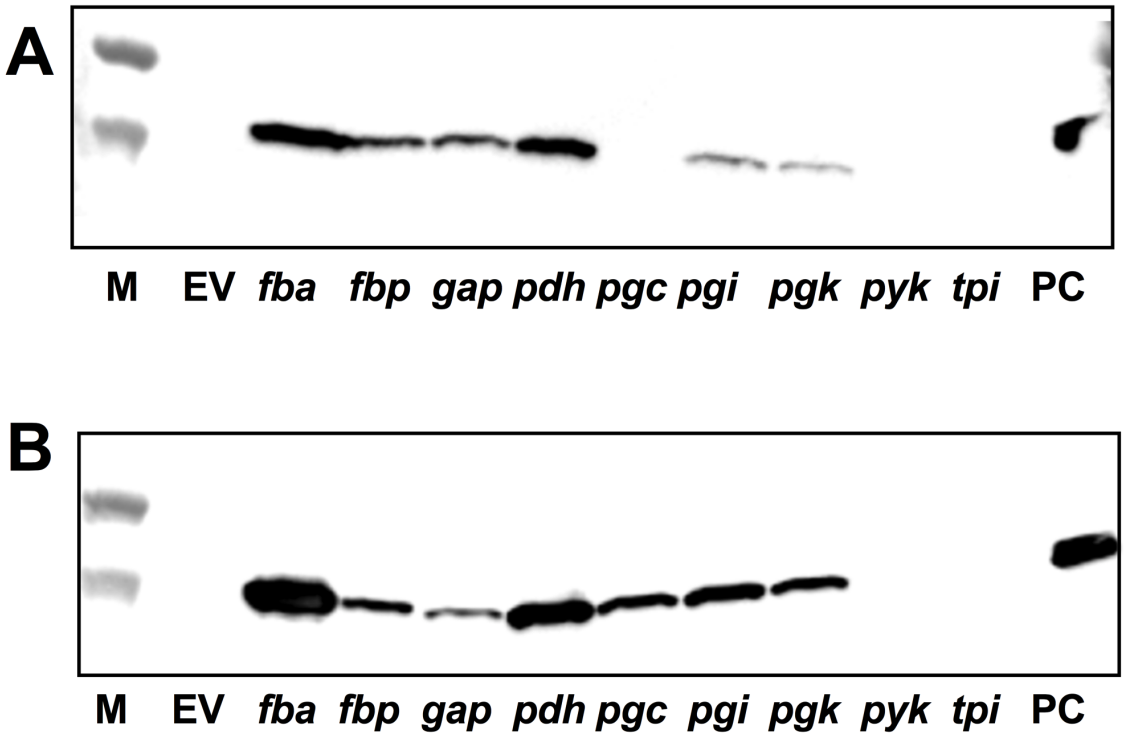


Fig. S37. Western blot of anti-GFP of cell-extracts from cell-free and *in vivo* samples. (A) Cell-free and (B) *in vivo*. For the Western blotting, equivalent samples were loaded onto on SDS-PAGE. Abbreviations: M; PageRuler Plus (ThermoFisher), cropped to 35 and 25 kDa bands, EV; empty vector and PC; positive control of purified His_6-GFP. Data is representative of a single measurement and verifies protein detection and that there is no degradation. *pyk* and *tpi* samples were not detected from either cell-free or *in vivo* samples. Weak fluorescence was detected for the *pgc* sample in cell-free, but the band was not observed under Western.

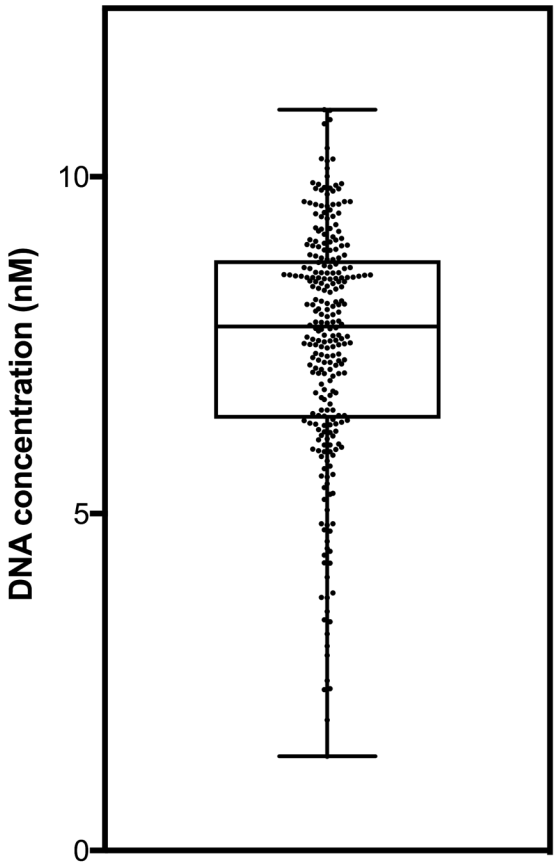


Fig. S38. Box and whisker plot of DNA concentrations from RBS variants in 2 μL screening cell-free reactions (Fig. 6B). Data statistics: *n* = 264, 1.399 nM minimum, 11.00 nM maximum, 7.78 nM median and 7.52 ± 1.8 nM mean.

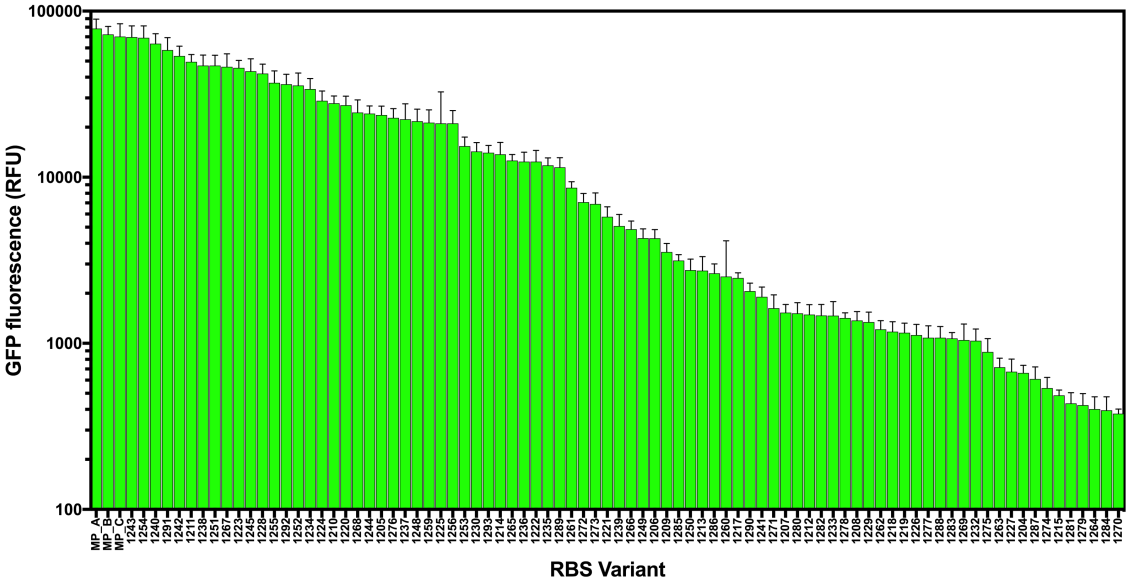


Fig. S39. Ranked RBS activities of selected RBS-4, -5, -6 and -7 variants. Please refer to Table S12-13 for RBS sequence and activities (RFU). For experimental details see *SI Appendix* text.

Table S1. Plasmid list.

Plasmid name	Description	Source
pSSBm85	Optimized xylose inducible promoter P_{xylA}^{opt}; P_{xylA}^{opt}-*gfp-stop*	Stammen *et al*., 2010
pKMMBm2	Derived from pSSBm85 carrying a *xylR-mCherry* gene encoding the fusion repressor protein XylR-mCherry; P_{xylA}^{opt}-*gfp*	Münch *et al*., 2015
pKMMBm5	Derived from pSSBm85 but with the *xylR* gene removed; P_{xylA}^{opt}-*gfp*	Münch *et al*., 2015
p3STOP1623hp	Optimized xylose inducible promoter plasmid backbone	Stammen *et al*., 2010
pMM1525	Plasmid for the secretion (signal peptide of the lipase A) of recombinant proteins in *B. megaterium* using the xylose inducible promoter; P_{xylA}-sp_{lipA}-*mcs*	Malten *et al*., 2007
pRBBm99	Derivative of p3STOP1623hp lacking all regulatory elements of the xylose inducible promoter	Münch *et al*., 2015
pET14b	Plasmid for the T7 production of recombinant proteins carrying an N-terminal His_6-tag	Novagen
pBP-BBa_B0015	iGEM Bba_B0015 double terminator in EcoFlex backbone	(4)
pRBBm257	*B. megaterium* constitutive promoter regulating the expression of operon *bmd_1326-1329*	This study
pRBBm258	*B. megaterium* constitutive promoter controlling GFP; P_{PDH}-*gfp*	This study
pRBBm266	*B. megaterium* constitutive promoter controlling mCherry; P_{PDH}-*mCherry*	This study
pRBBm299	*B. megaterium* constitutive promoter controlling codon optimised mCherry; P_{PDH}-*mCherry*	This study
pSWBm25	*B. megaterium* xylose promoter controlling mCherry; P_{PDH}-*mCherry*	This study
pRBEc2	pET14b-xylR	This study
pSJM579	pTU1-A-T7-PET_RBS-His_6-GFP-Bba_B0015	This study
pSJM673	pKMMBm5 encoding the malachite green aptamer (MGapt)	This study
pSJM687	pTU1-A-T7-PET_RBS-His_6-mCherry-Bba_B0015	This study
pSJM734	pTU1-A-T7-PET_RBS-GFP-MGapt-Bba_B0015	This study

Table S2. PCR oligonucleotides (restriction sites highlighted).

Oligo name	Sequence (5'-3')	Template	Ref
MGapt_F_SphI	CACGCATGCTCGATATAGCAATACGGATCCCGACTGGC GAGAGCCAGGTAACGAATGGATCCCCAGGCATCAAATA AAACG	pBP-BBa_B0015	(4)
MGapt_R_AgeI	GTGACCGGTAACATATAAACGCAGAAAGG		
RB128_F_BamHI	TATCAGGATCCTGTTATTATTCAAATTGCAG	pMM1525	(3)
RB129_R_BamHI	TATCAGGATCCCTAACTTATAGGGGTAACAC		
RB216_AflII_F	TATCACTTAAGGAAGGCCTGTGAAGTTTAC	*B. megaterium* DSM319 gDNA	(26)
RB217_SpeI_R	TATCAACTAGTCATCTAAGTCACCTCTTCC		
RB233_SpeI_F	ACATACTAGTAAGGGCGAGGAGGATAAC	pKMMBm2	(2)
RB234_SphI_R	ACATGCATGCTTACTTGTACAGCTCGTC		
SW55_BglI_F	ATCAAGATCTATGGTGAGCAAGGGCGAG		
SW56_EahI_R	ATCACGGCCGTTACTTGTACAGCTCG		

Table S3. EcoFlex PCR oligonucleotides (restriction sites highlighted).

Oligo name	Sequence (5'-3')	Template	Ref
Shuttle vector backbone			
Tet_RebB_F;	AGGTCTCATCGACCACCTGACGTCTAAGAAACC	pMM1525	(3)
Tet_RebB_R	AGGTCTCAAACACCTTCCTTAAGGAACGTACAGACG		
RebB_BsaI_X (BsaI mutagenesis)	CTTTTCTCTCCATGGTCTGACTTTTCCACTTTTTGTCTTG		
GFP_MGapt transcript			
GFP_MGapt_F	AGGTCTCACATATGACTAGTTCGAAGATCTGG	pSJM673	This work
GFP_MGapt_R	AGGTCTCATCGACTTGGACCGGTAACATATAAACG		
B. megaterium DSM319 RNA polymerase Sigma A subunit – *rpoD*			
RpoD_BsaI_F	AGGTCTCACATATGGCTGAAAAATCAGCTCAATC	Genomic DNA	This work
RpoD_BsaI_R	AGGTCTCATCGATTATTCTAAGAAATCTTTTAAACGTTTGCTTC		
mCherry codon optimised for *B. megaterium* DSM319 – synthesised by Life Technologies™ Geneart™ (Regensburg, Germany) and cloned into pRBBm299 separately with BglII and SphI			
mCherry_Bm_NdeI_F	CACCATATGGTATCTAAAGGTGAAGAAGATAATATGG	pRBBm299	This work
mCherry_Bm_BamHI_R	CACGGATCCTTATTTATATAATTCATCCATACCACCTGTAGAATG		
Glycolysis and gluconeogenesis promoters – template *B. megaterium* genomic DNA			
tpi_BsaI_F	AGGTCTCACTATGTTTATGTTTCAGTAAGACGGAG		
tpi_BsaI_R	AGGTCTCAGTACTTTGTTTATTGTGGGTTTTCCCTTC		
pyk_BsaI_F	AGGTCTCACTATTAAAAGGGCTGTCTCTTAGG		
pyk_Bsa_R	AGGTCTCAGTACTTTTAATGTAGGATGCCCATTTTC		
pgk_BsaI_F	AGGTCTCACTATGAATAACTTAAAATTCGCTTC		
pgk_BsaI_R	AGGTCTCAGTACTTAAGTATTTATGGTACTAAGCAAGTAGG		
pgi_BsaI_F	AGGTCTCACTATACGCTGGAATTTCCTATAC		
pgi_BsaI_R	AGGTCTCAGTACATTAAAATATCTTTATTTTTCACTTTATCG		
pgc_BsaI_F	AGGTCTCACTATAAAATGAGGCTGAGACAAAAG		
pgc_BsaI_R	AGGTCTCAGTACATTAGGCAATTAGTTTCCCTG		
pdh_BsaI_F	AGGTCTCACTATGAAGGCCTGTGAAGTTTACAG		
pdh_BsaI_R	AGGTCTCAGTACTTTCATTTAAGCCGTATAAATTCATTTTAC		
fbp_BsaI_F	AGGTCTCACTATGTTCACAAATCACTACATTCG		
fbp_BsaI_R	AGGTCTCAGTACTTTCAACGTTCACCATTTAGTATAATC		
fba_BsaI_F	AGGTCTCACTATGGTTTTGGTAGGGTTTTCATC		
fba_BsaI_R	AGGTCTCAGTACTTAAAATCAACCTTAGATTAGATGC		

Table S4. Sequencing oligonucleotides.

Sequencing		
pRBBm258_for	GTAAATGCATTGACGAACTGTTTTAG	pRBBm258-MGapt
pMM1520_for	ATGATGAGATAAAGTTAGTTTATTGG	pKMMBm5-MGapt
pMM1520_rev	GTTTGCGCATTCACAGTTCTCC	pRBBm258-MGapt / pKMMBm5-MGapt
TetA_RebB_seq1	GAATATAATGGGCTATCATTCCACC	pBP-TetA_RebB
TetA_RebB_seq2	CAATGTGGAATTGGGAACGG	pBP-TetA_RebB

Table S5. Comparison of cell-free and live cell systems using *E. coli* data collated from several studies and presented by Garamella *et al* (27).

	E. coli* – from Garamella *et al (27)		***B. megaterium* – this study**	
Parameter				
	in vivo	**Cell-free**	***in vivo***	**Cell-free (low-high range)**
Cell-size **l × w × d (μm)**	1 × 1 × 0.5	NA	4-8 (variable) × 1.5 × 1.5	NA
Protein concentration (mg/ml)	200-320	10	n.d.	10
	in vivo	**Cell-free**	***in vivo***	**Cell-free**
Transcription				
RNA polymerase apo (nM)	1900	60-75	n.d.	285-628 (2 s.f)
Sigma σ (nM)	730	< 35	N/A	N/A
Sigma A (nM)	N/A	N/A	n.d.	14-96 (2 s.f)
mRNA elongation rate (nuc/s)	39-55	> 5	n.d.	5.81-5.87
mRNA mean lifetime (min)	9.8	16-17	n.d.	15.59
Translation				
30S+50S (μM) **[active fraction]**	46.5	< 2.3	n.d.	1.43-2.48 (from model)
30S (μM) [total]	n.d.	n.d.	n.d.	1.1-3.6 (LC-MS)
50S (μM) [total]	n.d.	n.d.	n.d.	1.1-3.7 (LC-MS)
tRNA synthetase (nM) [total]	n.d.	n.d.	n.d.	86-397 (LC-MS)
Peptide chain elongation rate (AA/s)	18	> 1.5	n.d.	0.03-0.66 (from model – mCherry-GFP)

n.d. – not determined

N/A – not applicable

Table S6. Preparation of amino acids for 50 mL of stock solution.

Amino Acid	MW	mg	mM
L-Alanine	89.09	26.7	6
L-Arginine hydrochloride	210.66	63.2	6
L-Asparagine	132.12	39.6	6
L-Aspartic acid	133.10	39.9	6
L-Cysteine	121.16	36.4	6
L-Glutamic acid	147.13	44.1	6
L-Glutamine	146.14	43.8	6
Glycine	75.07	22.5	6
L-Histidine hydrochloride	191.62	57.5	6
L-Isoleucine	131.18	39.4	6
L-Leucine	131.17	32.8	5
L-Lysine hydrochloride	182.65	54.8	6
L-Methionine	149.21	44.8	6
L-Phenylalanine	165.19	49.6	6
L-Proline	115.13	34.5	6
L-Serine	105.09	31.5	6
L-Threonine	119.12	35.7	6
L-Tryptophan	204.23	61.3	6
L-Tyrosine	181.19	54.4	6
L-Valine	117.15	35.2	6

Table S7. LC-MS/MS of singly charged fragment (y) ions for the targeted QconCAT and Sigma A peptides.

Protein	No.	Peptide	y18	y17	y16	y15	y14	y13	y12	y11	y10	y9	y8	y7	y6	y5	y4	y3	y2	y1
30S ribosomal protein S4	1	ISEYGLQLQEK									1194.6	1107.568	978.5255	815.4621	758.4407	645.3566	517.298	404.214	276.1554	147.1128
30S ribosomal protein S4	2	EALEVNNFVPDYLTVDAEK	2037.018	1965.98	1852.896	1723.854	1624.785	1510.742	1396.7	1249.631	1150.563	1053.51	938.4829	775.4196	662.3355	561.2879	462.2195	347.1925	276.1554	147.1128
30S ribosomal protein S7	1	AFDLVQER												906.468	759.3995	644.3726	531.2885	432.2201	304.1615	175.119
30S ribosomal protein S7	2	WLVNYAR													735.4148	622.3307	523.2623	409.2194	246.1561	175.119
50S ribosomal protein L2	1	SAGTSAQVLGK									931.5207	860.4836	803.4621	702.4145	615.3824	544.3453	416.2867	317.2183	204.1343	147.1128
50S ribosomal protein L2	2	VATIEYDPNR										1078.516	1007.479	906.4316	793.3475	664.3049	501.2416	386.2146	289.1619	175.119
50S ribosomal protein L4	1	AQDNSIVVLEALSFDAPK		1845.959	1717.901	1602.874	1488.831	1401.799	1288.715	1189.646	1090.578	977.4938	848.4512	777.4141	664.3301	577.298	430.2296	315.2027	244.1656	147.1128
50S ribosomal protein L4	2	GGGVVFGPVPR									984.5625	927.5411	870.5196	771.4512	672.3828	525.3144	468.2929	371.2401	272.1717	175.119
alanyl-tRNA synthetase	1	VEGLLEEIR											958.5204	829.4778	772.4563	659.3723	546.2882	417.2456	288.203	175.119
alanyl-tRNA synthetase	2	IVSESGIGAGTR								1033.527	934.4588	847.4268	718.3842	631.3522	574.3307	461.2467	404.2252	333.1881	276.1666	175.119
RNAP - beta subunit	1	LAETLVDPETGEIIAEK			1714.875	1643.837	1514.795	1413.747	1300.663	1201.595	1086.568	989.515	860.4724	759.4247	702.4032	573.3606	460.2766	347.1925	276.1554	147.1128
RNAP - beta subunit	2	DIPNVGEDALR									1083.579	970.4952	873.4425	759.3995	660.3311	603.3097	474.2671	359.2401	288.203	175.119
methionyl-tRNA synthetase	1	VAQVTQVEPVK									1098.615	1027.578	899.5197	800.4512	699.4036	571.345	472.2766	343.234	246.1812	147.1128
methionyl-tRNA synthetase	2	EVPFGSDGVFTPEAFVER		1853.907	1754.838	1657.786	1510.717	1453.696	1366.664	1251.637	1194.615	1095.547	948.4785	847.4308	750.3781	621.3355	550.2984	403.23	304.1615	175.119
Sigma A	1	GVLTYEEIAER									1222.631	1123.563	1010.479	909.4312	746.3679	617.3253	488.2827	375.1987	304.1615	175.119
Sigma A	2	EQLEDVLDTLTDR							1417.717	1289.658	1176.574	1047.532	932.5047	833.4363	720.3523	605.3253	504.2776	391.1936	290.1459	175.119

Table S8. Peptide ions and retention times for calibration and quantitation in cell-extracts.

Protein	Peptide	Retention time (min)	Charge state	14N precursor *m/z*	15N precursor *m/z*
30S ribosomal protein S4	ISEYGLQLQEK	6.10	2	654.346	661.325
30S ribosomal protein S4	EALEVNNFVPDYLTVDAEK	8.98	2	1083.534	1094.501
30S ribosomal protein S7	AFDLVQER	5.95	2	489.256	495.239
30S ribosomal protein S7	WLVNYAR	6.47	2	461.251	467.233
50S ribosomal protein L2	SAGTSAQVLGK	3.72	2	509.780	516.261
50S ribosomal protein L2	VATIEYDPNR	4.80	2	589.296	596.276
50S ribosomal protein L4	AQDNSIVVLEALSFDAPK	9.62	2	959.002	969.471
50S ribosomal protein L4	GGGVVFGPVPR	6.22	2	521.296	528.275
alanyl-tRNA synthetase	VEGLLEEIR	7.32	2	529.298	535.281
alanyl-tRNA synthetase	IVSESGIGAGTR	4.18	2	573.809	581.287
RNAP - beta subunit	LAETLVDPETGEIIAEK	7.43	2	914.483	923.457
RNAP - beta subunit	DIPNVGEDALR	6.34	2	599.807	607.285
methionyl-tRNA synthetase	VAQVTQVEPVK	4.62	2	599.346	606.325
methionyl-tRNA synthetase	EVPFGSDGVFTPEAFVER	9.29	2	991.978	1002.447
Sigma A	GVLTYEEIAER	6.55	2	640.330	647.310
Sigma A	EQLEDVLDTLTDR	9.13	2	773.883	782.358

Table S9 – Model parameters and inferred values of the xylose-inducible promoter system in cell-free (see Main text Figs 2 & 3, and S18-22).

Parameter name	Units	Description	Notes	Inferred values (DNA titration only)			Inferred values (varying DNA/XylR/xylose)		
				5% quantile	Median	95% quantile	5% quantile	Median	95% quantile
tl_rate_gfp	aa min^{-1}	translation elongation rate	Previous estimates for translation elongation rate are about 18 aa/s (18×60 = 1080 aa/min) *in vivo* and 1.5 aa/s (90 aa/min) *in vitro* (28). It is highly likely that this will vary depending on codon usage in the coding sequence.	5.43	5.87	6.31	5.44	5.81	6.12
tl_escape_rate_gfp	min^{-1}	translation initiation rate	This is the overall translation initiation rate defined to mean the binding of the 50s subunit to the already formed 30s:mRNA complex followed by GTP hydrolysis and Pi release. Previously, 50S binding was found to be fast, GTP hydrolysis rate about 30 s^{-1}, however, Pi release is about 1.5 s^{-1} (90 min^{-1}) and dominates the overall rate (29).	1.16	1.21	1.26	1.25	1.28	1.32
tx_rate	bp min^{-1}	transcription elongation rate	Previously, mRNA elongation rate in cell-free has been estimated at about 3 bp/sec (180 bp/min) (30), but is estimated to be much faster in cells at up to 90 bp/sec (5400 bp/min) (31).	554	616	688	488	529	572
tx_escape_rate_gfp	min^{-1}	promoter escape rate	This is the composite rate constant of the rate limiting step in transcription initiation. Going by the terminology in Hsu LM, 2002 paper (32), this is likely dominated by either either be k_2 or k_E. k_E half life values can range from 5 s (8.316 min^{-1}) to 10 min (0.0693 min^{-1}).	2.67	2.81	2.94	2.84	2.94	3.05
rna_deg_rate_gfp	min^{-1}	mRNA degradation rate	mRNA half-life has been previously found to be around 5 min (0.14 min^{-1}) for *E. coli*: 21 min (0.032 min^{-1}) for *S. cerevisiae*: 600 min (0.0012 min^{-1}) for human HepG2/Bud8 cells, so rate constant is 0.693 / 5 = 0.1386 min^{-1}(32, 33).	0.0302	0.0321	0.0342	0.0248	0.0261	0.0273
gfp_mat_rate	min^{-1}	GFP+ maturation rate	Wildtype GFP is estimated to have a maturation rate of 0.0186 min^{-1}, while the fastest folding GFP, GFPmut3 is estimated to have a rate of 0.132 min^{-1}. GFP+, used in this study, is synonymous to GFPuv3 which has been estimated to have a maturation rate of 0.066 min^{-1} (34).	0.0576	0.0603	0.0632	0.0573	0.0594	0.0615
ntp_deg_rate	min^{-1}	NTP degradation rate	NTP degradation to NMP by non-productive NTPase side-reactions	0.0437	0.0459	0.048	0.0413	0.0428	0.0444
Km_ntp_rnap	μM	Michaelis constant for NTP binding RNAP in trancription reactions		4550	4880	4990	4430	4660	4870
Km_ntp_ribo	μM	Michaelis constant for NTP binding the ribosome in translation reactions		40.8	53.4	67.1	43.8	53.7	63.8
Km_aa_ribo	μM	Michaelis constant for amino acids binding the ribosome in translation reactions		248	2500	4750	965	3100	4720
koff_prom_gfp	min^{-1}	koff for the RNAP:promoter complex	Previous work suggests that KM_dna is likely fall between 0.001-0.2 μM (35). If initial binding of RNAP to the promoter is assumed to be in rapid equilibrium so that kon can be set to an arbitrarily high value (1000000 uM min^{-1}), koff would lie in the range 1000 min^{-1} to 200000 min^{-1}.	100	101	106	100	103	106
koff_mrna_gfp	min^{-1}	koff for the ribosome:mRNA complex	Previous work suggests that KM_mRNA is likely to range from 0.003 to 0.3 uM. If initial binding of the ribosome to the RBS is assumed to be in rapid equilibrium such that kon can be set to an arbitarily high value (1000000 uM min^{-1}), koff would lie in the range 3000 to 300000 min^{-1}.	1250	2000	2810	3000	3640	4300
regen_rate	μM min^{-1}	Vmax for the NMP + IInd_nrg_src ->NTP reaction		33.9	39.2	42.9	29	31.2	33.8
Km_IInd_nrg_src	μM	Michaelis constant for IInd_nrg_src in NTP regeneration		3320	4520	4960	444	630	883

		reaction							
Km_nmp_regen	μM	Michaelis constant for NMP in NTP regeneration reaction		2.82	44	268	1.86	31.1	162
Kd	μM		Kd of binding of XylR repressor to the promoter	n/a	n/a	n/a	0.0129	0.0136	0.0142
Kx	μM		Kd of binding of xylose to the XylR repressor protein	n/a	n/a	n/a	12.2	12.9	13.5
n	n/a		Hill coefficient of binding of XylR to the promoter	n/a	n/a	n/a	1.74	1.77	1.8
m	n/a		Hill coefficient of binding of xylose to XylR	n/a	n/a	n/a	1.44	1.47	1.49
ribo_deg_rate	min^{-1}	Inactivation rate of ribosomes (8)		0.0136	0.0137	0.0139	0.00544	0.00587	0.0063
$[NTP]_0$	μM	Initial NTP concentrration	Experimentally this was set to 5000 μM	4550	4880	4990	5940	5970	6000
$[Ribo]_0$	μM	Initial ribosome concentration	Upper bound set by mass spec experiments in this paper	2.41	2.48	2.5	1.36	1.43	1.5
$[IInd\_nrg\_src]_0$	μM	Initial secondary energy source concentration (in μM NTP equivalents)	3PGA was used as the secondary energy source in these experiments	5270	5570	5830	4380	4530	4680
$[RNAP]_0$	μM	Initial RNAP concentration	Upper bound set by mass spec experiments in this paper	0.0322	0.036	0.0404	0.0187	0.0205	0.0227

Table S10 – Model parameters and inferred values of the cell-free promoter competition experiment. (see Main text, Fig. 4, and S29-35)

Parameter name	Units	Description	Notes	5% quantile	Median	95% quantile
tl_rate_gfp	aa min^{-1}	translation elongation rate	Previous estimates for translation elongation rate are about 18 aa/s (18×60 = 1080 aa/min) *in vivo* and 1.5 aa/s (90 aa/min) *in vitro* (28). It is highly likely that this will vary depending on codon usage in the coding sequence.	31.6	35.3	39.6
tl_rate_rfp	aa min^{-1}	translation elongation rate	as above	1.87	2.02	2.2
tl_escape_rate_gfp	min^{-1}	translation initiation rate	This is the overall translation initiation rate defined to mean the binding of the 50s subunit to the already formed 30s:mRNA complex followed by GTP hydrolysis and P_i release. Previously, 50S binding was found to be fast, GTP hydrolysis rate about 30 s^{-1}, however, P_i release is about 1.5 s^{-1} (90 min^{-1}) and dominates the overall rate (29).	0.19	0.202	0.216
tl_escape_rate_rfp	min^{-1}	translation initiation rate	as above	0.625	0.698	0.771
tx_rate	bp min^{-1}	transcription elongation rate	Previously, mRNA elongation rate in cell-free has been estimated at about 3 bp/sec (180 NTPs/min) (30), but is estimated to be much faster in cells at up to 90 bp/sec (5400 bp/min)(31)	705	730	759
tx_escape_rate_gfp	min^{-1}	promoter escape rate	This is the composite rate constant of the rate limiting step in transcription initiation. Going by the terminology in Hsu LM, 2002 paper (32), this is likely dominated by either be k_2 or k_E. k_E half-life values can range from 5 s (8.316 min^{-1}) to 10 min (0.0693 min^{-1}).	2.14	2.2	2.26
tx_escape_rate_rfp	min^{-1}	promoter escape rate	as above	1.93	2.04	2.17
rna_deg_rate	min^{-1}	mRNA degradation rate	Here we assume that the degradation rate is the same for all mRNA species. mRNA half-life has been previously found to be around 5 min (0.14 min^{-1}) for *E. coli*: 21 min (0.032 min-1) for *S. cerevisiae*: 600 min (0.0012 min^{-1}) for human HepG2/Bud8 cells so rate constant is 0.693 / 5 = 0.1386 min^{-1} (32, 33).	0.0211	0.0216	0.0221
gfp_mat_rate	min^{-1}	GFP+ maturation rate	Wildtype GFP is estimated to have a maturation rate of 0.0186 min^{-1}, while the fastest folding GFP, GFPmut3 is estimated to have a rate of 0.132 min^{-1}. GFP+, used in this study, is synonymous to GFPuv3 which has been estimated to have a maturation rate of 0.066 min^{-1} (34).	0.0605	0.0654	0.0693
rfp_mat_rate 1	min^{-1}	mCherry maturation rate 1	dsRed, from which mCherry is derived, has been proposed to have a three step maturation (36) with roughly three equal rates (0.00456 min^{-1}), although other models have been proposed (37).	0.0201	0.0218	0.0236
rfp_mat_rate 2	min^{-1}	mCherry maturation rate 2	as above	0.0201	0.0218	0.0236

rfp_mat_rate 3	min^{-1}	mCherry maturation rate 3	as above	0.0201	0.0218	0.0236
ntp_deg_rate	min^{-1}	NTP degradation rate	NTP degradation to NMP by non-productive NTPase side-reactions	0.0717	0.0755	0.0794
Km_ntp_rnap	μM	Michaelis constant for NTP binding RNAP in trancription reactions		168	225	295
Km_ntp_ribo	μM	Michaelis constant for NTP binding the ribosome in translation reactions		382	451	536
Km_aa_ribo	μM	Michaelis constant for amino acids binding the ribosome in translation reactions		177	2070	4650
koff_prom_gfp	min^{-1}	koff for the RNAP:promoter complex	Previous work suggests that a KM_dna is likely fall between 0.001-0.2 μM . If initial binding of RNAP to the promoter is assumed to be in rapid equilibrium so that kon can be set to an arbitrarily high value (1000000 uM min^{-1}), koff would lie in the range 1000 min^{-1} to 200000 min^{-1} (35).	306	557	876
koff_prom_rfp	min^{-1}	koff for the RNAP:promoter complex	as above	253	456	708
koff_mrna_gfp	min^{-1}	koff for the ribosome:mRNA complex	Previous work suggests that KM_mRNA is likely to range from 0.003 to 0.3 μM. If initial binding of the ribosome to the RBS is assumed to be in rapid equilibrium such that kon can be set to an arbitrarily high value (1000000 uM min^{-1}), koff would lie in the range 3000 to 300000 min^{-1}.	3660	6190	8900
koff_mrna_rfp	min^{-1}	koff for the ribosome:mRNA complex	as above	161000	243000	294000
regen_rate	μM min^{-1}	Vmax for the NMP + IInd_nrg_src ->NTP reaction		350	730	976
Km_IInd_nrg_src	μM	Michaelis constant for IInd_nrg_src in NTP regeneration reaction		6.51	91.3	462
Km_nmp_regen	μM	Michaelis constant for NMP in NTP regeneration reaction		16.7	476	1600
ribo_deg_rate	min^{-1}	Inactivation rate of ribosomes (8)		9.52E-06	0.000127	0.000531

[NTP]0	μM	Initial NTP concentrration	Experimentally this was set to 5000 μM	4330	5080	5840
[Ribo]0	μM	Initial ribosome concentration	Upper bound set by mass spec experiments in this paper	2.46	2.49	2.5
[RNAP]0	μM	Initial RNAP concentration	Upper bound set by mass spec experiments in this paper	0.0786	0.0797	0.08
[IInd_nrg_src]0	μM	Initial secondary energy source concentration (in μM NTP equivalents)	3-PGA was used as the secondary energy source in these experiments	7870	8690	9690

Table S11 – Separation method for SRM LC-MS/MS analysis.

Time (minutes)	% Solvent A	% Solvent B	Flow rate (mL/min)
0	97	3	0.2
12.5	60	40	0.2
14.5	10	90	0.2
16.5	0	100	0.2
17.50	97	3	0.2

Table S12 – RBS-6 library parts characterised in cell-free, *in vivo* and prediction software.

		Cell-Free		Cells		Prediction software	
		RFU		RFU/OD_{600}			
Plasmid ID	**RBS Sequence**	Mean	SD	Mean	SD	**RBS calculator**	**UTR designer**
pRBBm258	AGGAGGTGAATGTCAT	78538.0	10975.7	302.6	192.9	1008.4	458652.0
pSM1211	GAGGAGGTATTAT	49532.3	5381.9	617.4	23.2	7517.8	2712828.8
pSM1223	AAGGGGGTATTAT	45537.0	5050.5	691.9	143.1	18244.4	2789763.8
pSM1228	GGAGGAATTTATA	42143.1	5904.8	150.1	27.9	3384.2	952164.3
pSM1224	GGAGGAAAATTAA	28905.9	4216.2	78.8	15.2	2089.5	1369611.8
pSM1210	GAGGAGGTATTAT	27935.9	2963.6	140.3	49.3	1349.4	447502.7
pSM1220	AGGGGAGATTAAT	27130.4	3670.2	96.1	15.7	2085.9	427125.7
pSM1205	GGGGAGGTAATAT	23737.3	3091.8	138.0	46.9	9781.1	2293781.9
pSM1225	AAGGGGGTTATAT	21144.8	11562.7	221.5	54.3	20489.7	1594654.3
pSM1230	GGAGAGATATATT	14291.4	1904.5	18.5	4.4	1705.9	188061.2
pSM1214	GAGGGGGATAATA	13755.6	2482.9	146.8	32.4	2875.8	2358832.9
pSM1222	AGGGAGAAATAAAT	12429.0	2097.7	25.2	4.6	2156.1	524541.1
pSM1221	AGGAGAGTATTAA	5779.4	870.1	7.9	2.0	1519.3	226687.8
pSM1206	AGGAGAATATTAA	4290.8	557.5	6.2	1.9	2595.7	354608.3
pSM1209	GGGGGAGTAATTT	3555.1	440.4	9.2	2.5	2569.3	162064.4
pSM1213	AAAGAGGTATTAT	2738.3	589.6	3.5	0.4	1813.3	244877.3
pSM1217	GAGAGGGATATAT	2475.1	188.6	6.0	0.9	854.4	134449.4
pSM1207	AGAGGGGTAATAA	1531.3	184.9	1.9	0.4	2324.3	237415.9
pSM1212	AGAGAAAAAATAT	1490.5	222.4	1.4	0.2	330.1	17457.0
pSM1233	AAAAAAATTTTTT	1466.9	317.0	1.4	0.2	136.7	862.3
pSM1208	GGGGAAATAAATA	1375.5	180.0	1.9	0.3	556.7	116905.1
pSM1229	GAAGGAATAATAT	1342.8	200.4	1.5	0.1	452.2	82034.8
pSM1219	AGAGGAATATTAT	1158.6	165.7	1.7	0.2	999.5	116905.1
pSM1232	AAGAGGAATAAAT	1038.5	182.7	1.8	0.3	633.1	126003.4
pSM1227	AGAAGGATAATAA	674.6	129.5	5.8	2.2	903.4	62626.2
pSM1204	GAAAGGATAATAT	664.0	73.4	5.8	1.4	238.0	51867.9
pSM1215	GGGGGAAAAATTA	486.8	36.7	34.8	8.4	2402.4	761292.5

A selection of hits were chosen from the initial RBS library screens. DNA was repeat purified from a single colony, sequenced and then tested again in cell-free activity assays at the 2 μL scale under a normalized DNA concentration of 10 nM. Four measurements for each DNA sample were tested in each assay. Mean and standard deviation of GFP fluorescence at the steady-state is shown as previously described (38). Data is ranked based on cell-free activity. To determine limit of detection (LOD – 461.63), this is defined as three standard deviations from the mean negative control background fluorescence (Mean – 347.3, s.d. – 34.43).

Table S13 - RBS-4, -5, -7 and -8 library parts. Experimental details described with accompanying Table S12 text. *Cells shaded in grey fall below the LOD.

			Cell-Free		Prediction Software	
Plasmid ID	**RBS Sequence**	**Spacer**	Mean	SD	RBS Calculator	UTR Designer
pSM1243	GGGGGGATTAAT	5	69838.9	11710.3	3385.1	861937.0
pSM1254	AAGGGGGAATATATA	8	69153.1	12431.4	6145.9	1448396.9
pSM1240	GGGGGGAATAAATT	7	63683.8	9509.9	1648.4	1064858.9
pSM1291	GGAGGGGTTTT	4	58368.6	10981.0	4272.1	535808.7
pSM1242	GGAGGGGAATTT	5	53757.1	7863.8	3786.5	1116081.1
pSM1238	GGAGGGTTTTT	4	47093.8	7402.4	3620.2	670147.1
pSM1251	GGGGAGGTAATTAAT	8	47056.9	7256.1	8293.5	1665761.4
pSM1267	GGGGGGATTTAAAT	7	46297.8	9045.3	2756.1	389108.3
pSM1245	GGGAGGGTTTTT	5	43475.6	8274.6	2583.1	815052.3
pSM1255	AGGGGGATATATTTT	7	37087.5	6675.7	5577.8	565996.9
pSM1292	GGAGGGATAAA	4	36294.9	5444.7	4172.3	2230524.9
pSM1252	AAAGGAGATATTTAT	7	35722.1	6764.7	5770.6	620368.6
pSM1234	GGGGGGATATTT	5	33990.6	5337.4	2959.5	1639878.2
pSM1268	AGGAGGGAATATTT	7	24542.5	4730.9	4630.9	2329986.2
pSM1244	GGAGAGGATTTT	5	24187.9	2692.6	1283.0	125160.5
pSM1276	AAGGAGGATATAAT	7	22790.1	3099.8	10946.9	3714776.6
pSM1237	AAGGAGGTAATT	5	22354.9	5337.7	19813.9	3198887.6
pSM1248	AGGAGGAAAATA	5	21737.8	3994.4	6632.8	3208430.7
pSM1259	GGGGGAGTTAAATAA	8	21350.3	4104.8	1018.7	133998.9
pSM1256	GGAGGAGTTTAAAAT	8	21100.1	4100.8	1027.4	215559.5
pSM1253	GAAGGAGTATATTTT	8	15396.1	2065.7	1386.9	126003.4
pSM1293	GGGGGAGATTA	4	14060.8	1483.5	3477.4	1019402.8
pSM1265	AGGGGGGAATTTAA	7	12628.3	1074.2	3560.3	1283574.2
pSM1236	AAGGAGAATAAAT	6	12433.4	1725.9	3392.1	633694.7
pSM1235	AAGGAGGTTAAA	5	11802.8	1280.6	17904.2	2630169.0
pSM1289	GGGGGGAATTT	4	11477.8	1637.4	2829.8	988341.6
pSM1261	GGAGGAGTAAATTAA	8	8651.8	753.2	500.7	421743.3
pSM1272	GAGGAGATATAATT	7	7091.3	911.1	1407.5	554714.6
pSM1273	GAGGGGGTTTTAAT	7	6906.8	1136.4	4339.9	733702.6
pSM1239	GGAGAAGAAAAT	5	5094.3	883.2	2435.7	239727.7
pSM1266	GGGGGAAAATTTTT	7	4872.0	580.7	2544.4	478568.2
pSM1249	GAAGAGGTTATATAT	8	4303.6	583.6	1081.8	28082.5
pSM1285	GGGGGAGTTAT	4	3152.0	271.3	1591.4	465891.2
pSM1257	AGAGGGGTTAATTTT	8	2816.3	452.7	2381.3	72024.7
pSM1250	AGAGGGGTTAATTTT	8	2764.3	453.5	2381.3	72024.7
pSM1286	GGAGGGATAAT	4	2642.9	366.8	4331.3	2230524.9
pSM1258	GAAGAGGTTATATAT	8	2560.6	958.1	1081.8	28082.5
pSM1260	AAAGAGGTAATTTAA	8	2528.8	1619.9	2701.1	130741.6
pSM1290	GAGGGGGATAT	4	2064.4	244.3	1339.1	781570.0
pSM1241	GAGGGGAATTTA	5	1905.8	280.5	694.8	251072.9
pSM1271	AGAAAGGATTATTA	7	1629.3	328.7	635.0	24377.1
pSM1280	GGAGAAGAAAA	4	1515.9	242.0	2617.3	266307.7
pSM1282	GGGGGAGTATT	4	1473.6	242.0	1071.0	616218.9
pSM1278	GGAGAAAAATAATA	7	1421.1	108.4	1949.7	147860.4
pSM1262	AGAAGGGTAAATTTT	8	1213.6	156.0	1393.0	22705.7
pSM1218	AGAGAGAAATAAT	6	1177.6	174.0	819.3	47348.4
pSM1226	GAAAGAATAAAAT	6	1123.1	176.9	149.7	4002.5
pSM1277	AGAAGGAATTTATT	7	1080.9	195.7	1036.1	60311.8
pSM1288	AAGGGGGTTTT	4	1080.8	182.3	9166.9	642617.5
pSM1283	AGGGGGGTATT	4	1071.9	90.6	1770.1	1503428.4
pSM1269	GGAGAGGTTTATAT	7	1046.3	261.4	1642.5	85948.4
pSM1275	GAGGGAGAAATTTA	7	887.8	181.9	1890.0	201341.5
pSM1287	GAGGAGGTAAT	4	610.3	110.9	4328.6	1093222.2
pSM1274	AGAGAGGAATATTT	7	537.6	86.6	1557.6	172927.5
pSM1281*	GGGGGAAATAT	4	435.3	69.3	2941.3	1205635.7
pSM1279*	GGAGAAATTAT	4	424.5	73.9	3269.7	281626.6
pSM1264*	GGGAAAATAAAATT	7	402.1	75.4	157.8	47774.1
pSM1284*	GGAGAGAATAT	4	395.4	81.1	1745.1	231557.3
pSM1270*	GAAAAAATTTTTTT	7	377.8	24.6	172.5	1018.7

Supplementary References

1. Stammen S, *et al*. (2010) High-yield intra- And extracellular protein production using *Bacillus megaterium*. *Appl Environ Microbiol* 76(12):4037–46.
2. Münch KM, *et al*. (2015) Polar fixation of plasmids during recombinant protein production in Bacillus megaterium results in population heterogeneity. *Appl Environ Microbiol* 81(17):5976–5986.
3. Malten M, *et al*. (2006) A *Bacillus megaterium* plasmid system for the production, export, and one-step purification of affinity-tagged heterologous levansucrase from growth medium. *Appl Environ Microbiol* 72(2):1677–9.
4. Moore SJ, *et al*. (2016) EcoFlex - A Multifunctional MoClo Kit for *E. coli* Synthetic Biology. *ACS Synth Biol*. doi:10.1021/acssynbio.6b00031.
5. Finger C, Gamer M, Klunkelfuß S, Bunk B, Biedendieck R (2015) Impact of rare codons and the functional coproduction of rate-limiting tRNAs on recombinant protein production in *Bacillus megaterium*. *Appl Microbiol Biotechnol* 99(21):8999–9010.
6. Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 25(4):402–408.
7. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):45e–45.
8. Failmezger J, *et al*. (2016) Site-Specific Cleavage of Ribosomal RNA in *Escherichia coli*-Based Cell-Free Protein Synthesis Systems. *PLoS One* 11(12):e0168764.
9. Hoops S, *et al*. (2006) COPASI - A COmplex PAthway SImulator. *Bioinformatics* 22(24):3067–3074.
10. Sameer A, Mierle K No Title. *Ceres Solver*. Available at: http://ceres-solver.org.
11. Lange KL, Little RJA, Taylor JMG, Lange L, Taylor MG (2013) Robust Statistical Modeling Using the t Distribution. *J Am Stat Assoc* 84(408):881–896.
12. Vihola M (2012) Robust adaptive Metropolis algorithm with coerced acceptance rate. *Stat Comput* 22(5):997–1008.
13. Gelman A, Rubin DB (1992) Inference from Iterative Simulation Using Multiple Sequences. *Stat Sci* 7(4):457–472.
14. Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. *R News* 6(1):7–11.
15. Weaver LJ, *et al*. (2015) A kinetic-based approach to understanding heterologous mevalonate pathway function in *E. coli*. *Biotechnol Bioeng* 112(1):111–9.
16. Al-Majdoub ZM, Carroll KM, Gaskell SJ, Barber J (2014) Quantification of the Proteins of the Bacterial Ribosome Using QconCAT Technology. *J Proteome Res* 13(3):1211–1222.
17. Pratt JM, *et al*. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc* 1(2):1029–1043.
18. Sun ZZ, *et al*. (2013) Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J Vis Exp*

50762(79):e50762.
19. Kwon Y-C, Jewett MC (2015) High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci Rep* 5:8663.
20. Cai Q, *et al*. (2015) A simplified and robust protocol for immunoglobulin expression in *Escherichia coli* cell-free protein synthesis systems. *Biotechnol Prog* 31(3):823–831.
21. Rygus T, Hillen W (1991) Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. *Appl Microbiol Biotechnol* 35(5):594–9.
22. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25(5):955–64.
23. Krebs HA, Ruffo A, Johnson M, Egglestone LV, Hems R (1953) Oxidative phosphorylation. *Biochem J* 54(1):107–16.
24. Lou C, Stanton B, Chen Y-J, Munsky B, Voigt CA (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat Biotechnol* 30(11):1137–42.
25. Sawano a, Miyawaki a (2000) Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic Acids Res* 28(16):E78.
26. Eppinger M, *et al*. (2011) Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J Bacteriol* 193(16):4199–213.
27. Garamella J, Marshall R, Rustad M, Noireaux V (2016) The all *E. coli* TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth Biol* 5(4):344–55.
28. Underwood KA, Swartz JR, Puglisi JD (2005) Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnol Bioeng* 91(4):425–435.
29. Tomsic J, *et al*. (2000) Late events of translation initiation in bacteria: a kinetic analysis. *EMBO J* 19(9):2127–2136.
30. Siegal-Gaskins D, Tuza ZA, Kim J, Noireaux V, Murray RM (2014) Gene circuit performance characterization and resource usage in a cell-free "breadboard." *ACS Synth Biol* 3(6):416–425.
31. Dennis PP, Ehrenberg M, Fange D, Bremer H (2009) Varying rate of RNA chain elongation during rrn transcription in *Escherichia coli*. *J Bacteriol* 191(11):3740–3746.
32. Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S, Cohen SN (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* 99(15):9697–702.
33. Wang Y, *et al*. (2002) Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci U S A* 99(9):5860–5.
34. Iizuka R, Yamagishi-Shirasaki M, Funatsu T (2011) Kinetic study of de novo chromophore maturation of fluorescent proteins. *Anal Biochem* 414(2):173–8.
35. Karzbrun E, Shin J, Bar-Ziv RH, Noireaux V (2011) Coarse-grained dynamics of protein synthesis in a cell-free system. *Phys Rev Lett* 106(4):48104.
36. Verkhusha V V, Akovbian NA, Efremenko EN, Varfolomeyev SD,

Vrzheshch P V (2001) Kinetic analysis of maturation and denaturation of DsRed, a coral-derived red fluorescent protein. *Biochemistry (Mosc)* 66(12):1342–51.
37. Hebisch E, Knebel J, Landsberg J, Frey E, Leisner M (2013) High Variation of Fluorescence Protein Maturation Times in Closely Related *Escherichia coli* Strains. *PLoS One* 8(10):e75991.
38. Salis HM (2011) The ribosome binding site calculator. *Methods Enzymol* 498:19–42.