Comparison and Calibration of Different Reporters for Quantitative Analysis of Gene Expression Supporting Materials

Hernan G. Garcia, Heun Jin Lee, James Q. Boedicker and Rob Phillips

November 17, 2010

1 Promoter activities

The promoter activities in figure 1 are also shown in table S3. These correspond to various measurements found in the literature and were obtained as follows. Some promoters such as *lac* [1] and *rrnB* [2] had expression levels directly reported in Miller Units corresponding to single-copy chromosomal integrations. However, the data for the P_L and the T7 P_{A1} promoters was available in pBLA units rather than Miller units. This is measured by comparing the rate of mRNA synthesis of the promoter of interest with the rate of mRNA synthesis of the β-lactamase promoter [3]. Lanzer and Bujard [4] estimated the relation between pBLA and Miller Units, to be around 5000 MU/pBLA units.

The measurements for the P_L, T7 P_{A1} [3] and pBAD [5] promoters were performed on plasmids bearing a ColE1 origin of replication. Earlier measurements have estimated this origin of replication to result in a plasmid copy number of approximately 60 per cell [6]. As a result, we estimate the level of expression of a single copy of each of the promoter on plasmids by dividing their expression by the plasmid copy number.

Finally, the levels of expression calculated in Miller Units were converted to an absolute number of molecules using the absolute LacZ calibration described in the text of $0.6 \frac{\text{LacZ tetramers/cell}}{\text{MU}}$.

2 Supplementary Materials and Methods

2.1 Plasmids

Plasmid pZS22-YFP was kindly provided by Michael Elowitz. The EYFP gene comes from plasmid pDH5 (University of Washington Yeast Resource Center [7]). The main features of the pZ plasmids are located between unique restriction sites [6]. The sequence corresponding to the *lacUV5* promoter [8] between positions -36 and +21 was synthesized from DNA oligos and placed between the EcoRI and XhoI sites of pZS22-YFP in order to create pZS25O1+11-YFP. Note that we follow the notation of Lutz and Bujard [6] and assign the promoter number 5 to the *lacUV5* promoter. The O1 binding site in pZS25O1+11-YFP was changed to O2, O3 and to Oid using Site Directed Mutagenesis (Quikchange II, Stratagene), resulting in pZS25O2+11-YFP, pZS25O3+11-YFP and pZS25Oid+11-YFP. These plasmids are shown diagrammatically together with the promoter sequence in fig. S7.

The *lacZ* gene was cloned from *E. coli* between the KpnI and HindIII sites of all the single site constructs mentioned in the previous paragraph. The O2 binding site inside the *lacZ* coding region was deleted without changing the LacZ protein [9] using Site Directed Mutagenesis. Successful mutagenesis was confirmed by sequencing the new constructs around the mutagenized area.

After we had generated these constructs and integrated them on the *E. coli* chromosome (as described below) we determined that the different LacZ constructs had acquired some mutations. On average there were three different point mutations in each construct, though pZS25O3+11-lacZ had lost both the KpnI and HindIII sites. All these constructs still expressed functional LacZ. This problem did not present itself in the case of the EYFP constructs. We attribute this higher number of mutations in part to possible problems in the PCR amplification of the *lacZ* coding region. Another possible explanation is related to having a longer plasmid with the *lacZ* gene as opposed to the EYFP gene (3213 bp versus 714 bp). However, it must be noted that none of their EYFP counterparts had any mutations in the coding region giving less strength to this argument since a simple estimate assuming the same proportion of mutations would have resulted in roughly 1/4 the mutations seen in the LacZ case.

Plasmid pZS21-lacI was kindly provided by Michael Elowitz. This plasmid has kanamycin resistance. The chloramphenicol resistance gene flanked by FLIP recombinase sites was obtained by PCR from plasmid pKD3 [10].

The insert was placed between the SacI and AatII sites of pZS21-lacI to generate pZS3*1-lacI. The ribosomal binding sequence of pZS3*1-lacI was weakened by performing the mutation AGAGGAGAAAGG → AGA**TTT**GAAAGG (Alon Zaslaver, personal communication) resulting in pZS3*1RBS1-lacI. Higher levels of Lac repressor with respect to wild-type can be confirmed by comparing the expression of a construct such as pZS25O1+11 in the two different backgrounds.

Plasmid pET11a-His-YFP was used for the EYFP over-expression and purification described below. His-YFP was created by PCR amplifying the EYFP gene from pZS25O1+11-YFP adding a His-tag at the N-terminus of EYFP and the restriction sites for NheI and BamHI (see table S2 for primer sequences). This insert was ligated in pET11a (New England Biolabs). The resulting plasmid was transformed into strain BL21(DE3).

Plasmid pLAU53-NoLacI-TetR-YFP was constructed from plasmid pLAU53 ([11], kindly provided by Paul Wiggins). Its Lac repressor-CFP fusion was deleted by making use of the EcoRI restriction sites flanking the coding sequence. The plasmid was digested and the relevant fragment gel purified and re-ligated to obtain pLAU53-NoLacI. Though this plasmid already has TetR-YFP fusion, the EYFP it contains differs from the EYFP used in this work in some key amino acids [12]. As a result we swapped the EYFP in the plasmid for the one used in our reporter constructs. The EYFP gene was amplified from plasmid pZS25O1+11-YFP using primers HG22.03 and HG25.11 (see table S2). These primers added flanking restriction sites for HindIII and XhoI. Plasmid pLAU530-NoLacI was digested with the same restriction enzymes and the relevant fragment gel purified and ligated with the PCR product to generate pLAU53-NoLacI-TetR-YFP.

2.2 EYFP purification

His-tagged YFP was expressed in *E. coli* BL21(DE3) cells harboring the pET11a-His-YFP expression plasmid and purified using Ni-NTA affinity chromatography (Qiagen) according to the manufacturer's protocol.

2.3 *In vitro* EYFP calibration

The purified protein was diluted by 10^5 to 10^6 in PBS. Single YFP molecules bound nonspecifically to low autofluorescence (Corning D263) coverglass were imaged using a 473nm laser in epifluorescence. The rest of the imaging was done as described for the *in vivo* calibration in the main text.

2.4 Gene expression measurements

Three replicates of each strain were grown in different tubes in order to be able to obtain a mean expression level and its standard deviation. However, the day-to-day variation tended to be more significant than the variation from sample to sample on a given day. Therefore all data points shown in this experiment are the result of averaging over mean values obtained on at least three different days. The error bars are the standard deviation over these days.

The level of gene expression as a function of different IPTG concentrations was measured for strains HG104::galK<>25Oid+11, HG105::galK<>25Oid+11, HG205::galK<>25Oid+11, HG104 + pZS25Oid+11, HG105 + pZS25Oid+11 and HG205 + pZS25Oid+11 for EYFP and LacZ. For the repression measurements in EYFP strains HG104 was replaced by MG1655 and strain HG105 by TK140 as described in the main text.

The results of the induction measurements are shown in figure S8, where the levels of gene expression are normalized by their maximum. The levels of expression for both reporters were then combined to generate figures 4 and S10. The level of fluorescence was normalized by the mean fluorescence per cell of strain HG105::galK<>25O2+11-YFP, which became our fluorescence standard.

2.4.1 Single cell microscopy

Cells bearing EYFP were imaged at 100x magnification. In order to check for uniformity of the epi-illumination field we first imaged 0.5 μm fluorescent beads (TetraSpeck, Invitrogen) resulting in a typical inhomogeneity throughout the field of view of less than 5%. The cells were immobilized between a number 1.5 coverglass and a pad of 1.5% low melting temperature agarose in PBS. Images of the cells in phase contrast and fluorescence were taken. The time between the initial placement of cells on the pad and the last picture taken was about five hours. No detectable difference in the level of gene expression that could be attributed to these five hours on the pad was observed.

We used automated microscopes to take 20 snapshots per strain. Fluorescence was quantified using either a Hamamatsu Orca-285 or a Roper Scientific CoolSnap camera. With every data acquisition we quantified the reference strain HG105::galK<>25O2+11-YFP. This allowed us to directly compare the result form different the microscopy setups.

The phase contrast images were used for automatic segmentation of the

cells using custom Matlab code, based on code kindly provided by Michael Elowitz. The total fluorescence per cell was calculated by integrating the fluorescence per pixel over the cell area. The average fluorescence per cell was calculated by averaging the cell's intensity over the cell's area, as determined by the segmentation. Cell and pad autofluorescence were determined by looking at strains bearing the no reporter construct.

2.4.2 β-galactosidase assay

LacZ activity was measured by the classic colorimetric assay [13, 14] with some slight modifications as follows. A volume of the cells between 2.5 μl and 200 μl was added to Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol, pH 7.0) for a total volume of 1 ml. The volume of cells was chosen such that the yellow color would develop in no less than 15 minutes. For the case of the no-reporter constructs 200 μl of cell culture was used. Additionally, we included a blank sample with 1 ml of Z-buffer. The whole assay was performed in 1.5 ml Eppendorf tubes.

In order to lyse the cells, 25 μl of 0.1% SDS and 50 μl of chloroform was added and the mixture was vortexed for 10 s. Finally, 200 μl of 4 mg/ml 2-Nitrophenyl β-D-galactopyranoside (ONPG) in Z-buffer were added to the solution and its color, related to the concentration of the product ONP, monitored visually. Once enough yellow developed in a tube the reaction was stopped by adding 200 μl of 2.5 M Na_2CO_3 instead of adding 500 μl of a 1 M solution as done in other protocols. At this point the tubes were spun down at $> 13,000$ g for three minutes in order to reduce the contribution of cell debris to the measurement.

200 μl of solution was read for OD420 and OD550 on a Tecan Safire2 and blanked using the Z-buffer sample. The OD600 of 200 μl of each culture was read with the same instrument. The absolute activity of LacZ was measured in Miller Units using the formula

$$\mathrm{MU} = 1000 \frac{\mathrm{OD}_{420} - 1.75 \times \mathrm{OD}_{550}}{t \times v \times \mathrm{OD}_{600}} 0.826, \tag{S1}$$

where t is the reaction time in minutes and v is the volume of cells used in ml. The factor of 0.826 is not present in the usual formula used to calculate Miller Units. It is related to using 200 μl Na_2CO_3 as opposed to 500 μl. When using 500 μl, the final volume of the reaction is 1.725 ml (1ml Z-buffer, 25 μl 0.01% SDS, 200 l ONPG, 500 μl Na_2CO_3). However, when using only 200 μl of Na_2CO_3 the total volume is 1.425 ml. The factor of 0.826 adjusts for the difference in concentration of ONP.

All reactions were performed at room temperature. No significant difference in activity was observed with respect to performing the assay at 25C in an incubator.

2.4.3 Plate reader measurements

Cells were grown using the protocol described above in section "Gene expression measurements". When the culture reached an OD600 higher than 0.3 we loaded 200 μL of each culture onto a 96-well plate with a flat, clear bottom. The plate was measured in a Tecan Safire II. Fluorescence was measured from the top (height set manually to 5250 μm) with an excitation wavelength of 505 nm, and an emission wavelength of 535 nm, both with a bandwidth of 12 nm. The gain was automatically adjusted from the brightest well. Absorbance at 600 nm was also measured. After subtracting the readings from a blank sample (media without cells) we calculated the fluorescence per absorbance unit. The cell autofluorescence was obtained by performing this measurement on a strain without a fluorescent reporter. Its fluorescence per absorbance unit was subtracted from all other samples. Finally, all resulting fluorescence values were normalized by the fluorescence of strain HG105::galK<>25O2+11-YFP.

3 *In vivo* and *in vitro* single molecule measurements

Fluorescent molecules were tracked using the Matlab code "PolyParticleTracker" [15]. This code finds local maxima in an image and follows these maxima over successive frames of the same field of view. When the molecule bleaches, however, there is nothing to track and the code stops quantifying the fluorescence from that region. Additionally, for some frames, the code's selection criteria (involving area, skewness, etc.) fail to continue tracking particles which were clearly there.

In order to circumvent these two issues we modified the code such that whenever the tracking of a particle was lost it would report its last known position. Additionally, we only kept traces where the particles had been found successfully in at least the first three initial frames. This particle finding scheme resulted in a significant amount of false positives per field of view because random fluctuations in the field of view would be transiently recognized as real particles. On average, we would track on the order of 100 puncta per field of view of which no more than 10 would correspond to

real molecules. In order to distinguish the molecules from the vast number of false positives we manually screened all molecules using custom Matlab code. We confirmed that indeed there was a molecule at the position as shown by the snapshots in figure 3(A,C). We also checked that the position of the particle did not change by more than two to four pixels over the course of the analyzed trace.

The fluorescence of the selected molecules was then integrated over a box of a certain area. We present a detailed discussion regarding the choice of this parameter further below. For now, the following examples will be given using an area of 9×9 pixels, but the general conclusions are independent of this choice. Our traces consisted of 200 continuous frames with an exposure of 250 ms. However, in most cases we would only keep the fragment of the trace around the photobleaching steps as shown in figure S1(A). In figure S1(B) we present the tracking of the centroid and in figure S1(C) we show a set of snapshots over the selected time window.

The steps in the traces such as the one shown in figure S1(A) were fitted to a step or multiple step function using least squares minimization. We manually called the number of steps within a trace and the position of the transitions as starting parameters for the fit. We compared this scheme to using an automated Hidden Markov Model (HMM) approach [16]. We analyze the complete traces corresponding to molecules we had manually selected using both our manual fit scheme and and HMM. For the HMM analysis we used vbFRET, a Matlab package [16]. A sample trace with both fitting schemes is shown in figure S2(A). Notice that in the case of HMM we do not constrain the time window for each trace. By doing this over our whole data set we can compare the distributions of steps obtained by the manual fit to the distribution of discrete levels found by HMM. This is shown in figure S2(B), where it is clear that both distributions are comparable. This gives us confidence that our manual fitting approach is not significantly different from other automated schemes.

One of the main problems in measuring low fluorescent signals *in vivo* is the contribution from the cell autofluorescence [17]. When taking a time lapse this cell autofluorescence will bleach, resulting in a time-varying background. An example of this effect is shown in figure S3(A), where we present the fluorescence per pixel as a function of time of a small box located inside and outside one of the cells. We see that there is a change of about 40 counts per pixel of the cell autofluorescence over the time trace. For a box with an area of $9 \times 9 = 81$ this corresponds to a total fluorescent signal of approximately 3,000 counts, which is comparable to the magnitude of the steps we are trying to detect. A way to reduce the contribution of this effect is to

"pre-bleach" the field of view before actually taking data. Operationally, this can be done by keeping only photobleaching steps that occurred after a certain time. In figures S3(B,C) we show the effect of doing such filtering on the step size distribution. Notice that the distributions seem to converge once we only keep steps that occurred after 40 frames or later. As a result we apply this filter with a threshold of 40 frames for all our *in vivo* data.

In order to quantify both the *in vivo* and *in vitro* fluorescence of the single molecules and obtain traces such as those shown in figure 3(A,B) we integrated the signal over a box centered around their centroids. If the integration box has an area given by A and this area is bigger than the size of the diffraction limited spot corresponding to the molecule we are observing then the fluorescence before the photobleaching step is given by

$$Fluo_0 = YFP + Fluo_{back,0} \times A, \tag{S2}$$

where YFP is the fluorescence of the YFP molecule and $Fluo_{back,0}$ is the fluorescence per unit area coming from the background. This last magnitude will be a combination of the camera offset and the autofluorescence coming from the glass, buffer and, in the *in vivo* case, of the cell autofluorescence. Notice that we gave the background fluorescence the subscript "0" to denote that there might be a time dependence. After the the photobleaching step the fluorescence detected in that same area is

$$Fluo_f = Fluo_{back,f} \times A. \tag{S3}$$

In this last equation we have defined the fluorescence of the background after the photobleaching event. Of course, if the background does not change over the time course of the experiment then the difference $Fluo_f - Fluo_0$ corresponds to YFP, the fluorescence coming from a single molecule. If the level of background of baseline fluorescence is slightly different between the initial and final time points we get

$$Fluo_0 - Fluo_f = YFP + (Fluo_{back,f} - Fluo_{back,0})A. \tag{S4}$$

We worry that this effect could be present leading to a systematic error in the estimation of the fluorescence steps. This could be due to changes in background autofluorescence over time or to non-linearities in the detection of low fluorescence levels. This effect would manifest itself as a linear increase in step size with the area of integration. There are obvious limits to this equation. If the area is too small then it will not capture the total fluorescence of the diffraction limited spot. If the area is too large then it will

go beyond the cell itself and the background will be significantly different. In figure S4 we show the mean step fluorescence as a function of the size of the integration box. As expected, a small area (of 3×3 pixels^2) results in a small step size with respect to the other areas. The remaining area correspond to 5×5, 7×7 and 9×9 pixels^2. Given that a cell under our magnification conditions is about 8 to 9 pixels wide we view the area of 9×9 pixels^2 as the biggest box that can be fit within the cell. In the figure we show a fit to equation S4 and the mean resulting from averaging the over the three larger area boxes. Both the values calculated from the intercept of the linear fit and from averaging over the different data points give comparable results. We conclude that we cannot detect a difference between the background fluorescence before and after the bleaching step within our experimental error.

Interestingly, the *in vitro* values did not show a flat response of the step size as a function of the integration area in the same way than its *in vivo* counterpart did. In figure S5(A) we show the scaling of the *in vitro* step fluorescence with the area. Interestingly, we see a significant slope of (11.5 ± 0.4) au/pixel^2. If we are to believe the model in equation S4, such a slope would correspond to an underestimation of the value of the background before the bleaching step, $Fluo_{back,0}$. This shift can be clearly observed in the distributions as shown in figure S5(B). Additionally, it can be seen in the trace corresponding to a single molecule. This is shown in figure S5(C), where we shifted all the fluorescence levels after the bleaching step in order to show all the traces on the same plot. The fact that we see this effect when directly integrating the signal corresponding to a single trace without going through any automated analysis script suggests that this is not an artifact of our data analysis, but a true feature of the data.

We are unable to determine where this systematic error is coming from. This effect was also present when imaging the single molecules under Total Internal Reflection Fluorescence microscopy (TIRF), where there is a significant reduction of the background fluorescence coming from the buffer. Our main hypothesis is that it is due to a non-linearity in the acquisition process. However, we were unable to determine this unequivocally. Regardless of the origin of this systematic error if we assume that the measured step fluorescence as a function of the integration area follows a form such as the one shown in equation S4 we can use the intercept of the linear fit to account for the systematic shift and determine the real fluorescence per molecule. The intercept in the fit of figure S5(A) is 1470 ± 30 au/pixel^2. This implies that the fluorescence detected *in vitro* is 15% lower than its *in vivo* counterpart. If we were to take a box of a size comparable to the cell

size of 9×9 pixel^2 we would estimate the *in vitro* fluorescence to be higher than the *in vivo* fluorescence by about 40%. Interestingly, this result is consistent with recent measurements comparing the *in vivo* single molecule and *in vitro* bulk fluorescence of the fluorescent protein Venus [18].

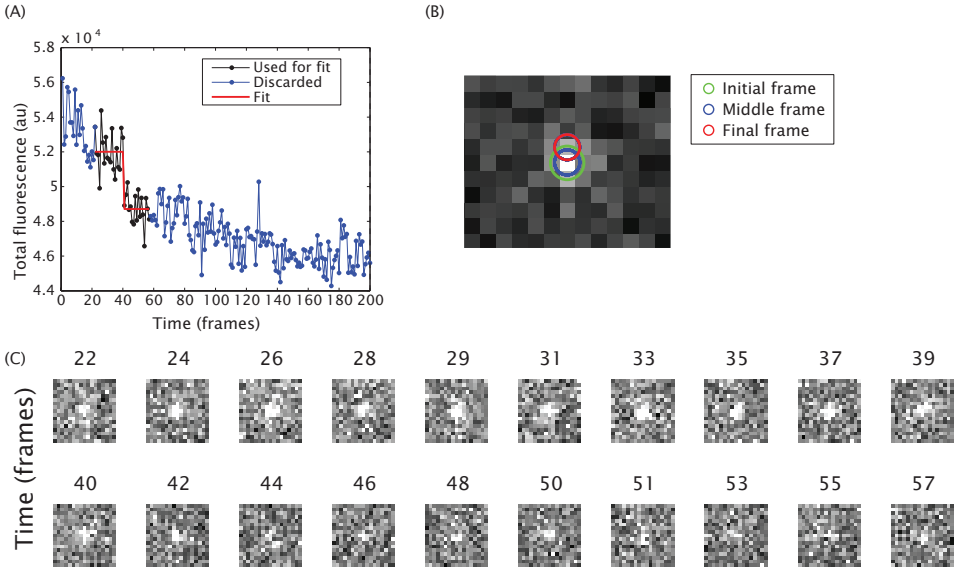


Figure S1: Selection scheme for single molecule traces. (A) A region of the trace presenting a discrete step or photobleaching is selected and fitted to a step function. (B) We confirm that the tracking of the molecule was successful by corroborating that there wasn't any significant movement of the particle over the selected time period. The image shown corresponds to the first selected frame and the circles show where the centroid of the particle was found as a function of time. (C) A window around the centroid at different time points is monitored to make sure there aren't any extraneous objects and that we do indeed have a diffraction limited spot within the integration area. The size of the window monitored is twice that of the integration window. In this example the window has a side of 17 pixels.

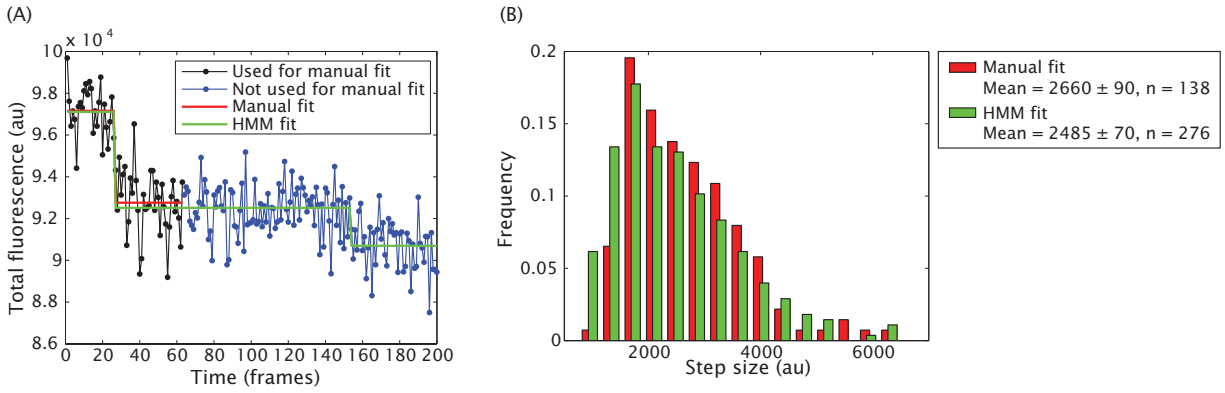


Figure S2: Comparison of different analysis schemes to obtain the fluorescence steps. (A) A single trace is shown where we fitted the step manually to a step function using least squares and using a Hidden Markov Model approach (HMM). For the latter we don't constrain the fit to a particular time window. (B) Histogram of steps obtained manually compared to a histogram of the different levels found by HMM over our whole data set.

4 Plate reader vs. microscopy for determining fluorescent levels

One of the advantages of using a fluorescent reporter to measure gene expression is that, unlike using LacZ, no further reactions are needed. Once the cells have reached the desired point in their growth all that remains is to quantify their fluorescence level. However, even though it can be highly automated, microscopy remains a slow technique when many strains are to be assayed. A compromise is to quantify fluorescence using bulk methods such as a plate reader. Such a device can query the level of fluorescence of multiple strains much faster. However, there is a price to be paid in the form of dynamic range. Whereas with microscopy we could detect down to 10 molecules/cell, with the plate reader used for this work (Tecan Safire II, see Supplementary Methods) the minimum level of fluorescence that could be detected reliably corresponded to about 50 molecules/cell. In figure S6 we show a direct comparison of the two techniques. Here, it is clear that both of them give the same result as long as the signal is not close to the detection limit.

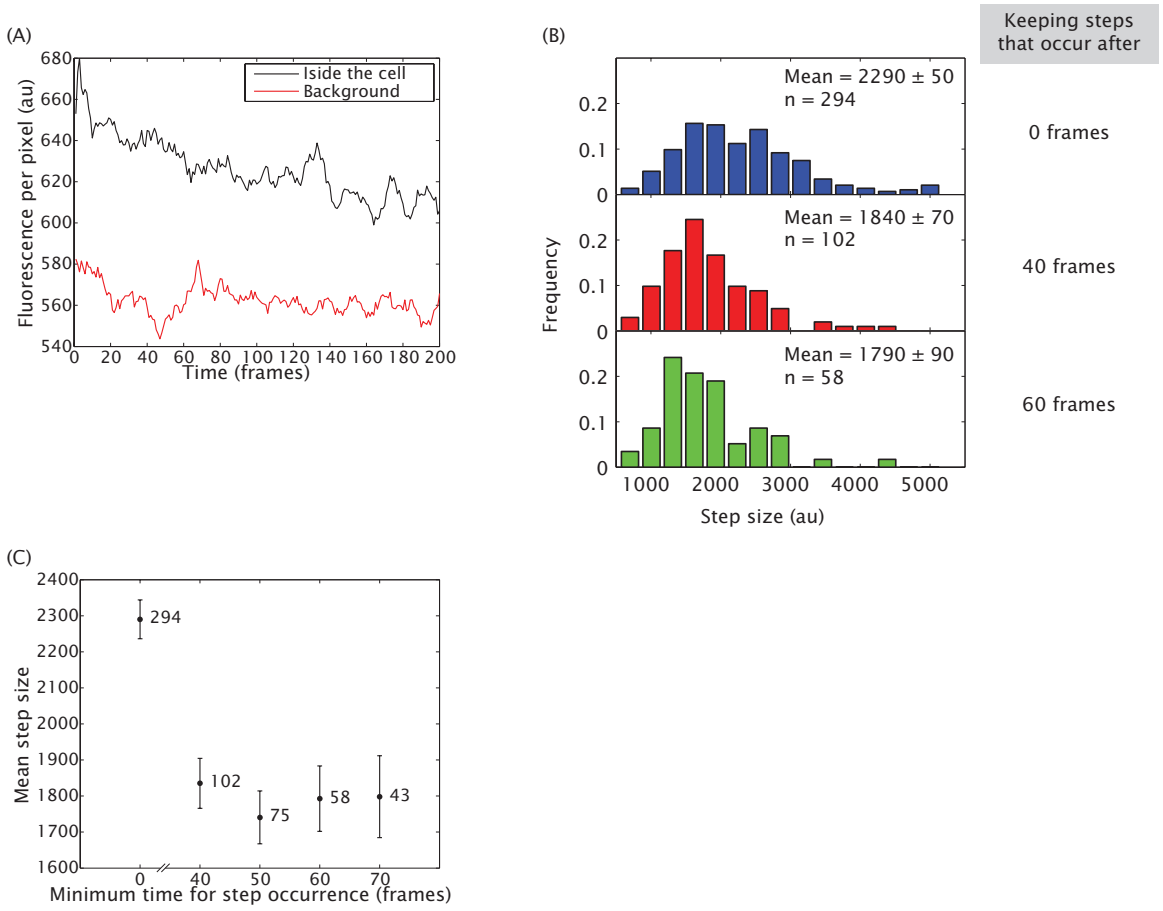


Figure S3: Effect of photobleaching on the *in vivo* calibration. (A) Fluorescence per pixel of a small region inside a cell without any fluorescent puncta and outside the cell. A moving average has been applied to smooth the traces. (B) Effect of keeping only steps that occurred after a certain time point on the step size distributions. (C) Mean step size as a function the minimum time of occurrence of steps. The error bars are the standard error of the means. The numbers next to the data points correspond to the number of steps analyzed.

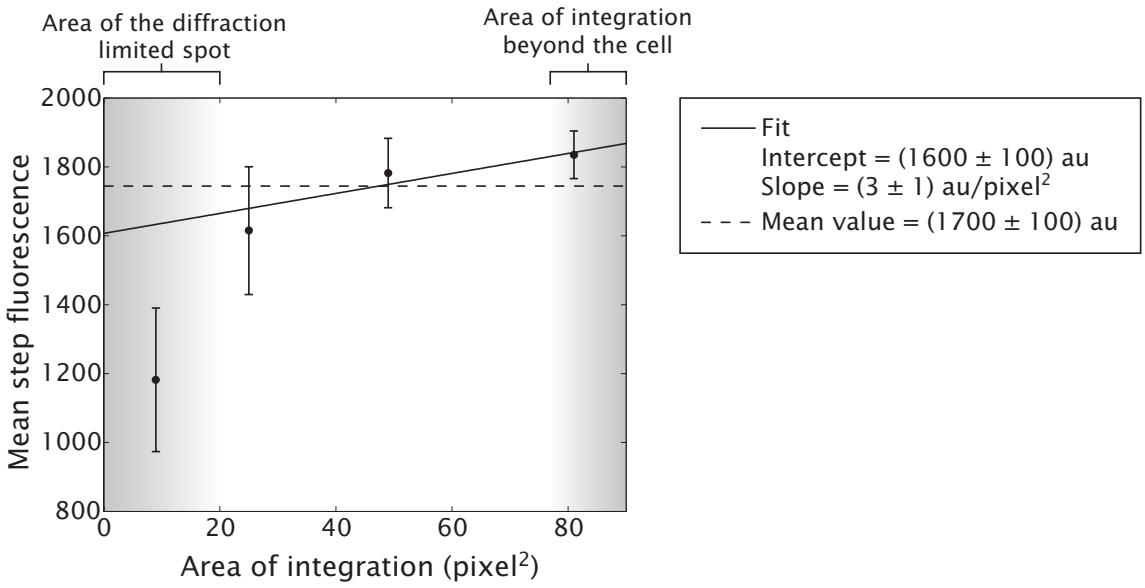


Figure S4: Scaling of the step size with the integration area for the *in vivo* calibration. The mean step size is shown as a function of the area of the integration box around the centroid of the molecule. The shaded regions mark the area sizes that are either too small to capture the fluorescence of the diffraction limited spot or too big such that the integration area would span beyond the cell. The linear fit corresponds to a fit to equation S4. This approach yields comparable values to just taking the average fluorescence step for any of the different area values.

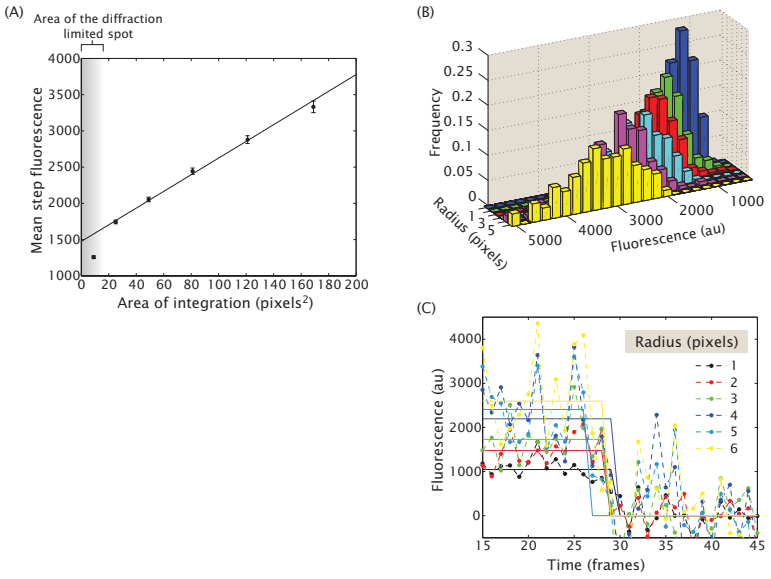


Figure S5: Scaling of the step size with the integration area for the *in vitro* calibration. (A) The mean step size is shown as a function of the area of the integration box. The fit corresponds to equation S4 with a slope of 11.5 ± 0.4 au/pixel^2 and an intercept of 1470 ± 30 au. (B) This difference in the mean step can be clearly observed at the distribution level. Here Area $= (2 \times \text{Radius} + 1)^2$. (C) Trace of fluorescence vs. time obtained for a single molecule for different choices of the size of the integration window. We have shifted the fluorescent level of all traces after photobleaching so that they would coincide for easier comparison.

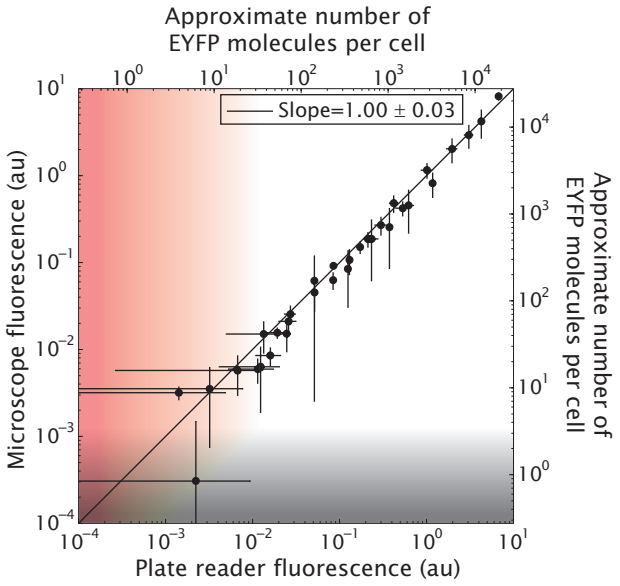


Figure S6: Comparison of microscopy and plate reader as methods to quantify gene expression. The same strains were quantified both using microscopy and using a plate reader (Tecan Safire II). The results show a 1:1 correlation between plate reader and microscopy, although the plate reader has a lower limit of detection which is greater than that of microscopy. Whereas using microscopy we can detect as few as roughly 10 EYFP molecules/cell, the plate reader can detect molecules only in excess of concentrations of 50 molecules/cell approximately. These detection limits are marked by the shaded regions.

References

[1] T. Kuhlman, Z. Zhang, Jr. Saier, M. H., and T. Hwa. Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 104(14):6043–8, 2007.

[2] S. Liang, M. Bipatnath, Y. Xu, S. Chen, P. Dennis, M. Ehrenberg, and H. Bremer. Activities of constitutive promoters in *Escherichia coli*. *J Mol Biol*, 292(1):19–37, 1999.

[3] U. Deuschle, W. Kammerer, R. Gentz, and H. Bujard. Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *EMBO J*, 5(11):2987–94, 1986.

[4] M. Lanzer and H. Bujard. Promoters largely determine the efficiency of repressor action. *Proc Natl Acad Sci U S A*, 85(23):8973–7, 1988.

[5] R. R. Seabold and R. F. Schleif. Apo-AraC actively seeks to loop. *J Mol Biol*, 278(3):529–38, 1998.

[6] R. Lutz and H. Bujard. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res*, 25(6):1203–10, 1997.

[7] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–5, 2005.

[8] Benno Mller-Hill. *The lac Operon: a short history of a genetic paradigm*. Walter de Gruyter, Berlin; New York, 1996.

[9] S. Oehler, E. R. Eismann, H. Kramer, and B. Muller-Hill. The three operators of the lac operon cooperate in repression. *EMBO J*, 9(4):973–9, 1990.

[10] K. A. Datsenko and B. L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, 97(12):6640–5, 2000.

[11] I. F. Lau, S. R. Filipe, B. Soballe, O. A. Okstad, F. X. Barre, and D. J. Sherratt. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol*, 49(3):731–43, 2003.

[12] R. Y. Tsien. The green fluorescent protein. *Annu Rev Biochem*, 67:509–44, 1998.

[13] J. H. Miller. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

[14] N. A. Becker, J. D. Kahn, and L. J. Maher III. Bacterial repression loops require enhanced DNA flexibility. *J Mol Biol*, 349(4):716–30, 2005.

[15] S. S. Rogers, T. A. Waigh, X. Zhao, and J. R. Lu. Precise particle tracking against a complicated background: polynomial fitting with Gaussian weight. *Phys Biol*, 4(3):220–7, 2007.

[16] Jonathan E. Bronson, Jingyi Fei, Jake M. Hofman, Ruben L. Gonzalez Jr., and Chris H. Wiggins. Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *arXiv:0907.3156v1*, 2009.

[17] G. S. Harms, L. Cognet, P. H. Lommerse, G. A. Blab, and T. Schmidt. Autofluorescent proteins in single-molecule research: applications to live cell imaging microscopy. *Biophys J*, 80(5):2396–408, 2001.

[18] Yuichi Taniguchi, Paul J. Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X. Sunney Xie. Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science*, 329:533, 2010.

[19] P. Lu, C. Vogel, R. Wang, X. Yao, and E. M. Marcotte. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol*, 25(1):117–24, 2007.

5 Supplementary figures and tables

Table S1: List of *E. coli* strains used throughout this experiment. Chromosomal positions correspond to the sequence in GenBank accession no. U00096.

Strain	Alternative name	Genotype	Derived from	Comment
HG104	lacI+	Δ*lacZYA*	MG1655	Deletion from 360,483 to 365,579
HG105	lacI-	Δ*lacZYA*, Δ*lacI*	MG1655	Deletion from 360,483 to 366,637
HG205	lacI++	Δ*lacZYA*, Δ*lacI*, ybcN<>3*1RBS1-lacI	HG105	
TK140		Δ*lacI*	MG1655	[1]

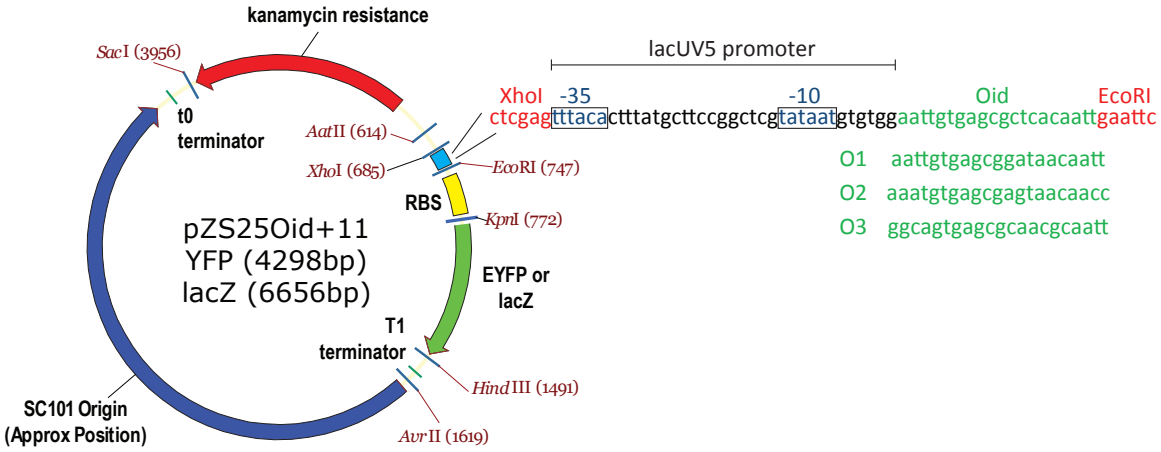


Figure S7: Plasmid diagram and promoter sequence. The main features of the plasmids pZS25O1+11-YFP and pZS25O1+11-lacZ are shown flanked by unique restriction sites. The particular promoter sequence based on the *lacUV5* promoter is shown together with the sequences of the different Lac repressor binding sites used.

Table S2: Primers used throughout this work. For integration primers, lowercase indicates the portion of the primer that is homologous to the *E. coli* gene where the integration is made and uppercase indicates primer homology to the plasmid where PCR was carried out.

Primer	Sequence	Comment
HG6.1	gtttgcgcgcagtcagcgatatccattttcgcgaatccggagtg taagaaACTAGCAACACCAGAACAGCC	Integration of the EYFP and *lacZ* reporter constructs into the *galK* gene.
HG6.3	ttcatattgttcagcgacagcttgctgtacggcaggcaccagct cttccgGGCTAATGCACCCAGTAAGG	
HG11.1	acctctgcggaggggaagcgtgaacctctcacaagacggcatca aattacACTAGCAACACCAGAACAGCC	Integration of pZS3*1RBS1-lacI into the *ybcN* gene.
HG11.3	ctgtagatgtgtccgttcatgacacgaataagcggtgtagccat tacgccGGCTAATGCACCCAGTAAGG	
HG22.03	attatagctagcatgggtcatcaccatcaccatcacggtcgtaa aggagaagaacttttcactgg	Make His-EYFP and insert into pET11a.
HG22.03R	tattaatggatccttatttgtatagttcatccatgccatgt	
HG25.02	atattaaagcttatttgtatagttcatccatgccatg	Fuse TetR for EYFP in pLAU53-NoLacI
HG22.11	attatctcgagttggtgcgtaaaggagaagaacttttcactgg	
HG22.04	gtttgcgcgcagtcagcgatatccattttcgcgaatccggagt gtaagaaTTAATGCGCCGCTACAGGG	Integration of pLAU53-NoLacI-TetR-YFP into the *galK* gene.
HG22.05	ttcatattgttcagcgacagcttgctgtacggcaggcaccagc tcttccgTACTTTTCATACTCCCGCCATTCA	

Table S3: Promoter activities from the literature measured using the LacZ assay. These activities from the literature have been measured for a range of different promoters and conditions. The promoter strengths quoted here are often approximate and should therefore not be considered as accurate. Refer to "Promoter activities" in these Supplementary Materials for details of how these activities were calculated. (1) Measured in M9 + 0.5% glucose. (2) For a cell doubling of about 1.25/h. (3) Calibrating P_{bla} activity and LacZ units [4]. (4) Assuming a plasmid copy number of 60 copies/cell [6].

Promoter	Strength (MU)	Reference	Comment
lac promoter, no IPTG	0.5-0.6	[1]	(1)
lac promoter, 1 mM IPTG	600-700	[1]	(1)
rrnB P1-P2	3,000	[2]	(2)
P_L, no cI	3,200	[3]	(3,4)
T7 P_{A1}	6,400	[3]	(3,4)
pBAD, no arabinose	7	[5]	(4)
pBAD, 2% arabinose	580	[5]	(4)

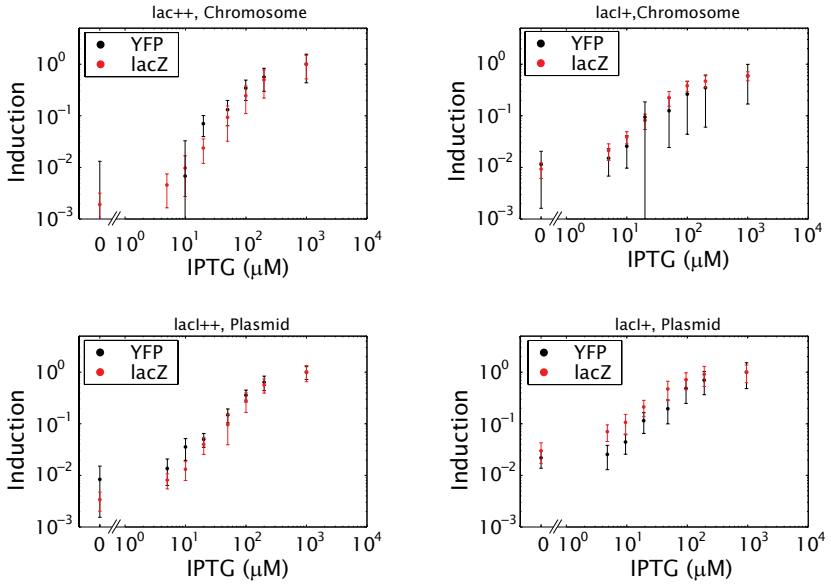


Figure S8: Induction curves used in this work. The level of gene expression of the EYFP and *lacZ* constructs is shown as a function of the IPTG concentration for the different construct locations (chromosome or low copy plasmid) and strain background. The level of expression is normalized by the corresponding maximum levels of activities. Error bars correspond to the standard deviation of measurements performed over at least four different days. Refer to the Materials and Methods for a description of the different strains.

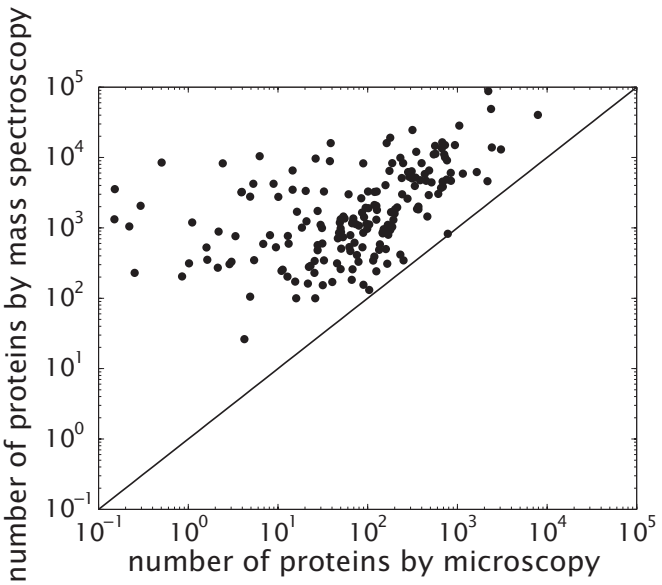


Figure S9: Comparison of different *E. coli* cell censuses. The number of proteins for a particular protein measured by mass spectroscopy [19] is compared to the same magnitude measured by fluorescence in single cells [18]. As a reference, a black line with a slope of one is plotted in order to emphasize the systematic disagreement between the two techniques.

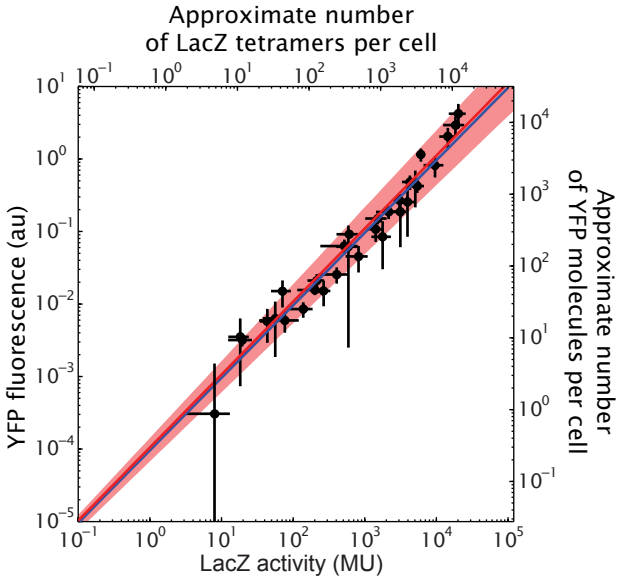
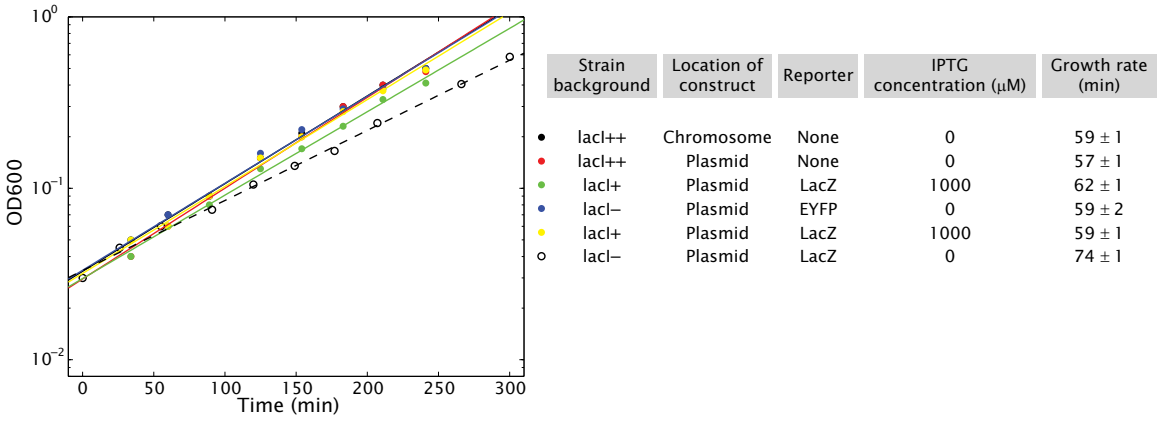


Figure S10: Relation between the mean cell fluorescence and the β-galactosidase activity. The total fluorescence per cell is plotted against the β-galactosidase activity. Each point corresponds to the same promoter bearing either EYFP or *lacZ* as a reporter in the same strain background and at the same concentration of IPTG. The blue line is a linear fit fixing the intercept to zero (see figure 4). The red line is a fit to a power law with a resulting exponent of 1.01 ± 0.05, consistent with a linear relation between the two reporters. The shaded area is defined by the standard error for the power law fit.



Strain background	Location of construct	Reporter	IPTG concentration (μM)	Growth rate (min)
lacI++	Chromosome	None	0	59 ± 1
lacI++	Plasmid	None	0	57 ± 1
lacI+	Plasmid	LacZ	1000	62 ± 1
lacI–	Plasmid	EYFP	0	59 ± 2
lacI+	Plasmid	LacZ	1000	59 ± 1
lacI–	Plasmid	LacZ	0	74 ± 1

Figure S11: Effect of reporter proteins on growth rate. The different levels of EYFP assayed in this work do not affect the growth of the cells significantly. However, high β-galactosidase levels slow down growth in a detectable fashion. This level of LacZ is reached when our plasmid reporter is present in strain lacI-.