Dissecting specific and global transcriptional regulation of bacterial gene expression

Luca Gerosa, Karl Kochanowski, Matthias Heinemann, Uwe Sauer

Corresponding author: Uwe Sauer, ETH

Review timeline:

Submission date: 18 December 2012
Editorial Decision: 08 February 2013
Revision received: 28 February 2013
Accepted: 06 March 2013

Editor: Thomas Lemberger

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 08 February 2013

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who accepted to evaluate the study. As you will see, the referees find the topic of your study of potential interest and are supportive. They raise however a series of concerns and make suggestions for modifications, which we would ask you to carefully address in a revision of the present work. The recommendations provided by the reviewers are very clear in this regard.

Thank you for submitting this paper to Molecular Systems Biology.

Referee reports:

Reviewer #1 (Remarks to the Author):

In this manuscript, Gerosa et al studied the interplay and the relative contribution of global transcription regulation and the specific regulated transcription. They chose to focus on the arginine biosynthesis genes in E. coli as a model system since these genes are regulated by a single repressor, argR, for which the regulatory regions on the different promoters are well characterized. The authors elegantly measured first the global transcription regulation by using genes that are under a global regulation only as well as the global regulation of the arginine biosynthesis genes in a delta argR strain. These experiments provided the contribution of the global transcriptional regulation. Using the extracted parameters (Vmax and Km) they then studied the specific argR transcriptional regulation.

The study revealed two main intriguing principles: (1) Regulated transcription, namely the repression, dominates most of the time, except for the transient states during a shift from external uptake to biosynthesis. (2) Global transcriptional regulation sets the maximal point to which
promoter activity can reach during the short transition times in metabolic states.

The drawn conclusions considerably enhance our understanding of the design principles underlying transcriptional networks, and the authors' approach lays the ground for additional similar studies to evaluate how common these transcriptional paradigms are. The experiments, their analyses and the modeling parts were carefully performed. In addition, the required controls were included and so are the necessary statistical tests.

There are, however, several issues I wish the authors could address:

1. The authors show in figure 1 that promoter activity follows a Michaelis-Menten relationship with growth rate. This is extracted from 18 different growth conditions each allowed the cells to grow in a different rate. Would promoter activities of the genes that are under global transcriptional regulation (including arg genes in delta argR background) follow growth rate during the course of the culture growth? That is, does culture growth rate (logOD over time) correlate with promoter activity? In particular, In figure 2, promoter activities are shown over many hours of growth and they decline towards the end. I assume this is because growth rate is reduced and cells enter stationary phase (Is there a lag phase at the beginning of the growth?). Does growth rate correlate with promoter activity over time throughout these growth phases?

2. On page 12 line 26 the authors state that activity peaks far above the steady state level. It would be nice to give the actual numbers/percent for the different genes with mean +/-std, and also to include how close (% of max) are the peaks to the transcription capacity?

3. On page 14 lines 17-18 the parameters for the model and fitting are given. How is the fitting procedure sensitive to the initial chosen parameters? In addition, why ArgR=800 and arg=71? If these arbitrary units correspond to concentration or absolute numbers, then I believe the values should be switched, as there are likely much more of arginine molecules in the cell than arginine repressor molecules.

4. In figure 4a, why to show the simulated constitutive argA activity rather than the measured one when comparing to measured regulated promoter activity?

5. In figure 4, it is not clear how the authors determined the different phases p1,p2 and p3. As there are no measurements of intracellular arginine concentration, do they rely on the drop in argR activity shown in Fig 4c? Can they show that growth rate is slowed down at the transition from p1 to p2 due to the drop in arg concentration and the consequent metabolic shift?

6. Fig S10, for some genes, the simulated results poorly correlate with the measured ones. In particular for argR, which is central to the analysis of the entire arg regulon. Can the authors explain why this is, and whether the poor argR match between simulations and measured data has any effect on the correctness and accuracy of the model?

7. In Tables S4 and S5, the predicted constitutive promoter activity gets pretty low correlation and R^2 values for several strains and conditions. For example, mean argG Pearson correlation is 0.28 and growth in Ace results in correlation of 0.27 only. Even more, R^2 values are low for most genes in many conditions. Why is that? I understand that when considering the entire data set the statistics becomes significant, but having such a low correlation for each strain/condition may suggest for additional parameters/inputs in the system that the simulation does take into account. Could the authors resolve this issue?

Typos:
1. page 28, line 17 - simulated.
The supplementary part has many typos:
1. page 4 line 2: protein concentration is a function....
2. page 4, 2nd para. Line5: exponential
Review: Gerosa et al.

The authors examine the interplay between circuit-specific and cellular-level mechanisms of transcriptional regulation during a cell's transient response to change in growth conditions. These two types of regulation have long been studied and discussed, but as pointed by the authors, a detailed quantitative examination of their relative role, in particular during non-steady-state conditions, has been lacking. The current manuscript describes such an investigation, combining quantitative experiments with theoretical analysis.

The subject matter and study method are all appropriate for MSB. The work appears well conceived and is clearly presented. I have a number of concerns that I'd like to see addressed before publication can be considered.

A recurring problem throughout the manuscript is the following: While the phenomenological models used by the authors do a very good job of describing the dependence of constitutive promoter activity on growth rate (Figure 1) and the transient response to change in growth conditions (Figures 2, 4, 5), the authors repeatedly insist on going beyond these phenomenological descriptions and invoke molecular mechanisms that are not directly supported by the data (and have not been directly shown in past studies). Specifically:

1) In deriving Equation 1, the authors state that "global regulation is predominantly exerted by free RNAP 70" (page 7). What is the evidence for the "predominance" of RNAP availability over other factors, such as nucleoid morphology and DNA topology? (see e.g. Kuhlman and Cox, MSB 2012).

2) The authors also state that "specific transcription factors [...] modulate the recruitment of RNAP 70 to the promoter" (page 7). Again, this is not generally true, as some bacterial transcription factors modulate other steps in the initiation of transcription, see e.g. H. Garcia et al., Cell Reports 2012; A. Sanchez et al., EMBO J 2011.

3) Later on, the authors take the success of their phenomenological model (Figures 1 & 2) as supporting evidence that "the free RNAP 70 concentration is proportional to the growth rate (P 70 )" (page 9). This seems an overreach. The success of the model is indeed encouraging, but it represents the success of describing promoter activity as a function of growth rate (Figure 1), independent of any specific mechanistic molecular interpretation.

4) In interpreting the Arginine circuit data, the authors state: "Calculations show the ArgR repressor bound to enzymatic promoters more than 80% of the time during arginine biosynthesis" (page 12). Here too a specific molecular interpretation, according to which gene activity in simply dependent to transcription-factor occupancy, is invoked. What evidence do they have for such a simple relation, and more broadly for the validity of Equation 3? (For a counter example where transcription-factor occupancy is not directly mapped to gene activity, see H. Garcia et al., Cell Reports 2012).

The point of items 1-4 above is not to disqualify the main narrative of the paper. On the contrary, the authors' success in predicting the transient gene-expression response (Figures 2, 4, 5) based on steady-state measurements (Figure 1) is significant and exciting. The point is that by needlessly
making molecular statements that are not supported by the data the authors weaken their case rather than strengthen it.

Additional comments:

5) The authors distinguish "steady state growth" from "dynamic growth" (page 4). The latter term is awkward, as all growth is dynamic. The term "transient growth" used later is also inappropriate, since this is not a case of transient growth between two periods of non-growth, but rather a transient change in growth rate.

6) The authors justify the use of GFP-based reporters for measuring transcriptional activity by pointing out that other gene expression parameters, especially translation rate and mRNA degradation rate, do not vary with growth rate (page 6). While the argument appears valid, I was wondering whether the authors (or any previous studies) have directly validated it by comparing the GFP-based estimates to direct RNA-level measurements such as QPCR or FISH (e.g. L. So et al., Nature Genetics 2011).

7) On the same issue: The authors list physiological parameters known to be growth-rate-independent, but the GFP maturation time is not discussed. Since variation in maturation time will directly affect the estimation of promoter activity, it's important to know whether it changes with growth rate, and how.

8) The authors describe the conversion of transcription-factor regulated promoters to "constitutive" ones, by altering their transcription-factor binding sites (page 8). Was the abrogation of regulation by the transcription factors verified by comparing expression level in the presence and absence of the transcription factor?

9) Promoter activity is stated in units of GFP/OD/h (page 9 and elsewhere). How was the absolute number of GFP molecules extracted from the data? If the units of fluorescence measurement are arbitrary rather than a real protein count, the notation should be changed.

10) The authors refer to the "acceleration" and "transition ... growth phases" (page 10). These terms are not standard and need to be defined.

11) Does Figure 2 contain only a subset of the promoters in Figure 1? Where is the remaining data? If it exists, it should be plotted in the Supplementary Material.

12) The notations in Figures 2, 4 & 5 are unclear: At what time point is the change in growth condition applied? Is it the time marked with an arrow in Figures 2 and 4? The exact time point, as well as the conditions before and after, need to be clearly marked in all figures.

13) In Figure 5, an in-plot legend needs to be added clarifying what the different colors and numbers denote.

14) "While reconstruction of topologies is steaming..." (page 16). Is this a typo?

Reviewer #3 (Remarks to the Author):

This manuscript addresses a timely and important (but often underappreciated) issue, namely the indirect effects of whole-cell physiological parameters on gene expression. The authors address this issue for biosynthetic genes (specifically argentine biosynthesis) where the coupling to global physiology (growth) is obviously unavoidable. In my view this is the first paper that addresses such issues in a real physiological context. The study combines modern experimental methods, in particular a library of fluorescent reporters, with theoretical analysis and the accumulated knowledge from traditional bacterial growth physiology in an original way. The paper is well written and the topic and scope of the study are very well suited for Molecular Systems Biology. This said, I have a few concerns with the methodology that need to be addressed and some results may not be as straightforward to interpret as the authors claim. These issues are described in detail below.

1) There are a few issues with the use of the GFP reporter library technique. Do all strains grow with...
the same rate under the same conditions? Is there an effect of GFP expression on growth? Growth in the 96 well plates is unavoidably anaerobic due to the sealing, while traditional growth studies were done in aerated cultures. I am aware that these are disadvantages of a technique that have to be weighed against its obvious benefits, but addressing them as far as possible would make comparison to the traditional studies easier.

2) At the end of the study the authors claim that their method is better than normalising to a constitutive control case as done in other recent studies (the very recent MSB paper by Berthoumieux et al is another example). I think it should be admitted that the method here is limited to special cases such as the one here, where all growth rate dependence can be attributed to one known parameter, namely RNA polymerase concentration. For example, growth rate independence of the plasmid concentration is required, which cannot be expected in general. In fact I was very surprised to see how constant the plasmid concentration is here.

3) Growth curves are not shown. Since growth is one of the key parameters here, I think growth curves are needed. This would allow to judge whether the cells actually ever reach adjusted balanced (exponential) growth and also to correlate dynamic changes in growth with changes in gene expression.

4) It is assumed based on earlier studies that translation rates are not affected by growth changes. The earlier studies have addresses steady state situations only. It is not clear whether that condition holds true in transient dynamics. Specifically it has been shown long ago that the expression of an abundant mRNA will transiently deplete the pool of free ribosomes (Vind et al. J Mol Biol 1993). For the assumption to be valid here the depletion must either be small so its effect can be neglected or occur on shorter times than what is studied here.

5a) The discussion of the two hypotheses about the growth rate dependence of the free RNA polymerase concentration on p.8 is not correct as it stands. First of all I think, biochemical Km values (in units of concentrations) and effective Km values obtained from the fit (in inverse time units) should be clearly distinguished to avoid confusion.

5b) In addition the interoperation of the data is more complex than suggested here. In both scenarios discussed here, the promoter activity is described by a Michaelis-Menten function of the free RNAP concentration (r). In one case (Liang et al), the RNAP concentration is taken as linear in the growth rate, r=μ/q, which results in a Michaelis-Menten dependence on growth rate with a effective Km (K_growth) given by the K_growth=Km*q (where Km is the Michaelis constant of the promoter). In that case clearly Km values will be promoter dependent and K_growth values will be as well.

In the other (Klumpp et al), it is assumed that the promoters are typically unsaturated and that the apparent saturation is due to the growth rate dependence of r, which happens to resemble a Michaelis-Menten function. The measured Michaelis constant (K_growth) will be independent of the promoter, but only if all promoters are indeed unsaturated. In the general case, both functions will be non-linear, but the result will again resemble a Michaelis-menten relation, a linear increase plus saturation, with an effective Km that is promoter dependent.

In short, what the authors can conclude is that some promoters become saturated as a function of growth rate, but they cannot make conclusions about the functional form of the RNAP growth rate dependence. I do not object to using the linear assumption but it is an assumption.
under a global regulation only as well as the global regulation of the arginine biosynthesis genes in a delta argR strain. These experiments provided the contribution of the global transcriptional regulation. Using the extracted parameters (Vmax and Km) they then studied the specific argR transcriptional regulation. The study revealed two main intriguing principles: (1) Regulated transcription, namely the repression, dominates most of the time, except for the transient states during a shift from external uptake to biosynthesis. (2) Global transcriptional regulation sets the maximal point to which promoter activity can reach during the short transition times in metabolic states.

The drawn conclusions considerably enhance our understanding of the design principles underlying transcriptional networks, and the authors’ approach lays the ground for additional similar studies to evaluate how common these transcriptional paradigms are. The experiments, their analyses and the modeling parts were carefully performed. In addition, the required controls were included and so are the necessary statistical tests.

We are grateful for the reviewer’s very positive assessment of our work.

There are, however, several issues I wish the authors could address:

1. The authors show in figure 1 that promoter activity follows a Michaelis-Menten relationship with growth rate. This is extracted from 18 different growth conditions each allowed the cells to grow in a different rate. Would promoter activities of the genes that are under global transcriptional regulation (including arg genes in delta argR background) follow growth rate during the course of the culture growth? That is, does culture growth rate (logOD over time) correlate with promoter activity? In particular, In figure 2, promoter activities are shown over many hours of growth and they decline towards the end. I assume this is because growth rate is reduced and cells enter stationary phase (Is there a lag phase at the beginning of the growth?). Does growth rate correlate with promoter activity over time throughout these growth phases?

Culture growth rate does indeed correlate with promoter activity over the time course of the experiments. We now added a figure comparing the measured promoter activity time courses for all 12 constitutive promoters in all 19 growth conditions (Figure S10) to the simulated promoter activities which were calculated directly from the corresponding growth rates and the promoter-specific parameters Vmax and Km. Thus, the good agreement of measured and simulated promoter activities suggests that growth rate is indeed a good predictor of global regulation throughout the different growth phases.

Moreover, as the reviewer rightfully suggested, the decline in promoter activity towards the end is most likely caused by the concurrent reduction in growth rate as cells approach stationary phase. To better illustrate this point, we added a figure showing the growth curves for both ArgR knockout strain and the wild type in each of the 19 growth conditions (Figure S1). This data shows that usually cells start growing immediately, albeit the growth rate typically shows an initial acceleration at the beginning of the time course, as well as a decline in growth rate as cells approach stationary phase. Please note that our experiments were designed to minimize potential lag phases by growing precultures in the same medium to adapt cells.

2. On page 12 line 26 the authors state that activity peaks far above the steady state level. It would be nice to give the actual numbers/percent for the different genes with mean +/-std, and also to include how close (% of max) are the peaks to the transcription capacity?

We agree that such a plot could be helpful and added Figure S14. The figure shows, for each arginine promoter in each of the 3 time course experiments, the promoter activity in the 3 phases: the average promoter activity (and standard deviation) during growth on external supplement of arginine (E) and arginine biosynthesis (B), and the maximal activity reached by the promoter during the transition phase (T). We also show the average constitutive promoter activity across the whole time course (and standard deviation) as a visual reference, so that a quantitative comparison between the level reached by regulated promoters and promoter capacity is now intuitive.

3. On page 14 lines 17-18 the parameters for the model and fitting are given. How is the fitting procedure sensitive to the initial chosen parameters? In addition, why ArgR=800 and arg=71? If these arbitrary units correspond to concentration or absolute numbers, then I believe the values
should be switched, as there are likely much more of arginine molecules in the cell than arginine repressor molecules.

The initial condition values for the ArgR concentration (800) and the arginine concentration (71) were correct as reported. The two variables have different units: ArgR repressor is given in GFP/OD and arginine concentration is in arbitrary units since we do not have absolute measurements for arginine. The two values should not be interpreted as comparable abundances. We note that absolute concentrations of arginine are not required in the model, since arginine concentration can freely be scaled with the corresponding kinetic parameters, $k_p$ and $k_{cat}$, $k_{arg}$ and $r_0$. To clarify this, we now included Table S8 providing the unit of measurements, as well as a description, for each parameter, variable and input in the ODE model.

The fitting procedure was performed by random sampling of initial arginine concentrations and fixing the ArgR concentration to the average of initial measurement in the three arginine depletion experiments (Figure 4). Thus, we cannot assess the sensitivity of the fitting procedure to initial parameters and cannot guarantee that the identified parameter set is unique or independent of initial conditions. However, we can estimate a more important sensitivity, the sensitivity to initial values for the simulation of the double perturbation experiment (Figure 5). We ran simulations by changing the ArgR and arginine (arg) initial values within a broad range (600<=ArgR<=100, 50<=arg<=90). We found that the overall model is still very accurate in predicting the promoter activity response of the pathway (Figure S17). Thus, simulations of the double perturbation experiment are robust to a wide range of initial values.

4. In figure 4a, why to show the simulated constitutive argA activity rather than the measured one when comparing to measured regulated promoter activity?

The reason is that the ArgR knockout strain has a different growth physiology compared to the wild type strain. We quantified the difference as an average 27% reduction in steady state growth rate (see new supplementary figure S1), also seen when comparing time-course growth in Fig.2 of the main text, or Fig.S2 and Fig S3. The growth difference is arguably given by the burden of a constantly up-regulated arginine pathway (which may explain the benefit of strong transcriptional regulation by the ArgR repressor also when arginine biosynthesis is required). Given that growth rate is the key parameter of constitutive promoter activity, constitutive promoter activity measured in an ArgR knockout strain can thus not be directly superimposed on the regulated promoter activity measured in the wild type strain. Calculation of the constitutive promoter activity using a parameterized quantitative model is, in our view, the only way to obtain an estimate of constitutive promoter activity that is comparable to measured regulated activity. This is the reason for reporting the simulated and not the measured constitutive promoter activity for argA (and for all the arginine promoters as well). We regard this as one of the advantages for accounting constitutive activity in the model explicitly, rather than correcting for it. To make this aspect more clear in the manuscript, we modified Figure 4a and now show promoter capacity only in the subplot presenting the simulation results.

5. In figure 4, it is not clear how the authors determined the different phases p1,p2 and p3. As there are no measurements of intracellular arginine concentration, do they rely on the drop in argR activity shown in Fig 4c? Can they show that growth rate is slowed down at the transition from p1 to p2 due to the drop in arg concentration and the consequent metabolic shift?

We determined the boundary between external supplement (E) and transition (T) phases by selecting the time point at which we observed up-regulation of the arginine promoters and a slowing down of the growth rate. The slowing down of growth rate is seen for glucose minimal medium and glucose minimal medium with amino acids, but is not evident in the galactose condition, possibly due to the already slow growth. For the second boundary between the transition (T) and the biosynthesis (B) phases, we selected the time point at which we observed the reaching of a new steady value in promoter activity and enzyme concentration. We have now added in Figure S13 an insert that shows the growth rate of cells during the three time course experiments. We would like to highlight that the three phases should be regarded as a visual guideline for interpreting the dynamic regulatory response, rather than precise time points indicating molecular events.

6. Fig S10, for some genes, the simulated results poorly correlate with the measured ones. In particular for argR, which is central to the analysis of the entire arg regulon. Can the authors
explain why this is, and whether the poor argR match between simulations and measured data has any effect on the correctness and accuracy of the model?

The observation is correct. There are two reasons why this discrepancy, in our opinion, does not in practice affect the accuracy and correctness of the overall model: i) inaccuracy is only in the order of \( \sim 2 \)-fold increased predicted promoter activity after arginine depletion and ii) changes in the repressor concentration are not the major driver of the transcriptional response, which is arginine concentration. To make this evident, we added Figure S16 in which we show the time courses for ArgR, ArgA, ArgR* activity and arginine concentration as predicted by the model. As evident, changes in arginine concentrations are larger and dominate the establishment of the ArgR* activity profile.

Nevertheless, we took up the challenge and tried to pin-point the possible reason for the inaccuracy in argR promoter activity predictions. The argR reconstructions are accurate during full repression (i.e. before arginine runs out) but less so during partial derepression after arginine depletion. In particular, the data reveal a small burst of promoter activity followed by a return to the full repression state, while simulations predict a higher burst and a subsequent maintenance of a higher derepressed level. Thus, possible causes are: 1) the existence of an unknown regulator for argR that is not included in the model, 2) inaccurate estimates of promoter-specific parameters for argR (Vmax, Km, Kr), or 3) a regulatory mechanism different from the assumed simple mutual exclusion between the ArgR repressor and a single RNAP.

We reasoned that the existence of an unknown regulator (case 1) is unlikely, since promoter activity for constitutive argR is well reconstructed in the corresponding experiments and we observed only global regulation (Figure 1, Figure S10, Table S6, Table S7). We tested the possibility of having obtained inaccurate estimates of parameters from steady state data (case 2) by re-estimating Vmax, Km and Kr parameters form the measured argR promoter activity and the ArgR* activity that explains enzymatic promoters in the double perturbation experiment. Since we could not identify any parameter set that better reproduced the data (data not shown), we conclude that the mismatch cannot be explained by the parameter estimation. This leaves a different mechanism of regulation for the argR promoter than the one assumed (case 3) as the most likely explanation for the mismatch. For example, is possible that the argR promoter contains multiples RNA polymerase binding sites that are differently affected by ArgR repression, that the ArgR repressor binds synergistically to the promoter, that the unbound form of ArgR repressor is also able to bind the promoter, or that the repression is not through a mechanism of mutual exclusion. While unraveling the exact mechanism of expression repression for the argR promoter is potentially interesting, it is outside the scope of this work. Most importantly, the exact molecular nature of the mechanism does not affect any of the conclusions in our manuscript!

7. In Tables S4 and S5, the predicted constitutive promoter activity gets pretty low correlation and \( R^2 \) values for several strains and conditions. For example, mean argG Pearson correlation is 0.28 and growth in Ace results in correlation of 0.27 only. Even more, \( R^2 \) values are low for most genes in many conditions. Why is that? I understand that when considering the entire data set the statistics becomes significant, but having such a low correlation for each strain/condition may suggest for additional parameters/inputs in the system that the simulation does take into account. Could the authors resolve this issue?

First of all we noted during the revision a small error in the calculation of Pearson correlations and coefficient of determinations (formerly Tables S4 and S5; now Tables S6 and S7). The new coefficients differ only 2% in average from the previous values, so quantitative and qualitative statements are unaltered. We nevertheless apologize for this minor correction.

To answer the point raised by the reviewer, we believe that mismatches between simulated and experimental data arise mainly because of noise, which is mostly due to the numerical derivation step needed to calculate promoter activity (the dGFP/dt component) and growth rate (dlog(OD)/dt).

In the case of the Pearson correlation, the promoter argG is the only one with low correlation values. argG is by far the lowest expressed among our promoter reporters and thus is particularly effected by noise (see also Figure S15a to observe noise effects for argG that are not present for other promoters). Regarding the coefficient of determination, we note that low (or negative) values arise when the mean of the promoter activity data provides a reconstruction that is equal (or superior) to our model prediction. As a consequence, obtaining high coefficient of determinations by predictions (not by fitting) is a very hard task, because the comparison made is between predictions without any parameter adjustment (our model) to the mean of the data which is known only after the actual
measurement. Thus, we regard obtaining low or negative coefficient of determination as not invalidating the predictive power of the model. In particular, the mean of promoter activity data is likely to be an already very good fit for constitutive promoters measured in growth conditions where the growth rate is constant for many hours. This seems to be the case for some of conditions with a low coefficient of determination, such as growth on acetate.

This said, we cannot exclude that there are no unaccounted additional biological parameters or inputs that could explain certain mismatches for certain combinations of promoters and conditions. We can only state that we were not able to identify any clear pattern that could be attributed to such putative biological signals. Thus, we regard mismatches as stemming from the inherent noise in measurements and numerically derivation.

Typos:
1. page 28, line 17 - simulated.
2. The supplementary part has many typos:
   1. page 4 line 2: protein concentration is a function....
   2. page 4, 2nd para. Line5: exponential
   3. page 4, 3rd para. Line 5: constant or a function of...
   4. page 4, 3rd para. Line 10: inserted promoter.
   5. page 4, 3rd para. Line 13: across growth rates both...
   6. Figure s1 - I could find data for only 12 conditions while there should be 13 conditions.
   7. page 6 line 3: (...scales on Y-axis)
   8. page 6 line 9: cell as a function of... was calculated by multiplying...
   9. page 7 line 2: Michaelis
10. Fig S3 and Fig S4 have the same legend so it should be corrected and so are the references to these figures on the top paragraph on page 8.
11. page 13, 2nd line : estimates of ...(missing)...is shown
12. Title of Fig S7a: constant

Thanks for the extra effort to help sort out typos that are now corrected. Regarding to point 6, glucose minimal medium with supplemented amino acids was by mistake mentioned twice in the list of growth conditions. We now removed it, as indeed Figure S5 (previously Figure S1) shows data for 12 and not for 13 growth conditions.
Reviewer #2 (Remarks to the Author):

The authors examine the interplay between circuit-specific and cellular-level mechanisms of transcriptional regulation during a cell’s transient response to change in growth conditions. These two types of regulation have long been studied and discussed, but as pointed by the authors, a detailed quantitative examination of their relative role, in particular during non-steady-state conditions, has been lacking. The current manuscript describes such an investigation, combining quantitative experiments with theoretical analysis.

The subject matter and study method are all appropriate for MSB. The work appears well conceived and is clearly presented. I have a number of concerns that I’d like to see addressed before publication can be considered.

We are very grateful for the reviewer’s positive comments on our work.

A recurring problem throughout the manuscript is the following: While the phenomenological models used by the authors do a very good job of describing the dependence of constitutive promoter activity on growth rate (Figure 1) and the transient response to change in growth conditions (Figures 2, 4, 5), the authors repeatedly insist on going beyond these phenomenological descriptions and invoke molecular mechanisms that are not directly supported by the data (and have not been directly shown in past studies).

A fair point. It was never our intention to unravel the molecular mechanisms but rather to build a phenomenological model consistent with the extensive molecular and physiological information available for E. coli. Since some mechanistic information was available in the literature and the subject had been extensively debated, we merely wanted to point out whether our results are consistent with certain views or how different underlying molecular mechanisms would effect (or rather not) our conclusions. We have now modified the text to make extra clear that we are not attempting to identify the molecular mechanisms themselves.

Thanks to this reviewer’s comment and similar comments from reviewer #3 (5a and 5b), we realized that issues about molecular interpretations arise uniquely around the role of free RNA polymerase and its functional dependency on the growth rate. Because these interpretations are not essential to our analysis or conclusions regarding global regulation, we decided to drop this interpretation when interpreting/discussing constitutive promoter activity.

We now describe the constitutive promoter activity always with regard to growth rate, and thus modified accordingly part of the model explanation. We changed Equation 2 and 3 by substituting $P_\sigma^{70}$ with the growth rate $\mu$ and modified the labels in the figure 1, 2 and 3 accordingly. We want to emphasize that this is merely a change of molecular interpretation of the same results because growth rate was always the actual measurement entering Equation 2. The conclusion that we are observing global regulation at the level of transcription is also unchanged, because it is based on the reported growth dependency within the gene expression cascade, independently from the free RNA polymerase interpretation.

Now we invoke the mechanistic interpretation of growth rate (as a proportional proxy to free RNP concentration) only when introducing the regulated promoter activity model in the arginine pathway. In our view a mechanistic motivation must be brought forward at this stage to explicitly clarify what are the assumptions used to derive Equation 3. Any other alternative formulation of Equation 3 would have to take on a number of similar assumptions.

We hope the reviewer will find that this modification (mainly from pag.6 line 24 to pag.7 line25) does help in clearly separating the phenomenological results from the mechanistic interpretations in our work.

Specifically:

1) In deriving Equation 1, the authors state that "global regulation is predominantly exerted by free RNP$\sigma^{70}$" (pager 7). What is the evidence for the "predominance" of RNP availability over other factors, such as nucleoid morphology and DNA topology? (see e.g. Kuhlman and Cox, MSB 2012).

We had made this statement based on the reported growth dependence of RNA polymerase availability and its global effect on the expression of constitutive promoters (e.g. from Ehrenberg et al., 2010 (PMID 19835927), Liang et al., 1999 (PMID 10493854) and Klumpp et al., 2008/2009.
(PMID 19073937, 20064380)). The evidence put forward in these studies suggests that free RNA polymerase availability constitutes at least a major determinant of global regulation. However, we agree with the reviewer, this evidence does not exclude other possible influential global and growth-dependent forms of regulation. Since we do not interpret transcriptional regulation in terms of free RNA polymerase anymore when discussing constitutive promoter activity, we removed this sentence.

2) The authors also state that "specific transcription factors [...] modulate the recruitment of RNAPσ70 to the promoter" (page 7). Again, this is not generally true, as some bacterial transcription factors modulate other steps in the initiation of transcription, see e.g. H. Garcia et al., Cell Reports 2012; A. Sanchez et al., EMBO J 2011.

The reviewer is absolutely right. We thank her/him for pointing out this aspect. We modified the respective statement (pag.6,line 25-27) as follows:

"...specific regulation of genes through transcription factors that can regulate the recruitment of RNAP polymerase (Bintu et al, 2005) as well as other steps in transcription (Garcia et al, 2012)(Sanchez et al, 2011)."

3) Later on, the authors take the success of their phenomenological model (Figures 1 & 2) as supporting evidence that "the free RNAPσ70 concentration is proportional to the growth rate (Pσ70 ∝ µ)" (page 9). This seems an overreach. The success of the model is indeed encouraging, but it represents the success of describing promoter activity as a function of growth rate (Figure 1), independent of any specific mechanistic molecular interpretation.

The reviewer is right, we now refrain from a strict molecular interpretation in terms of free RNA polymerase when discussing the constitutive promoter activity data. Given the observed difference in Km values, we now simply conclude that there is a promoter-specific component in the growth rate dependency of constitutive gene expression, hence the different Km values (pag.9,line 3-8). Free RNAP may well be the molecular basis for this dependency but it does not have to be.

4) In interpreting the Arginine circuit data, the authors state: "Calculations show the ArgR repressor bound to enzymatic promoters more than 80% of the time during arginine biosynthesis" (page 12). Here too a specific molecular interpretation, according to which gene activity in simply dependent to transcription-factor occupancy, is invoked. What evidence do they have for such a simple relation, and more broadly for the validity of Equation 3? (For a counter example where transcription-factor occupancy is not directly mapped to gene activity, see H. Garcia et al., Cell Reports 2012).

The reviewer is right in pointing out that transcription factor occupancy is not in all cases directly mapped to gene activity. However, we argue that in the specific case of ArgR this direct mapping is justified for two reasons. First, the ArgR binding sites on the promoters were shown to overlap with the RNA polymerase binding site (PMID 1640456). Second, it was shown by in vitro transcription assays that ArgR inhibits promoter activity by steric exclusion of RNA polymerase (PMID 17850814) for promoters whose ArgR and RNA polymerase binding sites overlap. Based on this experimental evidence, we conclude that at least for the promoters considered in this manuscript, mutual exclusion of ArgR and RNA polymerase is the most likely molecular mechanism of repression, which is adequately described by Equation 3 using the framework described in (PMID 15797194, 12702751).

The point of items 1-4 above is not to disqualify the main narrative of the paper. On the contrary, the authors' success in predicting the transient gene-expression response (Figures 2, 4, 5) based on steady-state measurements (Figure 1) is significant and exciting. The point is that by needlessly making molecular statements that are not supported by the data the authors weaken their case rather than strengthen it.

Thanks again for the valuable and constructive remarks that we fully agree with. In the revised version we addressed the main criticism about unneeded/unsubstantiated references to molecular mechanisms (in terms of free RNAP concentration).

Additional comments:
5) The authors distinguish "steady state growth" from "dynamic growth" (page 4). The latter term is awkward, as all growth is dynamic. The term "transient growth" used later is also inappropriate, since this is not a case of transient growth between two periods of non-growth, but rather a transient change in growth rate.

The reviewer is right, we changed occurrences of “dynamic growth” and “transient growth” to “dynamic changes in growth rate” in all instances.

6) The authors justify the use of GFP-based reporters for measuring transcriptional activity by pointing out that other gene expression parameters, especially translation rate and mRNA degradation rate, do not vary with growth rate (page 6). While the argument appears valid, I was wondering whether the authors (or any previous studies) have directly validated it by comparing the GFP-based estimates to direct RNA-level measurements such as QPCR or FISH (e.g. L. So et al., Nature Genetics 2011).

We have not directly compared GFP-based estimates to RNA-level measurements, and we are also not aware of studies which have done such a comparison at different growth rates. Comparing mRNA levels determined by FISH and RNA-Seq with levels of YFP-fusions of about 1000 E. coli proteins, Taniguchi et al., 2010 (PMID 20671182) found that mean mRNA levels and mean protein levels correlate (see Figure 3C of that work). Since only one condition was analyzed, this evidence is unfortunately of little help in assessing growth rate effects. The evidence for translation rate and mRNA degradation to be independent of growth rate are from classical works on steady state physiology and gene expression of E. coli, as summarized in Table S3. We agree that a more systematic comparison between GFP-based estimates and RNA-level measurements across growth rates would be extremely helpful to validate the assumption on translation rate constancy, but it has not happened so far.

7) On the same issue: The authors list physiological parameters known to be growth-rate-independent, but the GFP maturation time is not discussed. Since variation in maturation time will directly affect the estimation of promoter activity, it's important to know whether it changes with growth rate, and how.

Indeed there could be a potential contribution of GFP maturation time to promoter activity data, and we are not aware of any study directly investigating the growth rate dependency of GFP maturation times. Based on the extremely fast maturation time of the here used GFP variant (GFPmut2), we believe that potential differences in GFP maturation are unlikely to affect our findings. The reported maturation time of GFPmut2 is approximately 5-7 minutes (PMID 16862137, 21459075), which is shorter than our measurement interval (10 minutes). Therefore, differences in maturation time would not significantly change the interpretation of our time course experiments, as long as the maturation time does not differ several-fold between conditions, for which we have found not found any evidence in the literature. Moreover, the response of promoter activity to arginine depletion seems to be fast for all three steady state (Figure 4) and changing growth rates (Figure 5). We added a statement in the description of gene expression parameters, which now includes these comments about GFP maturation times.

8) The authors describe the conversion of transcription-factor regulated promoters to "constitutive" ones, by altering their transcription-factor binding sites (page 8). Was the abrogation of regulation by the transcription factors verified by comparing expression level in the presence and absence of the transcription factor?

The absence of regulation was verified for the constitutive forms of the pykF, kbl and epd promoters by comparing the constitutive and the regulated promoter activity (Figure 1). Additionally, we measured the promoter activity of the promoter pykF, which is controlled by Cra, in 8 different conditions and in the following combination: 1) the constitutive version of pykF in wild type strain, 2) the regulated version of pykF in Cra knockout strain, 3) the constitutive version of pykF in the wild type strain. We observed a very similar quantitative dependency between promoter activity and growth rate (Figure S4), allowing us to conclude that alteration of the transcription factor binding site was effective and did not change the kinetic properties of the promoter. In general, comparison between constitutive promoter activity obtained in wild type and
transcription factor knockout strains is difficult when knockouts have severe growth phenotypes in many conditions; i.e. there is simply not enough of a growth rate range to assess growth rate dependency of promoter activity. This was the case for the Crp and Lrp knockout strains that regulated the epd and kbl promoters. Thus, we performed such test only on pykF.

9) Promoter activity is stated in units of GFP/OD/h (page 9 and elsewhere). How was the absolute number of GFP molecules extracted from the data? If the units of fluorescence measurement are arbitrary rather than a real protein count, the notation should be changed.

The units of fluorescence measurement that we termed as “GFP” are arbitrary (the raw reading value of our plate reader) and thus do not represent a real protein count. We used this notation to ensure consistency with the various publications utilizing the same library of transcriptional reporters (PMID 15107854, 16862137, 19914165), which all used the GFP/OD/h (or GFP/OD notation). We have now added a comment in the Material and Methods section to clarify that GFP is in arbitrary units. We hope the reviewer agrees with this solution for the sake of consistency.

10) The authors refer to the "acceleration" and "transition ... growth phases" (page 10). These terms are not standard and need to be defined.

We changed the sentence to “from inoculation to stationary phase” (pag. 9, line 27).

11) Does Figure 2 contain only a subset of the promoters in Figure 1? Where is the remaining data? If it exists, it should be plotted in the Supplementary Material.

Figure 2 does indeed contain only a subset of the promoters in Figure 1 and also only a subset of the conditions that we tested. The full dataset is now plotted in Figure S10.

12) The notations in Figures 2, 4 & 5 are unclear: At what time point is the change in growth condition applied? Is it the time marked with an arrow in Figures 2 and 4? The exact time point, as well as the conditions before and after, need to be clearly marked in all figures.

The arrows in figures 2 and 4 did not denote any specific time points – we merely used them to relate the labels for conditions and promoters to the corresponding columns and rows in the subplot. We now removed the arrows and thank the reviewer pointing this ambiguous notation out. For all figures there is NO external change that could be marked. Growth changes exclusively because of the available nutrients in the medium; i.e. carbon sources in Figure 2 and initial amounts of arginine in Figures 4 and 5. The legends were reorganized for clarity.

13) In Figure 5, an in-plot legend needs to be added clarifying what the different colors and numbers denote.

We have added the in-plot legend as suggested.

14) "While reconstruction of topologies is steaming..." (page 16). Is this a typo?

The wording was changed.
Reviewer #3 (Remarks to the Author):

This manuscript addresses a timely and important (but often underappreciated) issue, namely the indirect effects of whole-cell physiological parameters on gene expression. The authors address this issue for biosynthetic genes (specifically argentine biosynthesis) where the coupling to global physiology (growth) is obviously unavoidable. In my view this is the first paper that addresses such issues in a real physiological context. The study combines modern experimental methods, in particular a library of fluorescent reporters, with theoretical analysis and the accumulated knowledge from traditional bacterial growth physiology in an original way. The paper is well written and the topic and scope of the study are very well suited for Molecular Systems Biology. We thank the reviewer for the favorable evaluation of our work.

This said, I have a few concerns with the methodology that need to be addressed and some results may not be as straightforward to interpret as the authors claim. These issues are described in detail below.

1) There are a few issues with the use of the GFP reporter library technique. Do all strains grow with the same rate under the same conditions? Is there an effect of GFP expression on growth?

Growth in the 96 well plates is unavoidably anaerobic due to the sealing, while traditional growth studies were done in aerated cultures. I am aware that these are disadvantages of a technique that have to be weighed against its obvious benefits, but addressing them as far as possible would make comparison to the traditional studies easier.

A very reasonable point. We evaluated such effects at the beginning of this work by performing a series of control experiments and analysis. We now included these control experiments under the SI section title “Strain, growth conditions and physiology”. We have added a text that in detail describes each experiment and the two corresponding new supplementary Figures S2 and S3. Briefly, we summarize the conclusions here:

1. No observable effect of GFP expression on growth, since we obtain very accurate similar growth rates for all strains, from low to high GFP expressing plasmids, including a strain carrying a plasmid that does not express GFP (see Figure S1).

2. We have extensive experience with the 96 well plate format and know from physiological experiments and flux analysis that our cultures were never in an anaerobic regime. In fact, they have the same physiology as in shake flasks (see Figure S5). Different from other laboratories, we do not cover our cultures with mineral oil because evaporation was negligible for the duration of our experiments.

2) At the end of the study the authors claim that their method is better than normalising to a constitutive control case as done in other recent studies (the very recent MSB paper by Berthoumieux et al is another example). I think it should be admitted that the method here is limited to special cases such as the one here, where all growth rate dependence can be attributed to one known parameter, namely RNA polymerase concentration. For example, growth rate independence of the plasmid concentration is required, which cannot be expected in general. In fact I was very surprised to see how constant the plasmid concentration is here.

The reviewer raises a valid question concerning the general applicability of our approach, but we agree only in part with the suggested restriction. It is true that explicit representation of global regulation requires existence and knowledge about a quantitative relationship between each parameter in the gene expression and growth rate. In our particular case, constant plasmid concentration across growth rates permitted to omit direct inclusion of the parameter g/v (gene concentration) in the model. Varying plasmid concentrations, however, could have explicitly been included in Equation 1 (and 2 and 3) by adding an additional growth rate dependent term g (gene concentration). Thus, our method is very well applicable also to other situations with multiple growth dependent parameters, but it does require that they are quantified. Absence or non-measurable quantitative relationships between any of the parameters and growth rate would indeed hamper application of our method, and the text (pag. 18 , line 6-7) was altered to make this clear: “...we developed a superior approach based on first principles to explicitly include the promoter-specific growth dependency into models, under the constraint that parameters in the gene expression system can be expressed as a function of the growth rate”.
3) Growth curves are not shown. Since growth is one of the key parameters here, I think growth curves are needed. This would allow to judge whether the cells actually ever reach adjusted balanced (exponential) growth and also to correlate dynamic changes in growth with changes in gene expression.

We have added two supplementary Figures (Fig. S2 and Fig. S3) that show optical density and growth rate of each of the 19 time-course growth experiments measured for the wild type and ArgR knockout strains. We have also highlighted the time periods that we selected as exponential growth phase. Standard deviations across replicates have been added to permit an evaluation of the accuracy in growth physiology.

4) It is assumed based on earlier studies that translation rates are not affected by growth changes. The earlier studies have addresses steady state situations only. It is not clear whether that condition holds true in transient dynamics. Specifically it has been shown long ago that the expression of an abundant mRNA will transiently deplete the pool of free ribosomes (Vind et al. J Mol Biol 1993). For the assumption to be valid here the depletion must either be small so its effect can be neglected or occur on shorter times than what is studied here.

The reviewer raises an interesting point. It is certainly conceivable that massive induction of a protein can deplete the pool of free ribosomes. We argue that such effects are not likely to occur in the experiments presented here. The lacZ high-copy expression constructs used in Vind et al. are very strong, yielding beta-galactosidase concentrations of up to 25% of total protein, which are orders of magnitude higher than the GFP expression by the low-copy reporters under natural promoters used in this manuscript.

5a) The discussion of the two hypotheses about the growth rate dependence of the free RNA polymerase concentration on p.8 is not correct as it stands. First of all I think, biochemical Km values (in units of concentrations) and effective Km values obtained from the fit (in inverse time units) should be clearly distinguished to avoid confusion.

We agree with the reviewer and apologize for the confusion. Following this comment and comments from reviewer #2, we now introduce and interpret the Michaelis-Menten model for constitutive promoter activity in terms of growth rate. So, we now clearly state that we infer effective Km parameters in inverse time units. Please see the response to the main comment of reviewer #2 for more details.

5b) In addition the interoperation of the data is more complex than suggested here. In both scenarios discussed here, the promoter activity is described by a Michaelis-Menten function of the free RNAP concentration (r). In one case (Liang et al), the RNAP concentration is taken as linear in the growth rate, r=μ=q, which results in a Michaelis-Menten dependence on growth rate with an effective Km (K_growth) given by the K_growth=Km*q (where Km is the Michaelis constant of the promoter). In that case clearly Km values will be promoter dependent and K_growth values will be as well.

In the other (Klumpp et al), it is assumed that the promoters are typically unsaturated and that the apparent saturation is due to the growth rate dependence of r, which happens to resemble a Michaelis-Menten function. The measured Michaelis constant (K_growth) will be independent of the promoter, but only if all promoters are indeed unsaturated. In the general case, both functions will be non-linear, but the result will again resemble a Michaelis-menten relation, a linear increase plus saturation, with an effective Km that is promoter dependent.

In short, what the authors can conclude is that some promoters become saturated as a function of growth rate, but they cannot make conclusions about the functional form of the RNAP growth rate dependence. I do not object to using the linear assumption but it is an assumption.

The point is absolutely correct. From our results, we can only conclude that constitutive promoter activity has a promoter-specific component in its observed non-linearity with growth rate. The revised version does not assume a relationship between free RNA polymerase and growth rate anymore. We now simply suggest that different effective Km parameters are consistent with a promoter saturation by RNA polymerase at high growth rates. The proportionality between free RNA polymerase and growth rate is now only used to develop Equation 3 for regulated arginine promoters, and we now clearly state that it is an assumption. Please see the response to the main comment of reviewer #2 for more details.