Robustness of DNA looping across multiple cell divisions in individual bacteria

Chang Chang†, Mayra García-Alcalá‡, Leonor Saiz‡, José M. G. Vilar†,‡, and Philippe Cluzel‡

Edited by Terence Hwa, Department of Physics, University of California San Diego, La Jolla, CA; received January 3, 2022; accepted July 9, 2022

DNA looping has emerged as a central paradigm of transcriptional regulation, as it is shared across many living systems. One core property of DNA looping–based regulation is its ability to greatly enhance repression or activation of genes with only a few copies of transcriptional regulators. However, this property based on a small number of proteins raises the question of the robustness of such a mechanism with respect to the large intracellular perturbations taking place during growth and division of the cell. Here we address the issue of sensitivity to variations of intracellular parameters of gene regulation by DNA looping. We use the lac system as a prototype to experimentally identify the key features of the robustness of DNA looping in growing Escherichia coli cells. Surprisingly, we observe time intervals of tight repression spanning across division events, which can sometimes exceed 10 generations. Remarkably, the distribution of such long time intervals exhibits memoryless statistics that is mostly insensitive to repressor concentration, cell division events, and the number of distinct loops accessible to the system. By contrast, gene regulation becomes highly sensitive to these perturbations when DNA looping is absent. Using stochastic simulations, we propose that the observed robustness to division emerges from the competition between fast, multiple rebinding events of repressors and slow initiation rate of the RNA polymerase. We argue that fast rebinding events are a direct consequence of DNA looping that ensures robust gene repression across a range of intracellular perturbations.

memory | noise | single-cell | microfluidics | robustness

Some genetic regulatory systems in bacteria are known to use only a few repressors to maintain low levels of expression, such as lac, ara, and lysogenic regulations (1–3). The diversity of these examples underlines the importance for these systems to have selected certain molecular mechanisms for efficiently maintaining low expression levels together with low levels of repressors. The lac operon is arguably among the most studied genetic regulatory systems of this class and is known to utilize a higher-order structure of DNA, DNA looping, to efficiently repress the activity of the lac promoter using only a handful of copies of repressors (1, 4–6). While the strong repression mediated by DNA looping has clearly been established in vivo and in vitro, the fact that it relies on a small number of repressors to function makes this molecular mechanism potentially sensitive to intracellular perturbations. For example, a small number of repressors can fluctuate greatly at cell division, which may yield undesirable promoter leaks (7–9), and it is still an open problem to know whether DNA looping can maintain repression even across several divisions. Indeed, it is standard to assume that gene duplication and cell division may disrupt the looping structure and binding of the repressors to DNA (10), which would, consequently, limit the duration of repression intervals. Moreover, during cellular growth, DNA replicates and gene dosage increases as a function of time, which may dynamically alter the ratio of the number of DNA binding sites to that of repressors.

In light of these outstanding questions, we aimed at quantitatively characterizing how robust the repression of DNA looping is with respect to intracellular perturbations in individual growing bacteria. We combine several techniques to record and analyze the spontaneous leakiness of the lac system (Fig. 1). In our experiments, we monitor the spontaneous leakiness of the lac promoter, as a measure for the repression level of the promoter in the presence or absence of DNA looping. Using a microfluidic device, we record long time series associated with promoter leakiness in individual growing Escherichia coli cells across more than 40 generations. We use, as a starting point, the model by Vilar and Leibler (11) that proposed that the change of free energy associated with DNA looping formation is equivalent to the existence of a very large “local” repressor concentration, effectively hundred times larger than the “global” wild-type repressor concentration (12). One key prediction of this model is that repression by means of DNA looping is robust to fluctuations of repressor concentration, with low levels of repressors. The robust gene repression across a range of intracellular perturbations by DNA looping has clearly been established in vivo and in vitro, the fact that it relies on a small number of repressors raises the question of the robustness of such a mechanism with respect to the large intracellular perturbations taking place during growth and division of the cell. Here we address the issue of sensitivity to variations of intracellular parameters of gene regulation by DNA looping. We use the lac system as a prototype to experimentally identify the key features of the robustness of DNA looping in growing Escherichia coli cells. Surprisingly, we observe time intervals of tight repression spanning across division events, which can sometimes exceed 10 generations. Remarkably, the distribution of such long time intervals exhibits memoryless statistics that is mostly insensitive to repressor concentration, cell division events, and the number of distinct loops accessible to the system. By contrast, gene regulation becomes highly sensitive to these perturbations when DNA looping is absent. Using stochastic simulations, we propose that the observed robustness to division emerges from the competition between fast, multiple rebinding events of repressors and slow initiation rate of the RNA polymerase. We argue that fast rebinding events are a direct consequence of DNA looping that ensures robust gene repression across a range of intracellular perturbations.

Significance

It is well established that certain intracellular regulators can stabilize DNA loops to greatly enhance activation or repression of gene transcription. In vitro and in vivo ensemble measurements have determined that only a few copies of regulators are, in fact, needed to stably form DNA loops. In view of such a small number, we address the issue of sensitivity of gene regulation by DNA looping to variations of intracellular parameters in individual growing Escherichia coli bacteria. Surprisingly, we find that DNA looping from the lac system is robust to a range of perturbations, including divisions during which cells can maintain tight repression over many generations. We propose molecular hypotheses compatible with the observed robustness across a range of intracellular perturbations.
while repression in the absence of DNA looping is sensitive to fluctuations. Here, we experimentally investigate the consequences of this theoretical hypothesis in the broader context of cell division at the single-cell level.

**Experimental Design**

Due to the very low leaking rate, we monitor the promoter activity from single cells and across many division cycles using a microfluidics device, called the "mother machine" (13, 14). In this device, cells grow under chemostatic conditions. The mother cell is trapped at the bottom of a microfluidic channel, while daughter cells are washed away when they exit the channel. In our experiments, to estimate promoter activity, we use the production rate of a fluorescent reporter driven by a copy of the native lac promoter. To measure gene expressions with an improved temporal resolution and signal-to-noise ratio, we use a fast-maturing fluorescent protein [VenusNB, maturation half-time 4.1 ± 0.3 min (15)], together with an optimized ribosome binding site to maximize the yield of translation of the fluorescent reporter (16). Maximizing the yield of translation of fluorescent reporter per messenger RNA (mRNA) helps us to detect small transcriptional bursts that may not be directly detectable at the single-cell level. It was found earlier that a large fraction of cells do not contain even one copy of the fluorescent protein controlled by the lac operon (17); thus the fluorescent level of most of the cells is often nearing the level of autofluorescence. Under this condition of such a low expression level, we further developed a probabilistic algorithm that explicitly takes into account the fluctuations of autofluorescence background to robustly discriminate promoter activity from background noise (*SI Appendix*).

Cells were cultured overnight in M9 medium with 0.4% glycerol as a carbon source. The native lac operon is under the combinatorial control of both lac repressor and cAMP, Cyclic adenosine monophosphate, Receptor Protein (CRP) (18). It is well established that carbon sources such as glycerol or glucose are associated with relatively higher or lower cAMP levels (19), and that cAMP–CRP can either promote a high transcription initiation rate or mediate repression with DNA looping (18). However, it was shown that, in the absence of inducer (i.e., our conditions), both carbon sources yield similar low levels of promoter activity (18), and, therefore, the particular choice of carbon source is not expected to profoundly affect the spontaneous leakiness of the promoter (*SI Appendix*, Fig. S7).

Overnight, cells were loaded in the mother machine and cultured at exponential growth by steadily flushing them with fresh media (≥40 h; 30 °C; also see *SI Appendix*, Fig. S6 for control experiments at 35 °C). Phase contrast and fluorescence images (Zeiss Axiovert 200M microscopy) were captured for each field of view with a dwell time of 5 min (15). We used an open-source software, Molyso (20), to perform cell segmentation and lineage tracking, and further customized that software for proofreading (*SI Appendix*).

**Statistics of Promoter Leakiness with and without DNA Looping**

The regulatory regions of the lac operon consist of one main operator O₁, and of two auxiliary operators O₂ and O₃ (Fig. 1A). When a tetramer of LacI repressor binds two operators, for example, O₁–O₂ or O₁–O₃, it can form a stable DNA loop (Fig. 1B, Left). Previous population measurements suggest that, with only ~10 repressor tetramers (21), the lac operon can maintain a repression level ~100 times stronger than in the absence of DNA looping when there is only the single operator, O₁, present (1, 22). To compare the statistics of spontaneous leaky events with and without DNA looping, we investigated two *E. coli* strains: one that carries all three operators, denoted as the "Loops" strain and another one that only carries the main operator O₁, denoted as the "No-loop" strain (Fig. 1B and *SI Appendix*, Table S1). Without DNA looping, the promoter exhibits frequent transcriptional bursts (Fig. 2A, Left). By contrast, in the presence of DNA looping, the promoter leaks infrequently, and transcriptional bursts are separated by very

---

*Fig. 1.* Monitoring multigenerational leakiness of the endogenous lac promoter in single cells. (A and B) Wild-type lac operon maintains a low level of expression by the means of DNA looping. (A) Genetic organization of the lac system, where Oᵢ denote the operators i that are DNA-specific sequences where the repressor LacI binds. LacI is fully functional as a tetramer (36), and can bind to any of three possible operators. (B) Cartoon that illustrates how several operators can mediate the formation of DNA looping (Left) while one operator alone cannot (Right). (C) We developed an experimental platform to monitor promoter activity for low-expression systems across multiple cell divisions.
Next, we test one key prediction of the Vilar and coworkers model, that is, the promoter leakiness is insensitive to repressor concentration in the presence of DNA looping. Consequently, we investigate how sensitive repression mediated by DNA looping is to a high concentration of repressors versus the No-loop strain (Fig. 2A, Right). As for the distributions of OFF intervals, the OFF intervals still follow exponential distributions (Fig. 2C). Again, the burst size of 100x/Loops has a long tail, but not the 100x/No-loop strain. However, in the 100x/No-loop strain, promoter leakiness is very sensitive to the increase of the repressor concentration, with OFF intervals increasing to 178 [SE ± 8 min (or 2.6 cell cycles) and burst size to 210 [SE ± 9 a.u. By contrast, 100x/Loops is mostly insensitive to the increase in LacI repressor concentration and shows only slightly longer OFF intervals (292 [SE ± 15 min or 4.5 cell cycles) than those of the Loops strain. As for the distributions of the burst size, they are similar (223 [SE ± 13 a.u.), indicating that the typical burst size of the Loops strain has already been reduced to its lowest limit, which, we reasoned, may be associated with the synthesis of only one mRNA per pulse. The burst size of the 100x/No-loop strain also reached a similar limit in the presence of high repressor concentration. Our first observations are in agreement with the predictions of the Vilar and coworkers model. The observed insensitivity may directly stem from the saturation of the operator site in the presence of DNA looping with high “local” but low global concentration of the repressors. In principle, in the absence of DNA looping, this saturation may be achievable with high global concentration as well, but this condition might not be biologically favored, as it could become toxic to the cell.

However, that OFF intervals across many divisions in the Loops strains follow exponential distributions is unexpected, because it is the signature of a memoryless one-step process associated with a single rate. We hypothesize that this one-step process that controls the statistics of OFF intervals in the Loops strains is largely dominated by the unbinding of repressors from the operator $O_1$, regardless of whether the auxiliary operators $O_2$ and $O_3$ are bound or not.

To check this hypothesis, we genetically removed one additional operator ($O_2$ or $O_3$) from the Loops strain, and find that the OFF intervals of those strains that can form only one loop (either $O_1$–$O_2$ or $O_1$–$O_3$) follow exponential distributions as well (Fig. 3A). Furthermore, we observe that the $O_1$–$O_2$ one-loop strain exhibits qualitatively similar timescales of the OFF intervals (204 [SE ± 9 min) as the Loops strain that has the possibility to form multiple alternative loops (Fig. 3A). This result indicates that only one loop, $O_1$–$O_2$, dominates the statistics of the OFF intervals in the Loops strain and that the possibility of forming different loops only mildly affects the statistics of OFF intervals.

To evaluate the impact of cell division on the OFF intervals, we performed stochastic simulations of gene expression using the Vilar and coworkers model (12) (SI Appendix). We attempted to model cell division by periodically forcing the simultaneous unbinding of the repressors from the operators; however, this approach failed to reproduce long and exponentially distributed intervals as well as the robustness to repressor concentration in the Loops strain (SI Appendix, Fig. S8). By contrast, stochastic simulations in the absence of cell division can reproduce the experimentally observed statistics (Fig. 2C, Insets), indicating that
DNA looping confers strong robustness to the perturbations associated with cell division.

**A Theoretical Model without Cell Division.**

Most observations in our experiments can be qualitatively understood with a three-state model extended from the Vilar and coworkers model (11, 12): 1) state B, where the repressor LacI is bound to O1; 2) state E, where the operator O1 is empty and freed from RNA polymerase; and 3) state TS, where O1 is cleared from RNA polymerase and transcription starts (regardless of the states of O2 or O3). We further assume that only one transcript is produced in state TS and that the system returns to state E immediately. The transitions between states are described by

\[
B \xrightarrow{k_b} E \xrightarrow{k_e} TS,
\]

where \(k_b\) is the effective binding rate for the repressors to O1, \(k_e\) gives the unbinding rate for a repressor from O1 (0.10 min\(^{-1}\)), and \(k_e\) is the effective transcription rate (20 VenusNB \(\times\) min\(^{-1}\)). Without DNA looping, repressors follow a simple ON-OFF dynamics; thus \(k_b\) scales linearly with the repressor concentration \(n_0\) \((n_0 = 10\) molecules per cell in 1x LacI strain and \(n_0 = 1,000\) in 100x LacI strain). For a strain with DNA looping, when the system is in state E, one of its auxiliary operators is most likely bound to a repressor, given the free-energy difference between a bound and a free operator. For an E-\(\overline{B}\) transition, either a repressor from the rest of the cell, denoted as “global,” binds to O1 or the repressor that has already bound to an auxiliary operator binds rapidly to O1, denoted as “local.” The free-energy difference of the looping formation, \(e^{-\Delta G}\), is effectively equivalent to a very large “local” concentration \(n_L\) \((n_L = 0\) for No-loop, 1,080 for Loops) (11, 12). Considering both situations, we have \(k_b = k_{\text{int}}(n_R + n_L)\), where \(k_{\text{int}}\) is the binding rate for a single repressor (0.28 per molecule per min).

An OFF interval consists of one or multiple rounds of E-\(\overline{B}\) transitions before the system goes to state TS. The probability of \(l\) rounds to occur is \(P_l = e^{-\beta(1 - \alpha)}\), with \(\alpha = k_b/(k_b + k_e)\) as the probability of entering state B from E. Considering that \(k_b \ll k_e\), as implied by the physical parameters of the system, the timing at which \(l\) unbinding events happen is given by the composition of \(l\) exponential decays, which can be described by the Erlang distribution, \(w_l = k_e e^{-k_e(kt)}((1 - 1))/((1 - 1))\) resulting in a distribution of waiting times between transcriptional events (OFF intervals) \(w_0 = \sum_{l=0}^{\infty} P_l = e^{-\beta(1 - \alpha)}k_e k_b(1 - \alpha)\). Thus, the average duration of the OFF interval is \(\tau_{\text{OFF}} = \int_0^\infty tw_0\,dt = \frac{1}{k_b} (1 + (k_b/k_e))\). Theoretical calculations suggest that the ratio of \(\tau_{\text{OFF}}\) between the Loops strains with 1x and 100x repressor concentration is ~2, but that the ratio between the No-loop strains with 1x and 100x is ~14 (SI Appendix).

Consequently, the model qualitatively predicts the great sensitivity of the No-loop strains and the robustness of the Loops strains to repressor concentration.

On the other hand, the burst size of a pulse is proportional to the number of E-\(TS\) transitions before the system goes to state B, equivalently, the number of transcripts. Starting from state E, the probability of entering state TS is \(\beta = k_b/(k_b + k_t)\). The probability to produce \(r\) transcripts in a pulse is \(P_r = \beta^{r-1}(1 - \beta)\), a geometric distribution with an average number \(\langle r \rangle = \sum_r r P_r = \frac{1 - \beta}{1 - \beta} = \frac{k_t}{k_{\text{int}}(n_R + n_L)} + 1\). When \(n_R + n_L\) is large, \(\langle r \rangle \rightarrow 1\). Theoretical calculations suggest the No-loop strain is expected to have more than one transcript per pulse (estimated as approximately eight); by contrast, the other three conditions with DNA looping or a high concentration of repressors are predicted to have only about one transcript per burst. These predictions are in line with our experimental results.

Fig. 3. Robustness of repression from distinct configurations of DNA loops and cell cycle dependence of promoter activity. (A) One-loop vs. No-loop or Loops strains configurations. OFF intervals for one-loop strains with either (\(O_1-O_2\)) or (\(O_1-O_3\)) configuration combined with 1x and 100x LacI background all exhibit exponential distributions. We compare cumulative distributions (PDF x 2) for four one-loop strains with those of the Loops and No-loop strains. In each panel, the OFF intervals were normalized with the maximum value of the red dots. Number of OFF intervals included in the analyses is as follows: \(O_1-O_2\) one-loop (\(n = 1,239\)), \(O_1-O_2\) one-loop (\(n = 463\)), 100x/\(O_1-O_3\) one-loop (\(n = 207\)), and 100x/\(O_1-O_2\) one-loop (\(n = 583\)). Slopes before normalization are as follows: \(O_1-O_3\) one-loop \((-0.01162\) min\(^{-1}\)), \(O_1-O_2\) one-loop \((-0.00478\) min\(^{-1}\)), 100x/\(O_1-O_3\) one-loop \((-0.00534\) min\(^{-1}\)), and 100x/\(O_1-O_2\) one-loop \((-0.00355\) min\(^{-1}\)). (B) Promoter activity as a function of time span between two cell division events. Time span between two successive division events is normalized by the total duration of the division time. Promoter activity is normalized by the mean of each strain. Mean promoter activity is computed using \(N \geq 50\) cells. Error bars represent the SE. Inset gives unnormalized promoter activity for each strain.
observations showing that the burst size cannot be reduced further in the Loops strain, even when we drastically increase the repressor concentration.

**Correlation between Promoter Activity and Gene Dosage during Cell Growth**

As described by the Cooper–Helmstetter relation (23), gene expression depends on global factors such as gene dosage (24). As the cell grows, DNA replicates in such a way that the average copy number of chromosomes is maintained after division, but, between two division events, the gene dosage increases. In *E. coli*, it has been reported that the promoter activity of a gene with a high expression level is correlated with the phase of cell cycle (25), and this has been quantitatively measured in ref. 26. Under full induction (i.e., the removal of the repressors), expression is constitutive, and the promoter activity of the lac operon exhibits a flat region at the early phase of the cell cycle, and gradually increases to about twice its initial level (26). In our experiments, we monitor how the spontaneous leakiness of the Loops strain, they all show positive correlations with the lifetime of the DNA loop measured in vitro (30). Using Chen experiments, including memoryless distributions of the OFF intervals and extremely long OFF intervals with DNA looping, as well as its robustness to variations of repressor concentration. Given that bound repressors will unbind DNA during replication (10), and that a new copy of lac operon will be created, the structure of DNA looping may be affected, and the observation of multigenerational OFF intervals is unexpected, especially in the presence of very few repressors (~10 tetramers; see ref. 21). Remarkably, neither the variations due to gene dosage within cell cycles nor perturbations associated with division events limit the long OFF intervals and alter the associated simple exponential distribution.

In the No-loop strain, the ab initio search time for LacI to bind free O1 is >30 s (27, 28), but it only takes ~3 s for the RNA polymerase to start transcription (29). Thus, the No-loop strain is expected to be sensitive to the removal of the bound repressors onto DNA during growth, as observed in our experiments. By contrast, the timescale for long repression intervals observed in the Loops strain can be reconciled with the short lifetime of the DNA loop measured in vitro (30). Using Chen et al. (30) measurements and refs. 31 and 32, we estimated that the in vivo loop lifetime was of the order of 10^3 s, and that of the open loop was ~1 s. These values are similar to those used in our theoretical model that provides a simple hypothesis for the existence of OFF intervals longer than division cycles.

When O1 is unoccupied, we hypothesize that there is competition between RNA polymerase initiation and repressor rebinding events. Using the assumption of high “local” repressor concentration from the Vilar and coworkers model, we reason that the rebinding of the repressor onto the operator after either spontaneous unbinding or DNA replication is fast. For example, the typical timescale for the RNA polymerase initiation (~3 s) is about an order of magnitude slower than the timescale to reform a loop (~0.1 s, estimation from our model); consequently, transcription initiation is statistically allowed to happen only once every 30 spontaneous unbinding events. Along the same lines, we hypothesize that “local” repressor concentration is shared between the old and new DNA copies during or immediately after replication, which remains high relative to the DNA copy number that increases only by about a factor of 2. Additionally, the number of replication events that can potentially perturb repressor binding is small. Then, we shall expect the mean OFF intervals in the presence of DNA looping to be on the order of 30 times the loop lifetime (10^3 s) measured in vitro, that is, about seven cell cycles, in line with our experiments. While we assessed the statistical origin of the robustness of repression, it is still not clear how DNA looping can mechanistically conserve a high “local” concentration of repressors across divisions when DNA duplicates, and additional experiments, most likely at the single molecule level in live cells, may be able tackle this open question.

In summary, we report the robustness of DNA looping to intracellular perturbations across multiple cell cycles. While a small copy number of repressors is present in the cell, we find that repression with DNA looping is robust to variations of intracellular environment, such as repressor concentration, cell divisions, and detailed configurations of DNA loops. We speculate that similar robustness plays a crucial rule in other genetic regulatory systems beyond the lac operon.

**Materials and Methods**

All *E. coli* strains in this study were constructed using lambda red recombineering (33, 34), and the list of strains is available in SI Appendix, Table S1. For a typical experiment, cells were cultured in M9 media overnight, and then loaded into a mother machine microfluidics device for microscopy (Zeiss Axiovert 200M). Microscopy images were analyzed based on the software Molyso (20). Simulations of gene expression were performed following ref. 12. A full description of materials and methods can be found in SI Appendix.

**Data, Materials, and Software Availability.** Time series for the fluorescent signals have been deposited in Harvard Dataverse (35). The codes are available at https://github.com/changsysbio/ProbabilisticInferenceForPromoterActivity.

**ACKNOWLEDGMENTS.** We thank Guangwei Si, Junhao Zhu, Richard Losick, J. Mark Kim, Kritika Gupta, and Jimmy Martenson for fruitful discussions. We are indebted to Enrique Ballesa, who performed preliminary experiments for this study, and to Yuta Belmont for making a microfluidics wafer. This work was supported by NSF Grant 1615487 (to P.C.). J.M.G.V. acknowledges support from Ministerio de Ciencia e Innovación under Grant PGC2018-101282-B-I00 (MCI/AEI/FEDER, UE). L.S. is indebted to the Institute for Medical Engineering & Science at the Massachusetts Institute of Technology for hosting her for a sabbatical stay during the early stages of this work.

Author affiliations: *Department of Molecular and Cellular Biology, Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca 62210, Mexico; Department of Biomedical Engineering, University of California, Davis, CA 95616; Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02139; Biofisika Institutua (CSIC, UPV/EHU), University of the Basque Country UPV/EHU, Bilbao, 48080 Spain; and Ikerbasque, 48013 Bilbao, Spain