An Hfq-dependent post-transcriptional mechanism fine tunes RecB expression in *Escherichia coli*

Irina Kalita[1,2,3,*], Ira Alexandra Iosub[1,4,5], Lorna McLaren[1,2], Louise Goossens[1,2], Sander Granneman[1,4], and Meriem El Karoui[1,2,*]

[1]*Centre for Engineering Biology, University of Edinburgh, UK*
[2]*Institute of Cell Biology, University of Edinburgh, UK*
[3]*Max Planck Institute for Terrestrial Microbiology & Center for Synthetic Microbiology (SYNMIKRO), Marburg, Germany*
[4]*Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh, UK*
[5]*The Francis Crick Institute, London, UK*
[*]For correspondence: irykalita@gmail.com, meriem.elkaroui@ed.ac.uk

Abstract

All living organisms have developed strategies to respond to chromosomal damage and preserve genome integrity. One such response is the repair of DNA double-strand breaks (DSBs), one of the most toxic forms of DNA lesions. In *Escherichia coli*, DSBs are repaired via RecBCD-dependent homologous recombination. RecBCD is essential for accurate chromosome maintenance, but its over-expression can lead to reduced DNA repair ability. This apparent paradox suggests that RecBCD copy numbers may need to be tightly controlled within an optimal range. Using single-molecule fluorescence microscopy, we have established that RecB is present in very low abundance at mRNA and protein levels. RecB transcription shows high fluctuations, yet cell-to-cell protein variability remains remarkably low. Here, we show that the post-transcriptional regulator Hfq binds to *recB* mRNA and down-regulates RecB protein translation *in vivo*. Furthermore, specific disruption of the Hfq-binding site leads to more efficient translation of *recB* mRNAs. In addition, we observe a less effective reduction of RecB protein fluctuations in the absence of Hfq. This fine-tuning Hfq-mediated mechanism might have the underlying physiological function of maintaining RecB protein levels within an optimal range.

Key words: RecBCD enzyme, DNA repair, Hfq protein, post-transcriptional regulation, noise suppression.

Introduction

All living organisms rely on multiple molecular mechanisms to repair their chromosomes in order to preserve genome integrity. DNA double-strand breaks (DSBs), simultaneous lesions on both strands of the chromosome, are one of the most detrimental types of DNA damage because they can lead to loss of genetic information. If not repaired or repaired incorrectly, DSBs can trigger cell death, deleterious mutations and genomic rearrangements [1]. In addition to various exogenous sources such as ionizing radiation, mechanical stresses and DNA-damaging agents, DSBs can occur during the normal cell cycle as a results of endogenous metabolic reactions and replication events [2]. For instance, replication fork breakage has been shown to result in spontaneous DSBs in 18% of wild-type *Escherichia coli* (*E. coli*) cells in each generation [3], underscoring the crucial need to repair these breaks to ensure faithful DNA replication.

In *E. coli*, DSBs are repaired by homologous recombination using an intact copy of the chromosome as a template. Initially, the multi-subunit nuclease/helicase enzyme RecBCD detects a double-strand DNA (dsDNA) end and degrades both strands. Upon recognition of a short DNA sequence, known as a *chi* (χ) site (5′-GCTGGTGG-3′), the RecBCD complex stops cleaving the 3′-ended strand and initiates loading of the RecA protein onto the 3′-single-strand (ssDNA) overhang [4, 5]. This leads to the formation of a RecA-ssDNA nucleoprotein filament [6], which catalyses homology search and strand exchange [7].

Recent *in vitro* and *in vivo* experiments have highlighted the extraordinary processive DNA degradation activity of RecBCD upon recognition of a dsDNA end [8, 9, 10]. Indeed, RecBCD has been shown to process the chromosome for up to $\sim$100 *kb* at $\sim$1.6 *kb*/*s* in live bacteria [5]. Such a potent DNA degradation activity is controlled by *chi* recognition [11]. However, because this recognition is probabilistic [12], RecBCD can degrade large chromosome fragments before DNA repair is initiated. As a result, overproduction of RecBCD has been shown to sensitize wild-type cells to ultraviolet and gamma irradiations

[13]. This strongly suggests that an excess of RecBCD imposes a threat to cells upon DNA damage. In contrast, RecBCD expression is essential upon DSB induction [14, 15], hence too low levels might impair DNA repair. Thus, the observation that RecBCD can positively and negatively impact cell fitness upon DNA damage suggests that its activity needs to be tightly controlled.

Whilst the biochemical and *in vivo* activities of RecBCD have been extensively studied ([11, 16, 17, 18, 19], and [7, 20] for review), less is known about the regulation of its expression. RecBCD has been reported to be expressed at very low levels [21, 22] and is present in less than ten molecules per cell [23]. In *E. coli*, the RecBCD subunits are encoded in the same locus by two operons. One operon controls the expression of RecC whilst the other is polycistronic and encodes PtrA, RecB and RecD [24, 25] (Fig. 1A). PtrA is a periplasmic protease with no known function in DNA repair, and the positioning of its gene likely results from a horizontal gene transfer event that has interrupted the usual order *recC-recB-recD* [26]. As sometimes found in bacterial operons [27, 28], the ribosome binding site (RBS) of *recB* is located within the coding sequence of *ptrA*, suggesting a potential translational coupling mechanism between these genes. Previous attempts to explore RecBCD expression upon DNA-damage conditions did not reveal any transcriptional regulation [21, 29] and the *recBCD* genes were shown not to belong to the SOS-response regulon [30]. Thus, RecBCD transcription is considered to be constitutive.

Given that RecBCD is essential for cell survival upon DSB formation and toxic when overproduced, it is likely that its expression needs to be tightly regulated to avoid fluctuations leading to cellular death or deleterious genomic rearrangements. However, its very low abundance, often leading to high stochastic fluctuations, could result in large cell-to-cell variability [31, 32, 33]. Although these fluctuations might not play a significant role during normal cell growth, they may become crucial and determine cell fate upon stress conditions [34, 35, 36]. Thus, the low abundance of RecBCD enzyme molecules raises the following questions: (i) to what extent do RecBCD numbers fluctuate, (ii) are these fluctuations controlled, (iii) if this is the case, at which level (transcription, translation, mRNA or protein degradation) is the control operating.

The lack of evidence of RecBCD transcriptional regulation does not preclude that other mechanisms may be involved in regulating its expression. For example, RNA-binding proteins (RBPs) have been shown to fine tune gene expression both co- and post-transcriptionally during stress responses and adaptation to environmental changes [37, 38]. One of the most studied RBPs in *E. coli*, a Sm-like hexameric ring-shaped protein called Hfq, has been shown to regulate RNA stability and degradation, transcription elongation and translation initiation through various mechanisms (for review [39, 40, 41, 42]). In the most prevalent case, Hfq facilitates base-paring between small non-coding RNAs (sRNAs) and target mRNAs which often results in mRNA translation and/or stability modulation [43]. There is also increasing evidence that Hfq can control mRNA translation directly [44, 45]. This mechanism is typically based on altering the accessibility of the RBS of a target mRNA by direct binding near this region [46, 47, 48].

Here, we show that the expression of RecB is regulated post-transcriptionally in an Hfq-dependent manner in *E. coli*. By quantifying RecB mRNAs and proteins at single-molecule resolution, we found that constitutive transcription leads to low levels of *recB* mRNAs and a noisy distribution. In contrast, fluctuations in RecB protein levels are significantly lower than expected, given the very low number of proteins present in cells. We show that Hfq negatively regulates RecB translation and demonstrate the specificity of this post-transcriptional control *in vivo*. Furthermore, we provide evidence of the role of Hfq in reducing RecB protein fluctuations and propose that this could constitute a fine-tuning mechanism to control RecBCD molecules within an optimal range.

Results

recB mRNAs are present at very low levels and are short-lived

To quantitatively measure RecB expression, we first precisely quantified *recB* mRNAs using single-molecule fluorescent *in situ* hybridization (smFISH) [49]. We confirmed that our hybridization conditions are highly specific as *recB* mRNA quantification in a Δ*recB* strain resulted in a negligible error (∼0.007 molecules per cell, Fig. 1B). In wild-type cells, we detected a very low average level of *recB* transcripts: ∼0.62 molecules per cell. Notably, ∼70% of cells did not have any *recB* mRNAs (Fig. 1C). To confirm that this result was not because of low sensitivity of the smFISH protocol, we over-expressed *recB* mRNAs from an arabinose-inducible low-copy plasmid and detected the expected gradual amplification of the fluorescence signal with increased arabinose concentration (Fig. S2).

We next measured the degradation rate of *recB* transcripts in a time-series experiment. Initiation of transcription was inhibited with rifampicin and *recB* mRNA expression levels were quantified by sm-

FISH at subsequent time points. By fitting the mean mRNA value evolution over time to an exponential function, we estimated the *recB* degradation rate as $\gamma_m = 0.62\ min^{-1}$ ($CI_{95\%}$ [0.48, 0.75]) (Fig. 1D). This corresponds to a *recB* mRNA lifetime of 1.6 *min* and is consistent with the genome-wide studies in *E. coli* where transcript lifetimes were shown to be on average ∼2.5 *min* across all mRNA species [50, 51].

recB transcription is constitutive and weak

As noted previously, quantitative analysis of transcription at the single-cell level needs to take into account the variation of gene copy number during the cell cycle [52, 53, 54]. Therefore, we analysed RNA molecule abundance using cell size as a proxy for the cell cycle. Specifically, we binned all bacteria in groups based on cell area and averaged mRNA numbers in each of those size-sorted groups (Fig. 1E). A two-fold increase in mRNA abundance was observed within a narrow cell-size interval (∼3.0-3.5 μm^2), which is consistent with an expected doubling of the transcriptional output as a result of the replication of the *recB* gene. We further showed that *recB* transcription from sister chromosomes happens independently, in full agreement with the hypothesis of an unregulated gene (Fig. S3). Taken together, this evidence indicates that *recB* mRNA levels follow a gene-dosage trend, strongly suggesting that *recB* is constitutively expressed [52].

Next, to infer *recB* transcription rate, we used a simple stochastic model of transcription where mRNAs are produced in bursts of an average size b at a constant rate k_m and degraded exponentially at a rate γ_m (Suppl. Info.: Section 1). The expected probability distribution for mRNA numbers at steady-state is given by a negative binomial [55, 56, 57]. We used maximum likelihood estimation to infer the parameters of the negative binomial distribution from the experimental data (Fig. 1F, Methods). For the inference, we restricted our analysis to newborn cells (Fig. 1E) as this model does not take into account cell growth. We then inferred the burst size for the *recB* gene, $b \sim 0.95\ molec$, and the rate of transcription, $k_m = 0.21\ min^{-1}$ (Table 7, Methods). The latter is remarkably low and close to the lower-bound reported range of transcription rates in *E. coli*, 0.17-1.7 min^{-1} [58, 59, 60]. Therefore, the low abundance of *recB* mRNAs can be explained by infrequent transcription combined with fast mRNA decay.

Variations in mRNA molecule numbers, as a result of their low abundance and short lifetimes, can significantly contribute to gene expression noise [31, 61, 62]. To evaluate fluctuations of *recB* transcription, we calculated a commonly-used measure of noise, the squared coefficient of variation (CV_m^2), defined as the variance divided by the squared mean $\sigma_m^2/\langle m\rangle^2$. The average CV_m^2 in newborn cells was found to be CV_m^2 = 5.9, ∼1.9 times larger than the lower bound given by a Poisson process ($CV_{Poisson}^2 = 1/\langle m\rangle$ = 3.1). This suggests that *recB* mRNA production is noisy and further supports the absence of a regulation mechanism at the level of mRNA synthesis.

RecB proteins are present in low numbers and are long-lived

To determine whether the mRNA fluctuations we observed are transmitted to the protein level, we quantified RecB protein abundance with single-molecule accuracy in fixed individual cells using the Halo self-labelling tag [23] (Fig. 2A&B). As we previously reported, the translational RecB-HaloTag fusion does not affect the activity of the RecBCD enzyme [23]. We obtained RecB protein distributions for cells of all sizes and confirmed its low abundance: 3.9 ± 0.6 molecules per cell (Fig. 2C).

To test if RecB is actively degraded by proteases or mainly diluted as a result of cell growth and division, we estimated the RecB protein removal rate. This is usually obtained from either pulse-chase experiments using radio-labelled amino acids or detection of protein concentration decay after treatment with a translation inhibitor [63, 64, 65]. While the first approach is not amenable to single-cell assays, the latter may affect the measurements on long time scales because of the physiological impacts of stopping all protein production. Therefore, we used an alternative approach [66, 67] where we pulse-labelled RecB molecules with the HaloTag ligand, then removed the excess of the dye by extensive washing and measured the abundance of the labelled molecules over time. Assuming protein removal can be described by a first-order reaction, we fitted the mean protein counts over time to an exponential function on a subset of time points (Fig. 2D, shaded area). The decay rate was extracted from the fit and resulted in $\gamma_p = 0.015\ min^{-1}$ ($CI_{95\%}$ [0.011, 0.019]). Comparison to the population growth rate in these conditions (0.017 min^{-1}) suggests that RecB protein is stable and effectively removed only as a result of dilution and molecule partitioning between daughter cells.

RecB protein cell-to-cell fluctuations are lower than expected

We then extended the transcription model described above by considering protein synthesis and removal. This is known as the *two-stage* or *mRNA-protein* model of gene expression [57, 68, 69]. In this model, a protein is produced from an mRNA at a constant translation rate k_p while protein dilution

is implicitly modelled by a first-order chemical reaction with rate γ_p (Suppl. Info.: Section 1). As discussed earlier, we focused on newborn cells and estimated the rate of translation as 0.15 min^{-1} ($CI_{95\%}$ [0.12, 0.18], Table 7, Methods). This value fell below the lower bound of previously reported range of the translation rates in *E. coli* (0.25-1 min^{-1}) [58, 59]. Therefore, we concluded that *recB* mRNAs are translated at low levels compared to other proteins in the cell.

Next, we examined whether the two-stage model can reproduce the cell-to-cell variability in the protein levels observed in our experiments. To this end, we simulated single trajectories of protein abundance over time using the Stochastic Simulation Algorithm (SSA) [70] for the two-stage model with the estimated parameters (Table 7). Strikingly, the simulated distribution had larger fluctuations than the experimentally measured one (Fig. 2E). While both protein means were equal because of the first-moment-based parameter estimation we used (Methods), we found the standard deviations to be higher in simulation (2.2) than in experiments (1.5) (Fig. 2E). In other words, a two-stage protein production model predicts a higher level of fluctuations than what is observed experimentally.

This phenomenon can also be analysed using the squared coefficient of variation of the protein distribution CV_p^2, which is defined as $\sigma_p^2/\langle p\rangle^2$. For the two-stage gene expression model with bursty transcription, it has been established that the following expression gives the coefficient of variation (Suppl. Info.: Section 1) [57, 68, 69, 71]:

$$CV_p^2 = \frac{1}{\langle p\rangle} + \frac{1+b}{\langle m\rangle}\frac{\gamma_p}{\gamma_m+\gamma_p} \tag{1}$$

where $\langle p\rangle$ and $\langle m\rangle$ are the mean protein and mRNA counts (per cell), b is the average (mRNA) burst size, γ_m and γ_p are the mRNA degradation and protein removal rates, respectively. We found that the experimental coefficient of variation (CV^2_{Data} = 0.19±0.06) is approximately half of the analytically predicted one (CV^2_{Theory} = 0.45)(Fig. 2F). On the contrary and as expected, both the theoretically predicted CV^2_{Theory} and the one computed from the simulated data CV^2_{SSA} are in good agreement. This deviation from the theoretical prediction implies a potential regulatory mechanism of RecB expression that actively suppresses protein production variation to maintain protein levels within a certain range.

RecB translational efficiency is increased upon DSB induction

Next, we addressed whether there was any further evidence of this potential regulation once bacterial cells are exposed to stress. We hypothesized that RecB mRNA and/or protein levels may exhibit a concentration change upon DNA damage to stimulate DNA repair. To test this, we treated bacterial cells with a sub-lethal concentration of ciprofloxacin (4 ng/ml), an antibiotic that leads to DNA DSB formation [72]. After two hours of exposure to ciprofloxacin, we quantified proteins by Halo-labelling and mRNAs by smFISH (Fig. 3A). This induction duration is sufficient to detect changes in mRNA and protein levels (see Suppl. Info.: Section 2 for the mathematical model supporting this choice). Because *E. coli* cells filament upon activation of the DNA damage response (Fig. 3B&C), we measured single-cell distributions of mRNA and protein concentrations (calculated as molecule numbers normalized by cell area) rather than absolute numbers. Cell area is a reasonable proxy for cell volume as rod-shaped *E. coli* cells maintain their diameter while elongating in length upon ciprofloxacin treatment. RecB protein concentration did not change upon DSB induction (Fig. 3E). In fact, the average cellular RecB concentration was surprisingly conserved: 2.28 $mol/\mu m^2$ ($CI_{95\%}$: [2.21, 2.32]) in perturbed condition and 2.12 $mol/\mu m^2$ ($CI_{95\%}$: [2.01, 2.32]) in intact cells. Additionally, we analysed RecB concentration in size-sorted groups of cells (Fig. S5b). Regardless of moderate fluctuations in protein concentration over a cell cycle, both samples significantly overlap along the whole range of cell sizes.

In contrast, single-cell mRNA concentration quantification showed a sharp decrease in *recB* transcripts upon DNA damage (Fig. 3D). Specifically, mRNA concentrations upon DSB induction are ∼2.4-fold lower than in the unperturbed control (Fig. 3D: insert): 0.12 $mol/\mu m^2$ ($CI_{95\%}$: [0.06, 0.17]) upon DNA damage versus 0.28 $mol/\mu m^2$ ($CI_{95\%}$: [0.23, 0.32]) in the control. We further confirmed the decrease in all cell-size sorted groups (Fig. S5a). As previously reported, the *recB* gene is not part of the SOS regulon [30], hence its transcription is not necessarily expected to be upregulated upon DNA damage. However, downregulation of *recB* transcription is an unforeseen observation and may be a result of lower concentration of genomic DNA and/or shortage of resources (such as RNA polymerases) upon SOS response activation.

We then computed the average number of proteins produced per mRNA in both conditions. The ratio between the average protein and mRNA concentrations in damaged cells is 19.7 ($CI_{95\%}$: [17.5,

24.2]) whereas it is 7.9 proteins/mRNA ($CI_{95\%}$: [7.5, 8.3]) in intact cells (Fig. 3F). Thus, a $\sim$ 2.5 increase in translational efficiency was detected upon DSB induction (p-value < 0.001, two sample t-test). Taken together, our results show decoupling between mRNA and protein production upon DNA damage, suggesting the existence of a post-transcriptional regulatory mechanism that controls RecB protein level.

Hfq controls the translation of *recB* mRNA

To identify the post-transcriptional mechanism controlling *recB* expression, we focused on Hfq because (i) it has been shown to regulate a vast number of mRNAs [73, 74, 75] and (ii) its activity has been directly linked to the DNA damage and SOS responses [45, 76, 77].

To assess whether Hfq is involved in post-transcriptional regulation of RecBCD, we analysed the available transcriptome-wide dataset, wherein Hfq binding sites were identified in *E. coli* using CLASH (cross-linking and analysis of cDNA) [78]. In a CLASH experiment, Hfq-RNA complexes are cross-linked *in vivo* and purified under highly stringent conditions, ensuring specific recovery of Hfq RNA targets [78]. Analysis of the Hfq binding profile across *recBCD* transcripts showed the interaction of Hfq with the *recBCD* mRNAs (Fig. 4A). Hfq binding is mainly localized at two sites: one in close proximity to the RBS of the *ptrA* gene, and the second further downstream in the coding sequence of *recB* gene. Interestingly, multiple trinucleotide Hfq-binding motifs, A-R(A/G)-N(any nucleotide) [73, 79] were found $\sim$8 nucleotides upstream of the RBS of the *ptrA* sequence (Fig. S6), suggesting that Hfq may control translation initiation and/or decay of the operon mRNA.

To test this hypothesis, we quantified RecB protein and mRNA levels in an *E. coli* strain lacking Hfq. Δ*hfq* mutant has been reported to have an increased average cell size and various growth defects [80, 81, 82]. Indeed, we confirmed the increased cell size of our Δ*hfq* mutant compared to wild-type cells (Fig. S7a&b). In our conditions, no difference between the growth rates of WT and the Δ*hfq* mutant was observed in the exponential phase of growth (Fig. S7d).

Using Halo-labelling technique, we performed RecB protein quantification in Δ*hfq*. Remarkably, we observed an increase of RecB protein concentration in the Δ*hfq* cells (Fig. 4B&C). The Δ*hfq* mutant showed more fluorescent foci relative to the wild-type strain (Fig. 4B). Even after normalization by cell area to take into account the larger cell size in Δ*hfq* cells, RecB concentration distribution was shifted towards higher values (Fig. 4C). Quantification across five replicated experiments showed a statistically significant 30% increase in RecB protein concentration in the Δ*hfq* strain (Fig. 4C: insert). The concentration was 2.67 $mol/\mu m^2$ ($CI_{95\%}$: [2.54, 2.80]) in Δ*hfq* compared to 2.17 $mol/\mu m^2$ ($CI_{95\%}$: [2.01, 2.32]) in the control.

We further quantified *recB* mRNA molecules in the Δ*hfq* strain using smFISH and analysed mRNA concentration distributions (Fig. 4D&E). Although the average *recB* mRNA concentration is slightly decreased in the Δ*hfq*: 0.21 $mol/\mu m^2$ ($CI_{95\%}$: [0.08, 0.30]) *versus* 0.28 $mol/\mu m^2$ ($CI_{95\%}$: [0.23, 0.32]) in the wild-type cells, this difference is insignificant (p-value = 0.11, two sample t-test). These results were confirmed by RT-qPCR measurements, which did not show any significant changes for *recB* mRNA steady-state levels (Fig. S7c). Thus, we conclude that Hfq binding to the *recBCD* mRNA does not substantially alter the stability of the transcript.

The observed increase of RecB protein concentration without a significant change in *recB* mRNA steady-state levels is reflected in a rise in RecB translational efficiency (Fig. 4F). The average level of proteins produced per mRNA in the Δ*hfq* mutant is 14.0 ($CI_{95\%}$: [13.0, 18.1]), which is two-fold higher than in the wild-type cells (p-value < 0.05, two-sample t-test). Altogether, this demonstrates that Hfq alters RecB translation *in vivo*.

To test if Hfq regulation contributes to the noise suppression detected in the wild-type cells (Fig. 2F), we looked at the level of fluctuations of RecB copy numbers in Δ*hfq*. To this end, we computed the difference between the analytically predicted noise CV^2_{Theory} and the noise observed experimentally CV^2_{Exp}. Since in the Δ*hfq* mutant *recB* transcription is not affected, the theoretical noise was calculated using Eq. (1) assuming the same mRNA contribution (the second term of the equation). As seen in Fig. 4G, the wild-type strain shows effective noise suppression as the theoretically predicted noise is larger than the one detected experimentally. Interestingly, this suppression is less effective in the Δ*hfq* mutant. In other words, in the absence of Hfq, the mismatch between the model and the experimental data is decreased in comparison to the native conditions of the wild-type cells (p-value = 0.0088, two sample t-test). This provides evidence that, in addition to the average RecB abundance, Hfq affects RecB protein fluctuations.

RecB protein number distribution is recovered by Hfq complementation

To further confirm that Hfq regulates RecB expression, we complemented the Δ*hfq* mutant with a multi-copy plasmid that expresses Hfq, pQE-Hfq [83]. After complementation, the RecB protein concentration distribution was shifted towards lower values (Fig. 5B) whilst no change of RecB protein level was detected in the Δ*hfq* cells carrying the backbone plasmid, pQE80L (Fig. S8a). Although in our experimental conditions, Hfq is expressed at lower levels than in wild-type cells (transcription is $\sim$ a third of the wild-type level, Fig. S8b), RecB protein concentration was nearly fully recovered to the wild-type conditions: 2.57 $mol/\mu m^2$ ($CI_{95\%}$: [2.50, 2.65]). We avoided full induction of Hfq from the plasmid as it can interfere with its self-regulation and lead to cell filamentation. We did not detect significant changes in the *recB* mRNA levels when expressing Hfq from the plasmid (Fig. S8c), confirming that Hfq complementation impacts *recB* mRNA translation but not its stability.

Sequestering Hfq leads to an increase in RecB production

To further ascertain the role of Hfq in controlling RecB expression, we characterized RecB expression while sequestering Hfq proteins away from its *recB* mRNA targets. Hfq is known to bind $\sim$100 sRNAs [73, 75], some of which can reach high expression levels and/or have a higher affinity for Hfq and, thus, compete for Hfq binding [84, 85]. We reasoned that because *recB* mRNA is present in very low abundance, it could be outcompeted for Hfq binding by over-expression of a small RNA that efficiently binds Hfq, such as ChiX [85, 86]. We constructed a multi-copy plasmid pZA21-ChiX to over-express ChiX and obtained a $\sim$30-fold increase in ChiX expression relative to its native expression in wild-type cells (Fig. S9). Given its already high abundance in wild-type cells, ChiX overproduction has been demonstrated to effectively sequester a large number of Hfq molecules [46, 48, 87].

Wild-type cells carrying the pZA21-ChiX plasmid (or the backbone plasmid) were grown to mid-exponential phase, and RecB mRNA and protein levels were quantified with RT-qPCR and Halo-labelling, respectively. No significant changes were detected at the *recB* transcript levels when ChiX was overproduced (Fig. S9). In contrast, a significant increase in RecB protein production was detected upon ChiX over-expression (Fig. 5C). Indeed, RecB concentration in the cells where ChiX was over-expressed almost fully overlapped with the distributions in the Δ*hfq* mutant. This suggests that when fewer Hfq molecules are available for RecB regulation because ChiX is overproduced, RecB is translated at a higher rate. A similar experiment with a plasmid that carried a different sRNA, CyaR which is not expected to titrate Hfq, did not change RecB protein concentration. This indicates that ChiX over-expression specifically impairs RecB post-transcriptional regulation (Fig. S9).

To exclude the possibility of a direct ChiX involvement in RecB regulation, we characterized RecB expression in Δ*chiX*. No significant change relative to wild-type cells was detected in RecB mRNA or protein levels (Fig. S10). Thus, we concluded that ChiX does not affect RecB translation. Instead, titration of Hfq proteins from a shared pool by ChiX sRNAs leads to the increased availability of *ptrA-recB* RBS for ribosomes and, thus, results in more efficient RecB translation.

Deletion of an Hfq-binding site results in increased RecB translational efficiency

To further investigate the mechanism of RecB translation, we tested the specificity of Hfq interaction with the mRNA of *recB* without affecting other functions of Hfq. Based on the localization of the Hfq-binding uncovered in the CLASH data (Fig. 4A), we deleted a 36 nucleotide region (5′-TTAACGTGTTGAATCTGGACAGAAAATTAAGTTGAT-3′) located in 5′UTR of the operon, in front of *ptrA* gene. This region showed the highest enrichment of Hfq binding (Fig. S6). It is closely located to the RBS site and 76 nucleotides downstream of the promoter sequence of *ptrA-recB*. To get a complete picture of RecB expression in the mutant strain carrying this deletion (which we further refer to as the *recB*-5′UTR* mutant), we performed single-cell quantification of both mRNA and protein levels of RecB and quantified its translational efficiency.

Firstly, we quantified *recB* mRNAs molecules in the *recB*-5′UTR* strain using smFISH (Fig. 5D). We detected a significant decrease in the concentration of mRNAs in the strain with the modified sequence: 0.14 $mol/\mu m^2$ ($CI_{95\%}$: [0.07, 0.17]) compared to 0.28 $mol/\mu m^2$ ($CI_{95\%}$: [0.23, 0.32]) in the wild type. This is not an unexpected outcome as the 5′-untranslated region of a transcript can control its degradation and stability [88, 89]. Despite the mRNA steady-state level decreasing two-fold in the modified *recB*-5′UTR* strain, only a slight decrease in RecB protein concentration was detected in comparison with the wild type (Fig. 5E): 1.96 $mol/\mu m^2$ ($CI_{95\%}$: [1.56, 2.36]) in *recB*-5′UTR* and 2.17 $mol/\mu m^2$ ($CI_{95\%}$: [2.01, 2.32]) in the wild type.

We calculated the translational efficiency of *recB* transcripts based on our single-cell mRNA and protein quantification. The resulting translational efficiency in the *recB*-5′UTR* strain was found to be

significantly higher than in the wild-type cells: 16.3 ($CI_{95\%}$: [14.3, 19.7]) in *recB*-5′UTR*, whereas it is 7.9 ($CI_{95\%}$: [7.5, 8.3]) proteins/mRNA in the wild type (Fig. 5F). Remarkably, the RecB translational efficiency in the *recB*-5′UTR* strain is comparable with the values observed in the Δ*hfq* mutant (a two-sample t-test confirmed that the difference is not significant). Thus, when the main Hfq binding site was deleted, the lower abundance of the mRNA was compensated by more efficient translation. This demonstrates that the translation of *recB* mRNA is limited by a specific interaction with Hfq.

Discussion

Despite its essential role in DSB repair, over-expression of RecBCD decreases cell viability upon DNA damage induction [13]. Such a dual impact on cell fitness suggests the existence of an underlying regulation that controls RecBCD expression within an optimal range. Using state-of-the-art single-molecule mRNA and protein quantification, we (i) characterize RecB expression at transcriptional and translational levels in normal and DNA-damaging conditions and (ii) describe a novel post-transcriptional mechanism of RecB expression mediated by the global regulator Hfq *in vivo*. Utilizing stochastic modelling, we provide evidence that Hfq contributes to suppressing fluctuations of RecB molecules. To our knowledge, this is the first experimental evidence of an RNA-binding protein involved in noise suppression in bacterial cells under native conditions.

While our observations are consistent with recent studies demonstrating direct gene regulation by Hfq in a sRNA-independent manner [45, 46, 47, 48, 90], we cannot entirely exclude a small RNA being involved. Using the available Hfq CLASH and RIL-seq datasets [78, 91], CyaR and ChiX sRNAs were identified as potentially interacting with the *ptrA-recB-recD* mRNA. However, we did not find any significant changes in RecB expression when CyaR was overexpressed (Fig. S9) and we showed that ChiX was involved indirectly by sequestering Hfq (Fig. 5 & Fig. S10). Strictly speaking, we cannot totally rule out the possibility of another small RNA candidate, but the very low abundance of the *recB* transcripts is likely to make it particularly difficult to detect RNA-RNA interactions in bulk experiments, making it necessary to develop a specific approach with single-molecule resolution.

Hfq-mediated control of RecB expression may be just a hint of a larger bacterial post-transcriptional stress-response programme to DNA damage. Indeed, post-transcriptional regulation has already been shown to impact DNA repair and genome maintenance pathways [76, 77, 92, 93]. For instance, the DNA mismatch repair protein, MutS, was shown to be regulated by a sRNA-independent Hfq binding mechanism and its translation repression was linked to increased mutagenesis in stationary phase [45]. In addition, a modest decrease (∼30%) in Hfq protein abundance has been seen in a proteomic study in *E. coli* upon DSB induction with ciprofloxacin [94]. It is conceivable that even modest changes in Hfq availability could result in significant changes in gene expression, and this could explain the increased translational efficiency of RecB upon DBS induction (Fig. 3F). Therefore, Hfq-mediated regulation might play a larger role in DNA repair pathways than previously considered. As many metabolic pathways are regulated through Hfq [95, 96, 97], the control of RecB (and other DNA repair proteins) expression may even provide a way of coordinating DNA repair pathways with other cellular response processes upon stress.

Mechanisms similar to the one described in our study might underlay the expression networks of other DNA repair proteins, many of which are known to be present in small quantities. In this regard, it is worth emphasizing the valuable methodological insights which can be taken from our study. By utilizing single-molecule mRNA and protein quantification techniques, we show the importance of detailed investigation of both mRNA and protein levels when establishing novel regulation mechanisms of gene expression. Indeed, measuring the protein level only would not be sufficient to draw the correct interpretation, for example, in the case of *recB*-5′UTR* mutants (Fig. 5). Moreover, single-molecule measurements combined with stochastic modelling allow access to cell-to-cell variability, which can be a powerful and, possibly, the only approach to investigate the regulation mechanisms of gene expression for low abundant molecules.

In agreement with a previous report [13], we confirm that RecBCD overproduction led to less efficient repair when DSBs were induced using ciprofloxacin (Fig. S11). We note that another study did not detect any toxicity of RecBCD overproduction upon DNA damage [98]. However, this is likely because the level of UV irradiation was too low to detect a significant reduction in cell viability. We only observed the effect at a relatively high level of ciprofloxacin exposure (>10 ng/ml with the minimum inhibitory concentration at 16 ng/ml in our conditions).

In addition to the regulation of the average number of proteins per cell, fluctuation control might be

crucial for a quasi-essential enzyme with very low abundance, such as RecBCD. Our analysis of RecB fluctuations suggests that Hfq is playing a role in protein copy number noise reduction (Fig. 4G). Post-transcriptional regulation has been proposed theoretically as an effective strategy for reducing protein noise [62, 99, 100, 101, 102]. Conceptually, this is achieved via buffering stochastic fluctuations of mRNA transcription and enabling more rapid response. Based on these principles, suppression of protein fluctuations via post-transcriptional control has been demonstrated in synthetically engineered systems in bacteria, yeast and mammalian cells [103, 104, 105]. However, do cells employ these strategies under native conditions? Suppressed fluctuations of protein expression were experimentally detected for hundreds of microRNA-regulated genes in mammalian cells [106]. Remarkably, those genes were expressed at a very low level. This is precisely the case for the regulation of *recB* translation by Hfq, where a very abundant post-transcriptional regulator can effectively buffer the fluctuations of the low-abundant RecB enzyme. Taken together, these previous observations and our results could hint at a universal functional 'niche' for post-transcriptional regulators across a wide range of organisms.

Methods

Strains and plasmids

All strains and plasmids, used in the study, are listed in Table 1 and Table 2, respectively. The *E. coli* MG1655 and its derivatives were used for all conducted experiments except the one with arabinose-inducible expression where we used *E. coli* BW27783 background strain. The strains with RecB-HaloTag fusion were used in Halo-labelling experiments while the corresponding parental strains were used in smFISH experiments.

MEK1329 and MEK1938 were built using the plasmid-mediated gene replacement method [107] with pTOF24 derivative plasmids, pTOFΔ*recB* [108] and pTOF*recB*-5'UTR (this work), respectively. The deletion of *recB* gene and a 36 bp region (5'-TTAACGTGTTGAATCTGGACAGAAAATTAAGTTGAT-3') were verified with Sanger sequencing.

MEK1902 and MEK1457 were constructed with lambda red cloning technique using hfq_H1_P1 and hfq_H2_P2 primers (Table 3). MEK1888 and MEK1449 were constructed with lambda red using ChiX_H1_P1 and ChiX_H2_P2 primers (Table 3). *hfq* and *chiX* deletions were verified by PCR.

pIK02 was constructed to induce *recB* expression. The backbone was amplified with the primers oIK03/oIK04 from pBAD33 vector while the *recB* gene was amplified from *E. coli* chromosome using oIK10/oIK11 oligos (Table 3). The PCR products were ligated using Gibson. Plasmid construction was verified by PCR and sequencing.

pZA21-ChiX plasmid was constructed to allow overproduction of a small RNA ChiX. The plasmid, derived from the pZA21MCS (Expressys), was generated according to the protocol described here [78] using chiX_ZA21 and pZA21MCS_5P_rev primers (Table 3).

To construct pTOF*recB*-5'UTR, the pTOF24 plasmid was digested with XhoI and SalI and ligated using Gibson assembly with the gBlock *recB*-5'UTR (Table 4). The deletion of a 36 bp region was confirmed by sequencing.

Growth media and conditions

In microscopy and population experiments, cell cultures were grown in M9-based medium supplemented with 0.2% glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$ and either 10% Luria Broth (Figs 1&2) or MEM amino acids (from Gibco) (Figs 3–5). In the *recB* induction smFISH experiments, 0.2% glucose was replaced with 0.2% glycerol, and arabinose ranging from 10^{-5}% to 1% was added to the medium. For mRNA lifetime measurements with smFISH, rifampicin (Sigma-Aldrich) was added to cell cultures to a final concentration of 500 $\mu g/ml$. In the DSB induction experiments, a sub-lethal concentration of ciprofloxacin 4 ng/ml was used. Ampicillin (50 $\mu g/ml$), kanamycin (50 $\mu g/ml$) or chloramphenicol (30 $\mu g/ml$) were added to cultures where appropriate.

Cell cultures were grown in medium overnight (14–16 hours) at 37°C. The overnight cultures were diluted (1:300) and grown at 37°C until the mid-exponential phase (optical density OD_{600} = 0.2–0.3). Unless otherwise stated, cells were further treated according to the smFISH or Halo-labelling procedures described below.

In the *recB* induction smFISH experiments, arabinose was added after one hour of overday growth, the cultures were grown for one more hour, and then samples were collected. For the smFISH experiments with rifampicin, the antibiotic was added once the cultures had reached OD_{600} = 0.2. Cells were harvested at 1 min intervals and fixed in 16% formaldehyde. The following steps of smFISH procedure were followed as described below.

In the DSB induction experiments, ciprofloxacin was added at $OD_{600} \sim 0.1$ and cell cultures continued growing for one hour. For mRNA quantification with smFISH, the cultures were grown for one more hour before being harvested and fixed in the final concentration of 3.2% formaldehyde. For protein quantification, the Halo-labelling protocol was followed as described below with ciprofloxacin being kept in the medium during labelling and washing steps.

Single-molecule RNA FISH (smFISH)

Single-molecule fluorescent *in situ* hybridization (smFISH) experiments were carried out according to the established protocol [49]. Bacterial cells were grown as described above, fixed in formaldehyde solution, permeabilised in ethanol and hybridized with a *recB* specific set of TAMRA-labelled RNA FISH probes (LGC Biosearch Technologies). Probe sequences are listed in Table 6. Unbound RNA FISH probes were removed with multiple washes, and then samples were visualized as described below.

Halo-labelling

For single-molecule RecB protein quantification, we followed the Halo-labelling protocol described previously [23]. *E. coli* strains with RecB-HaloTag fusion, grown as described above, were labelled with JaneliaFluor549 dye (purchased from Promega) for one hour, washed with aspiration pump (4-5 times), fixed in 2.5% formaldehyde and mounted onto agar pads before imaging. In each experiment, a parental strain which does not have the HaloTag fusion ('no HaloTag') was subjected to the protocol in parallel with the primary samples as a control.

Microscopy set-up and conditions

Image acquisition was performed using an inverted fluorescence microscope (Nikon Ti-E) equipped with an EMCCD Camera (iXion Ultra 897, Andor), a Spectra X Line engine (Lumencor), dichronic mirror T590LPXR, 100X oil-immersion Plan Apo objective (NA 1.45, Nikon) and 1.5X magnification lens (Nikon). A TRITC (ET545/30nm, Nikon) filter was used for imaging Halo-labelling experiments while smFISH data were acquired with an mCherry (ET572/35nm, Nikon) filter.

Once a protocol (smFISH or Halo-labelling) had been carried out, fixed cells were resuspended in 1X PBS and mounted onto a 2% agarose pad. The snapshots were acquired in bright-field and fluorescence channels. In all experiments, an electron-multiplying (EM) gain of 4 and an exposure time of 30 ms were used for bright-field imaging. In the fluorescence channel, smFISH data were acquired with a 2 s exposure time and a gain of 4, while RecB proteins were visualized with the same exposure time but EM gain 200. For each XY position on an agarose pad, a Z-stack of 6 images centred around the focal plane (total range of 1 μm with a step of 0.2 μm) was obtained in both channels. A set of multiple XY positions were acquired for each slide to visualize $\sim$1000 cells per sample.

Cell segmentation

Microscopy bright-field images were used to identify positions of bacterial cells with an automated MATLAB-based pipeline [109]. Briefly, the segmentation algorithm is based on detection of cell edges by passing an image through a low-pass filter. In comparison to the original version, we applied this analysis to defocused bright-field images (instead of fluorescence signal from a constitutive reporter [109]) and tuned segmentation parameters for our conditions. The segmentation outputs were manually corrected by discarding misidentified cells in the accompanied graphical interface. The corrected segmentation results were saved as MATLAB matrices (cell masks) containing mapping information between pixels and cell ID numbers. The cell masks were used as an input for the spot-finding software described below.

Spot detection

The Spätzcell package (MATLAB) was utilized to detect foci in fluorescent smFISH images [49]. Principally, the analysis consists of the following steps: (i) identification of local maxima above a chosen intensity threshold, (ii) matching the maxima across frames in a Z-stack, (iii) performing 2D-Gaussian fitting of the detected maxima. Based on the peak height and spot intensity, computed from the fitting output, the specific signal was separated from false positive spots (Fig. S1a). The integrated spot intensity histograms were fitted to the sum of two Gaussian distributions, which correspond to one and two mRNA molecules per focus. An intensity equivalent corresponding to the integrated intensity of FISH probes in average bound to one mRNA was computed as a result of multiple-Gaussian fitting procedure (Fig. S1b), and all identified spots were normalized by the one-mRNA equivalent.

To detect fluorescent single-molecule foci in the Halo-labelling experiments, the first step of Spätzcell analysis was modified (similar to the modification for fluorescent protein detection implemented here [52]). For the Halo-labelling single-molecule data, we tested two modifications: (i) removing a Gaussian filter (that was used to smooth raw signal in the original version) or (ii) calculating the Laplacian of a Gaussian-filtered image. Both versions gave similar and consistent quantification results. Peak height intensity profiles allowed the separation of specific signal from false positive spots (Fig. S4). As a result of low-abundance of the protein of interest and single-molecule labelling, the analysis of integrated intensity was skipped for Halo-labelling images. All other steps of the Spätzcell software remained unchanged.

Total (mRNA or protein) molecule numbers per cell were obtained by matching spot positions to cell segmentation masks and quantifying the number of (mRNA or protein) spots within each cell. Molecule concentration was calculated for each cell as total number of (mRNA or protein) molecules (per cell) divided by the area of the cell ($1/\mu m^2$).

Bacterial RNA extraction and RT-qPCR

Bacterial cells were grown as described earlier, harvested in the equivalent of $volume \times OD_{600}$ = 2.5 ml and flash-frozen in liquid nitrogen. In the Halo-labelling experiments, samples for RNA extraction were grown at 37°C and collected before the fixation step. Total RNA extraction was performed with the guanidium thiocyanate phenol protocol [110].

Primers for real-time qPCR experiments are listed in Table 5. RT-qPCR quantification was carried out on 20 ng of RNA with the Luna One-Step RT-qPCR Kit (NEB). All qPCRs were performed in technical triplicates. The qPCR reactions were carried out on a LightCycler®96 (Roche). The *rrfD* gene (5S rRNA) was used as a reference gene to compute a fold-change relative to a control sample with the $2^{-\Delta\Delta Ct}$ method. Outlier samples with the standard deviation $std(Ct) > 0.3$ across technical replicates were removed from the analysis.

Population growth and viability assays

Bacterial cells, seeded by inoculating from fresh colonies, were grown in an appropriate medium at 37°C overnight. The cultures were diluted (1:1000) and grown in a shaking incubator at 37°C. Optical density measurements were taken overday at OD_{600} with the interval of 15 min. The exponential phase of growth, OD_{600} = [0.08, 0.4], was used for the fitting procedure. The fitting was performed with a linear regression model in MATLAB.

In the viability assays, cells were grown overnight in Luria Broth at 37°C in a shaking incubator. The optical density of the overnight cultures was normalized to OD_{600} = 1.0. 10-fold serial dilutions were then performed in 96-well plates for the range of OD_{600} from 1.0 to 10^{-7}. 100 μl of cells of a specific dilution were plated onto LB agar plates containing ciprofloxacin and ampicillin at a concentration of 0–16 ng/ml and 100 $\mu g/ml$, respectively. For each condition, two dilutions were optimally chosen to have 30-300 colonies per plate. After 24 hours of incubation at 37°C, the colonies on each plate were counted and colony forming units (CFU) per ml were calculated using the following formula:

$$\frac{\text{CFU}}{\text{ml}} = \frac{\text{Number of colonies per plate * dilution}}{100\ \mu l} \tag{2}$$

The survival factor (SF) was calculated by normalizing against the CFU/ml in the absence of ciprofloxacin.

Data analysis of CRAC data

The CLASH data (NCBI Gene Expression Omnibus (GEO), accession number GSE123050) were analysed in the study [78] with the entire pipeline available here [111, 112]. The SGR file, generated with `CRAC_pipeline_PE.py`, was used to plot the genome browser track in Fig. 4A.

Fitting and parameter estimation

The smFISH data from Fig. 1C was conditioned on cell size (cell area < 3.0 μm^2, *newborns*) and the updated distribution was fitted by a negative binomial distribution. In mathematical terms, a negative binomial $NB(r, p)$ is defined by two parameters: r, the number of successful outcomes, and p, the probability of success. Mean and variance of $X \sim NB(r, p)$ are given as:

$$\langle X \rangle = \frac{pr}{(1-p)} \qquad \sigma_X^2 = \frac{pr}{(1-p)^2} \tag{3}$$

The probability density function of the fitted distribution was found with the maximum likelihood estimation method in MATLAB. The fitted parameters, r and p, are given as $r = 0.335$ and $p = 0.513$.

The average mRNA burst size b and transcription rate k_m were calculated using the fitted parameters, r and p, and the *recB* mRNA degradation rate ($\gamma_m = 0.615\ min^{-1}$, measured in an independent time-course experiment: Fig. 1D) according to the following analytical expressions (Eq. (A4), Eq. (A5) & Eq. (3)):

$$b = \frac{1-p}{p} = 0.95\, molec \tag{4}$$

$$k_m = \gamma_m r = 0.21\, min^{-1} \tag{5}$$

The RecB protein distribution from Fig. 2C was conditioned on cell size (cell area < 3.0 μm^2, *newborns*). Then, the average protein count number in *newborns* (∼3.2 molecules per cell) was matched to the

theoretical expression for the protein mean Eq. (A4). The rate of translation k_p was calculated as follows (using protein removal rate, $\gamma_p = 0.015\ min^{-1}$, and mean mRNA number in *newborns*, $\langle m \rangle = 0.323$):

$$k_p = \frac{\gamma_p \langle p \rangle}{\langle m \rangle} = 0.15\, min^{-1} \tag{6}$$

All parameters of the model are listed in Table 7.

Statistical analysis

Unless otherwise stated, histograms and RT-qPCR results represent the average across at least three replicated experiments and error bars reflect the maximum difference between the repeats.

The Kullback–Leibler divergence (D_{KL}) between simulated and experimental distributions was calculated with the MATLAB function [113]. Comparison between the average protein or mRNA concentrations among different conditions were performed with a two-sample *t*-test and *P* values were calculated in the MATLAB built-in function `ttest2`.

Average translational efficiency was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m (Fig. 3F & Fig. 4F). The error bars are standard deviations across replicated experiments.

Stochastic simulations

Single trajectories of mRNA and protein abundance over time were generated for the reaction scheme Eq. (A1) with the Gillespie's algorithm, executed in MATLAB environment [114]. Parameters, used in simulations, are listed in Table 7. 10,000 simulations were performed for a given condition and probabilistic steady-state distributions were recovered for population snapshot data.

Data and code availability

All data that support this study including microscopy data, RT-qPCRs, growth rates, viability assays, and source codes for image analysis are available on Zenodo (10.5281/zenodo.8431113).The modified versions of the image analysis used for spot detection in Halo-labelling experiments and bright-field segmentation pipeline are also available on GitLab https://gitlab.com/ikalita/spotdetection_spatzcells and https://gitlab.com/ikalita/cellsegmentation_mek, respectively. The CRAC/CLASH dataset (NCBI Gene Expression Omnibus ID GSE123050) [78], used in this study, is available from GSE123050. The pipeline for the analysis of the CRAC/CLASH data is available from https://git.ecdf.ed.ac.uk/sgrannem and https://pypi.org/user/g_ronimo/.

Article and author information

Author contributions

Irina Kalita: Conceptualization, Methodology, Software, Investigation (experiments), Formal analysis (computational analysis & modelling), Writing – Original Draft Preparation, Writing – Review & Editing, Visualization;

Ira Alexandra Iosub: Conceptualization, Methodology, Software, Investigation (experiments), Formal analysis (computational analysis), Writing – Review & Editing, Visualization;

Lorna McLaren: Investigation (experiments), Writing – Review & Editing;

Louise Goossens: Investigation (experiments), Writing – Review & Editing;

Sander Granneman: Conceptualization, Methodology, Software, Formal analysis (computational analysis), Writing – Review & Editing, Supervision, Funding Acquisition;

Meriem El Karoui: Conceptualization, Methodology, Formal analysis (modelling), Writing – Original Draft Preparation, Writing – Review & Editing, Supervision, Funding Acquisition.

Acknowledgements

We thank David Leach, Gerald Smith, Teppei Morita, and Hiroji Aiba for providing strains. We also thank Léna Le Quellec, Sebastian Jaramillo-Riveri, Alessia Lepore, Benura Azeroglu, James Holehouse, and Rachel Jackson for their technical support and assistance with cloning, image analysis, and modelling. We thank Livia Scorza from the the Biological Research Data Management team for expert data curation. We are grateful to Ramon Grima, Rosalind Allen, David Leach, Peter Swain, and Johan Paulsson for fruitful discussions and generous advice. This work was supported by a Wellcome Trust Investigator Award 205008/Z/16/Z (to M.E.K.), a grant from the Wellcome Trust 102334 (to I.A.I.), the

Wellcome Trust Centre for Cell Biology core grant 092076, a Medical Research Council non Clinical Senior Research Fellowship MR/R008205/1 (to S.G.) and a Darwin Trust of Edinburgh postgraduate studentship (to I.K.).

Competing interests

The authors have no competing interests.

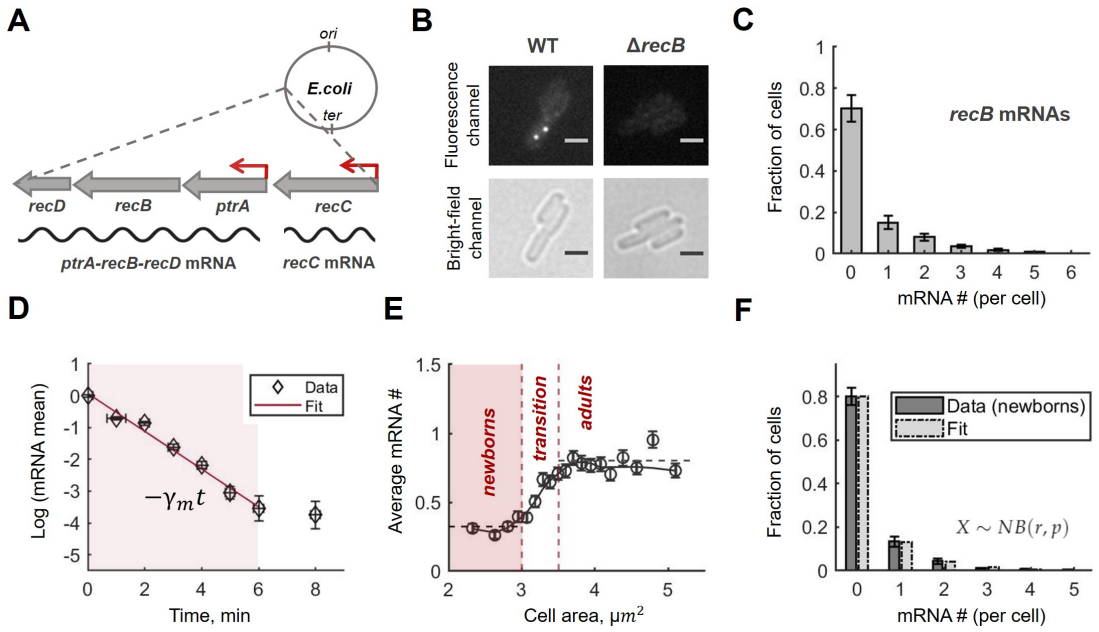

Figure 1: *recB* mRNAs are low abundant, short-lived and constitutively expressed.
A: Schematic description of the *recBCD* locus, its location on the *E. coli* chromosome and the corresponding mRNAs. **B:** Examples of fluorescence and bright-field images of *recB* mRNA FISH experiments in wild-type (WT) and Δ*recB* strains. Scale bars represent 2 μm. **C:** Total *recB* mRNA distribution quantified with smFISH and presented in molecule numbers per cell. The histogram represents the average across three replicated experiments; error bars reflect the maximum difference between the repeats. Total number of cells, included in the distribution, is 15,638. **D:** *recB* degradation rate measured in a time-course experiment where transcription initiation was inhibited with rifampicin. Mean mRNA counts, calculated from the total mRNA distributions for each time point, were normalized to the mean mRNA number at time $t = 0$ and represented in the natural logarithmic scale. Vertical error bars represent the standard error of the mean (s.e.m.); horizontal errors are given by experimental uncertainty on time measurements. Shaded area shows the time interval used for fitting. The red line is the fitted linear function, $-\gamma_m t$, where γ_m is the *recB* mRNA degradation rate. The final degradation rate was calculated as the average between two replicated time-course experiments (Table 7). **E:** *recB* mRNA molecule numbers per cell from the experiments in Fig. 1C shown as a function of cell area. The black circles represent the data binned by cell size and averaged in each group (mean +/- s.e.m.). The solid line connects the averages across three neighbouring groups. Based on the mean mRNA numbers, all cells were separated into three sub-populations: *newborns*, *cells in transition* and *adults*. **F:** Experimental data from Fig. 1C conditioned on cell size (<3.0 μm^2, *newborns*; total cell number is 2,180) fitted by a negative binomial distribution, $NB(r,p)$. The Kullback–Leibler divergence between experimental and fitted distributions is $D_{KL} = 0.002$.

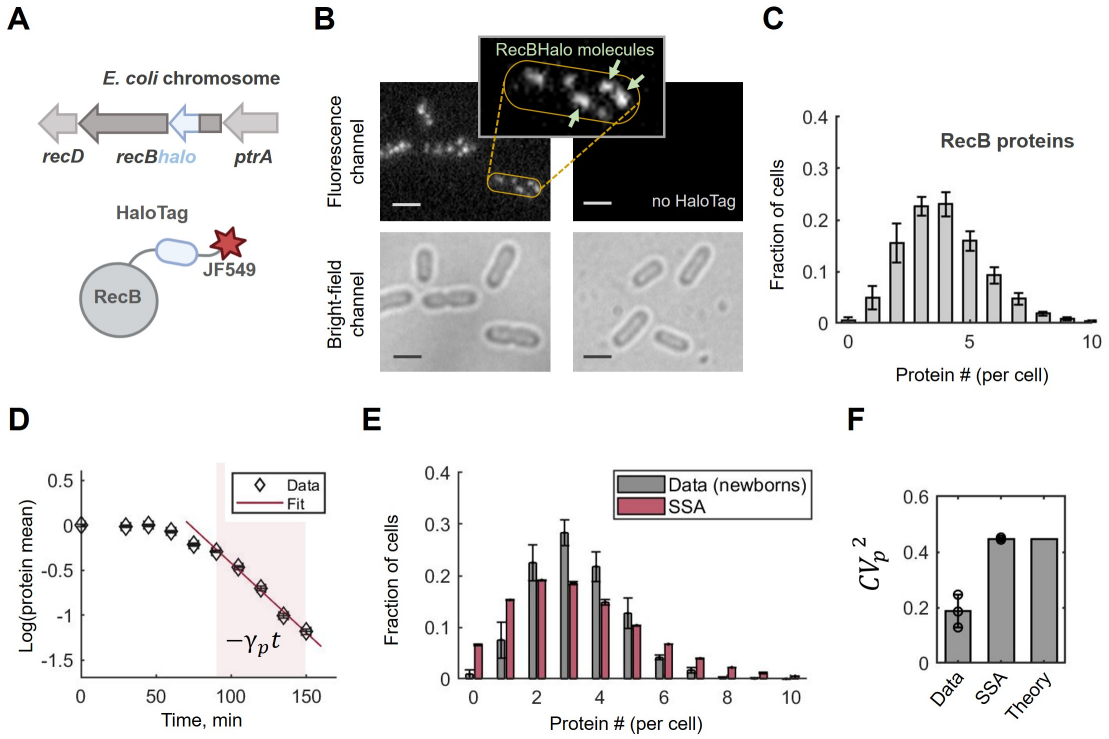

Figure 2: RecB proteins are low abundant, long-lived and show an evidence of noise reduction.
A: Schematic of RecB-HaloTag fusion at the endogenous *E. coli* chromosomal locus of the *recB* gene. RecB-HaloTag is conjugated to the Janelia Fluor 549 dye (JF549). **B:** Examples of fluorescence and bright-field Halo-labelling images for the strain with the RecBHalo fusion and its parental (no HaloTag) strain. Both samples were labelled with JF549 dye and the images are shown in the same intensity range. Scale bars represent 2 μm. **Zoom-in:** An example of a cell with five RecBHalo molecules (several single RecBHalo molecules are shown with light-green arrows). **C:** Total RecB protein distribution quantified with Halo-labelling and presented in molecules per cell. The histogram represents the average of three replicated experiments; error bars reflect the maximum difference between the repeats. Total number of analysed cells is 10,964. Estimation of false positives in no HaloTag control resulted in $\sim$0.3 molecules/cell. **D:** RecB removal rate measured in a pulse-chase Halo-labelling experiment where the dye was removed at time $t = 0$. Mean protein counts, calculated from the total protein distributions for each time point, were normalized to the mean protein number at time $t = 0$ and represented in the natural logarithmic scale. Shaded area shows the time interval used for fitting. The red line is the fitted linear function, $-\gamma_p t$, where γ_p is the RecB removal rate. The final removal rate was calculated as the average between two replicated pulse-chase experiments (Table 7). **E:** Comparison between the experimental RecB molecule distribution from Fig. 2C conditioned on cell size (in grey) and the results of Gillespie's simulations (SSA) for a two-stage model of RecB expression (in red). Parameters, used in simulations, are listed in Table 7. The Kullback–Leibler divergence between the distributions is $D_{KL} = 0.15$. **F:** Comparison of the coefficients of variation, $CV_p^2 = \sigma_p^2/\langle p \rangle^2$, for experimental (Data), simulated (SSA) data and analytical prediction (Theory). The experimental and simulated data from Fig. 2E were used; error bars represent standard deviations across three replicates; the black circles show the mean CV_p^2 in each experiment. Theoretical prediction was calculated using Eq. (1).

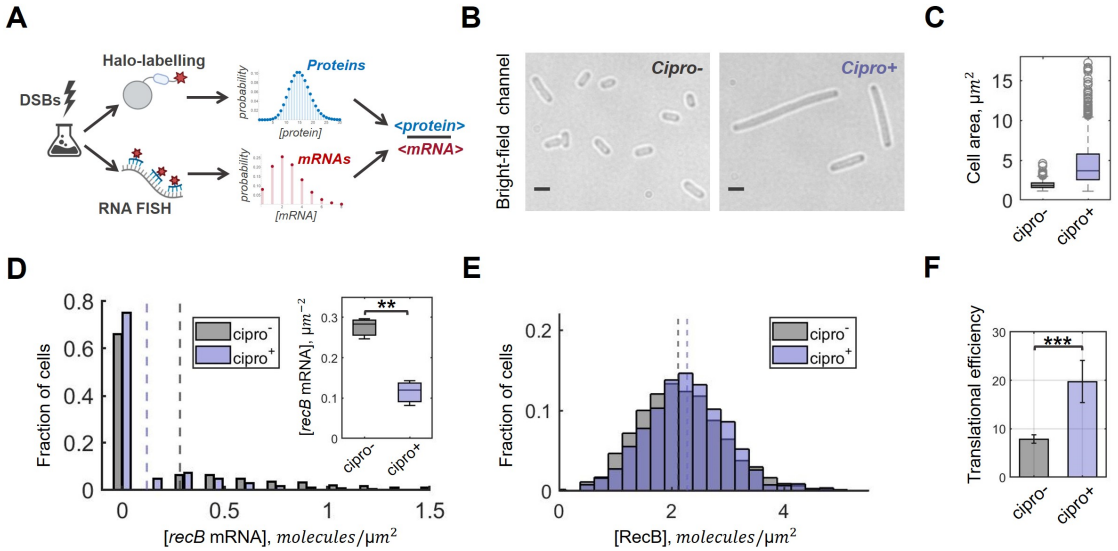

Figure 3: Translational efficiency of RecB is increased upon double-strand break (DSB) induction.
A: Schematic of the experimental workflow. DSBs were induced with 4 ng/ml of ciprofloxacin for two hours, and protein (mRNA) quantification was performed with Halo-labelling (smFISH). Mean protein and mRNA concentrations were calculated from the distributions, and the average protein-to-mRNA ratio (*translational efficiency*) was estimated. **B:** Examples of bright-field images for unperturbed ($Cipro^-$) and perturbed ($Cipro^+$) conditions. Scale bars represent 2 μm. **C:** Box plot with cell area distributions for perturbed (blue) and unperturbed (grey) samples. The medians of cell area in each sample are 1.9 μm^2 ($Cipro^-$) and 3.6 μm^2 ($Cipro^+$). **D:** *recB* mRNA concentration distributions quantified with smFISH in intact (grey) and damaged (blue) samples. The histograms represent the average of three replicated experiments. The medians of the *recB* mRNA concentrations are shown by dashed lines: 0.28 $mol/\mu m^2$ ($Cipro^-$) and 0.12 $mol/\mu m^2$ ($Cipro^+$). Total numbers of analysed cells are 11,700 ($Cipro^-$) and 6,510 ($Cipro^+$). **Insert:** Box plot shows significant difference between the average *recB* mRNA concentrations (P value = 0.0023, two-sample t-test). **E:** RecB concentration distributions quantified with Halo-labelling in unperturbed and perturbed conditions. The histograms represent the average of three replicated experiments. The medians are shown by dashed lines: 2.12 $mol/\mu m^2$ ($Cipro^-$) and 2.28 $mol/\mu m^2$ ($Cipro^+$). Total number of analysed cells are 1,534 ($Cipro^-$) and 683 ($Cipro^+$). The difference between the average RecB concentrations was insignificant (P value = 0.36, two-sample t-test). **F:** The average number of proteins produced per one mRNA in intact and damaged conditions. Average translational efficiency for each condition was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m. The error bars indicate the standard deviation of the data; statistical significance between the conditions was calculated with two-sample t-test (P value = 0.0001).

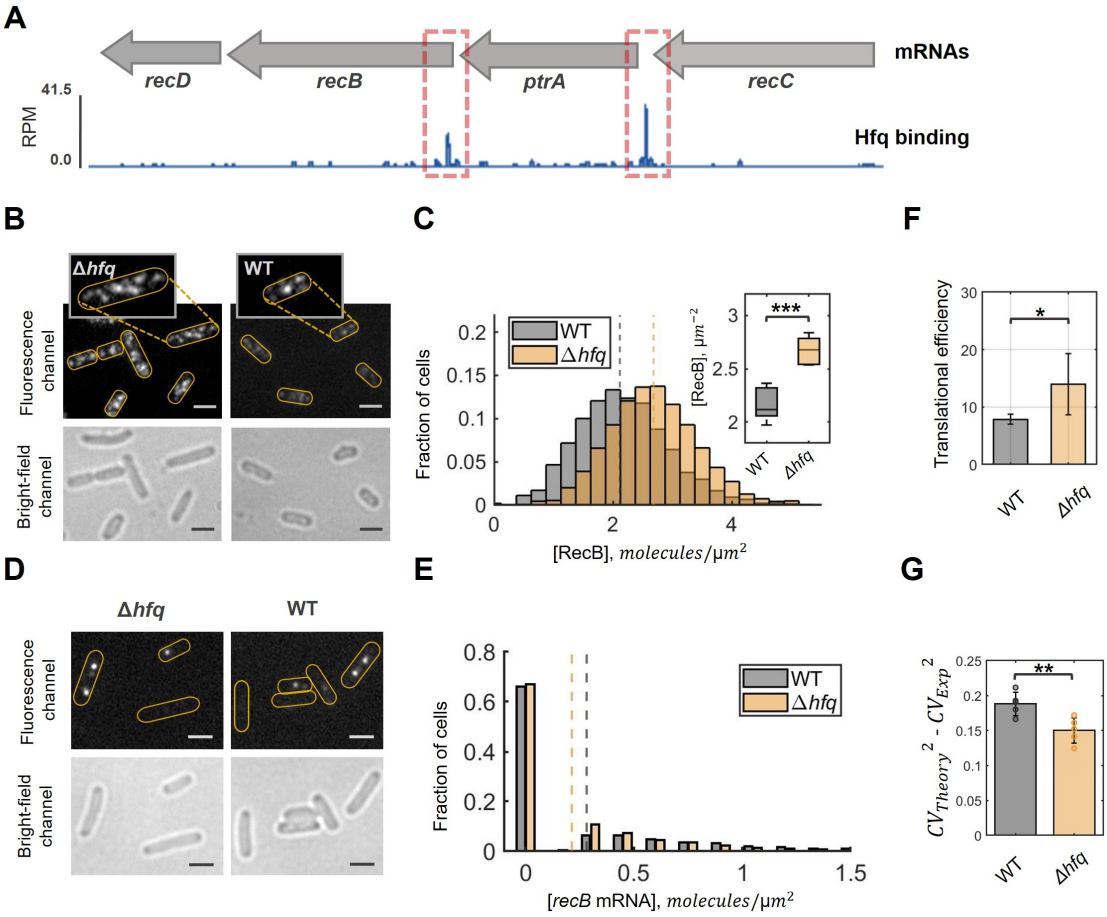

Figure 4: RecB expression is regulated by Hfq protein *in vivo*.
A: Genome browser track showing Hfq binding to *recC-ptrA-recB-recD* mRNAs. The coverage is normalized to reads per million (RPM). The major peaks of interest are highlighted by red dashed boxes. **B:** Examples of fluorescence and bright-field images of RecB quantification experiments in Δ*hfq* and wild-type strains. Yellow outlines indicate rough positions of bacterial cells in the fluorescence channel. Scale bars represent $2\mu m$. Both fluorescence images are shown in the same intensity range while different background modulation was applied in the zoom-in figures (for better spots visualization). **C:** RecB concentration distributions quantified with Halo-labelling in wild-type cells and the Δ*hfq* mutant. The histograms represent the average of five replicated experiments. The medians are shown by dashed lines: 2.12 $mol/\mu m^2$ for WT and 2.68 $mol/\mu m^2$ for Δ*hfq*. Total number of analysed cells are 3,324 (WT) and 2,185 (Δ*hfq*). **Insert:** Box plot shows significant difference (P value = 0.0007) between the average RecB concentrations in wild-type and Δ*hfq* cells verified by two-sample *t*-test. **D:** Examples of fluorescence and bright-field images of *recB* mRNA FISH experiments in Δ*hfq* and wild-type cells. Yellow outlines indicate rough positions of bacterial cells in the fluorescence channel. Scale bars represent 2 μm. **E:** *recB* mRNA concentration distributions quantified with smFISH in the Δ*hfq* mutant and wild-type cells. The histograms represent the average of three replicated experiments. The medians are shown by dashed lines: 0.28 $mol/\mu m^2$ for WT and 0.21 $mol/\mu m^2$ for Δ*hfq*. Total number of analysed cells are 7,270 (WT) and 5,486 (Δ*hfq*). The insignificant difference between the average *recB* mRNA concentrations in both strains was verified by two-sample *t*-test (P value = 0.11). **F:** RecB translational efficiency in wild-type and Δ*hfq* cells. Average translational efficiency (for each strain) was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m across replicated experiments. The error bars indicate the standard deviation of the data; statistical significance between the strains was calculated with two-sample t-test (P value = 0.014). **G:** The difference between theoretically predicted (CV^2_{Theory}) and experimentally measured (CV^2_{Exp}) coefficient of variation (squared). The theoretical values were computed according to Eq. (1). Error bars represent s.e.m. calculated from the replicated experiments. Statistical significance between the samples was calculated with two-sample t-test (P value = 0.0088).

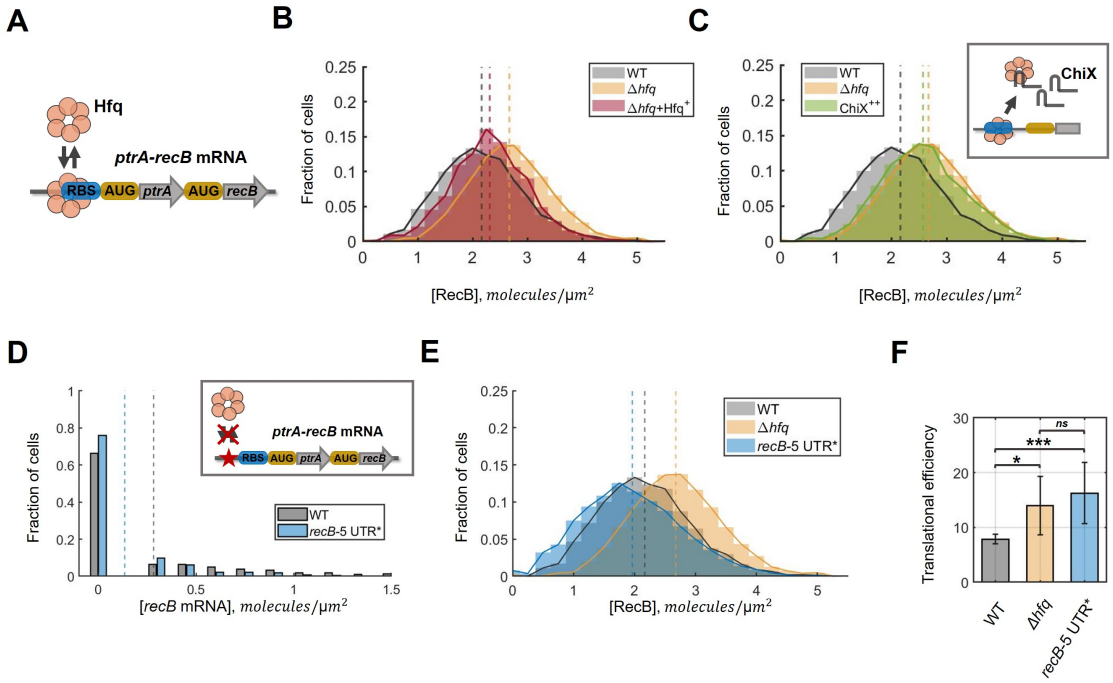

Figure 5: Specific alteration of RecB translation by Hfq *in vivo*.
A: A model of Hfq downregulating RecB translation by blocking the ribosome binding site of the *ptrA-recB* mRNA. **B:** Partial complementation of RecB expression to the wild-type level demonstrated by expression of a functional Hfq protein from a multicopy plasmid (pQE-Hfq) in Δ*hfq* (shown in red). The histogram represents the average of two replicated experiments. RecB concentration histograms in wild type and Δ*hfq* mutant from Fig. 4C are shown for relative comparison. Dashed lines represent the average RecB concentration in each condition. **C:** An increase in RecB protein production caused by sequestering of Hfq proteins with highly abundant small RNA ChiX (shown in green). The histogram represents the average of two replicated experiments. RecB concentration histograms in wild type and Δ*hfq* mutant from Fig. 4C are shown for relative comparison. Dashed lines represent the average RecB concentration in each condition. **D:** *recB* mRNA concentration distribution quantified with smFISH in the strain with the deletion of the main Hfq binding site *recB*-5′UTR* (shown in blue). The histogram represents the average of three replicated experiments. The *recB* mRNA distribution in the wild-type cells is shown for relative comparison. The medians are indicated by dashed lines. An approximate location of the removed sequence is schematically shown in the insert (red star). **E:** RecB protein concentration distribution quantified with RecBHalo-labelling in *recB*-5′UTR* strain (shown in blue). The histogram represents the average of three replicated experiments. RecB concentration histograms in wild type and Δ*hfq* mutant from Fig. 4C are shown for relative comparison. Dashed lines represent the average RecB concentration in each condition. **F:** Translational efficiency of RecB in wild type, Δ*hfq* mutant and the strain with the deletion of the main Hfq binding site, *recB*-5′UTR*. Average translational efficiency was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m across replicated experiments. The error bars indicate the standard deviation of the data; statistical significance between the samples was calculated with two-sample t-test. The P value for WT and *recB*-5′UTR* is 0.0007; while the difference between Δ*hfq* and *recB*-5′UTR* is non-significant (P value > 0.05).

Supplementary Information

1 A stochastic two-stage model of RecB expression

To model RecB expression, we applied a well-established stochastic mRNA-protein model of gene expression and inferred its parameters by fitting to the experimental data. Here, we briefly introduce the model, discuss its underlying assumptions and provide the analytical expressions used in the fitting procedure.

Model description

A stochastic two-stage model of gene expression has been widely utilized to study fluctuations underlying protein production in prokaryotic and eukaryotic cells [57, 68, 69, 115, 116, 117]. Generally, the model is based on four biological processes: transcription of a gene by an RNA polymerase, translation of an mRNA by a ribosome, mRNA degradation and protein decay. As elsewhere, we assume a cell to be a well-stirred system of biochemical molecules, the interaction between which can be described with a common approach of chemical reactions [118]. We focus on two species, mRNAs and proteins, and denote molecule numbers in a cell as m and p, respectively.

Considering gene expression in bacterial cells allows several further assumptions to be made. First, transcription of bacterial genes was demonstrated to occur in bursts [119, 120] and quite a few transcriptional models describing gene bursting have been proposed in the literature [69, 121, 122, 123]. Thus, mRNA production is described as a zeroth-order chemical reaction where a gene is transcribed in mRNA bursts with a constant rate k_m and the number of mRNAs per each burst n is sampled from a geometric distribution with a mean size b [55, 56, 61]. Secondly, based on a broadly accepted view of the exponential RNA degradation in bacteria [124, 125, 126], mRNA decay is modelled as a first order reaction with a constant rate γ_m. Furthermore, proteins in bacteria were shown to be diluted exclusively because of cell growth and division [127, 128]. This process can be implicitly modelled by a first-order chemical reaction with a constant rate equal to growth rate γ_p. Finally, we assume that a protein is produced from an mRNA with a constant rate k_p. Thus, the entire reaction network for the two-stage model with burst transcription is summarized in the following schematic:

$$\begin{aligned} \emptyset &\xrightarrow{k_m} n\,m \qquad & m &\xrightarrow{k_p} m + p \\ m &\xrightarrow{\gamma_m} \emptyset \qquad & p &\xrightarrow{\gamma_p} \emptyset \end{aligned} \tag{A1}$$

The dynamics of the mRNA and protein average levels (denoted as $\langle m \rangle$ and $\langle p \rangle$) is described by a set of the following rate equations:

$$\frac{d\,\langle m \rangle}{dt} = b\,k_m - \gamma_m\,\langle m \rangle \tag{A2}$$

$$\frac{d\,\langle p \rangle}{dt} = k_p\,\langle m \rangle - \gamma_p\,\langle p \rangle \tag{A3}$$

Steady-state mRNA and protein mean and variance

The stationary solutions for mean mRNA and protein numbers can be found by imposing the steady-state conditions ($d\,\langle m \rangle\,/dt = 0$ and $d\,\langle p \rangle\,/dt = 0$) to the rate equations (Eq. (A2) & Eq. (A3)):

$$\langle m \rangle = b\frac{k_m}{\gamma_m} \qquad \langle p \rangle = b\frac{k_m k_p}{\gamma_m \gamma_p} \tag{A4}$$

As a result of the linearity of the system, the mRNA and protein variances can be found exactly by van Kampen's Ω-expansion (also known as the Linear Noise Approximation and the Fluctuation-Dissipation Theorem) [31, 118, 129].

$$\sigma_m^2 = (1+b)\,\langle m \rangle \qquad \sigma_p^2 = \langle p \rangle + \frac{k_p^2(1+b)\,\langle m \rangle}{\gamma_p(\gamma_m + \gamma_p)} \tag{A5}$$

Finally, the mRNA and protein coefficients of variation (squared), defined as variance over squared mean, are given as follows [57, 68, 69, 71]:

$$CV_m^2 = \frac{\sigma_m^2}{\langle m \rangle^2} = \frac{1+b}{\langle m \rangle} \tag{A6}$$

$$CV_p^2 = \frac{\sigma_p^2}{\langle p \rangle^2} = \frac{1}{\langle p \rangle} + \frac{1+b}{\langle m \rangle}\frac{\gamma_p}{\gamma_m + \gamma_p} \tag{A7}$$

The analytical probability distribution of protein numbers under assumption of a long-lived protein relative to short-lived mRNA was derived here [57].

2 Deterministic model of RecB expression upon DSB induction

In this section, we analytically describe the dynamics of RecB mRNA and protein levels after a perturbation made by DSB induction with a sub-lethal dose of ciprofloxacin. Taking into account different time-scales of the response in mRNA and protein levels, we address the following question: how long does the DSB induction need to be applied for to see changes (if any) in both species? We focus on the average mRNA and protein concentrations instead of molecule numbers (because of cell elongation upon ciprofloxacin treatment (Fig. 3B&C) and modify a two-stage model of RecB expression by including deterministic time-dependency of cell volume, $V(t)$.

As discussed above, we consider the same set of reactions for two species of the system: mRNAs and proteins. While the processes of translation, active mRNA and protein degradation are described by first-order reactions with constant rates: k_p, γ_m and $\gamma_{p_{act}}$, respectively, mRNA transcription is given by a zeroth-order reaction. According to the law of mass action, we shall scale a zeroth-order reaction rate by the volume of the system, as $k_m V(t)$ [118, 130]. We define the number of mRNA and protein molecules at time t in a cell of volume $V(t)$ as $m(t)$ and $p(t)$, respectively. Then, the evolution of the average numbers of mRNA and protein molecules in time is described by the following rate equations:

$$\frac{d\langle m(t)\rangle}{dt} = k_m V(t) - \gamma_m \langle m(t)\rangle \tag{A8}$$

$$\frac{d\langle p(t)\rangle}{dt} = k_p \langle m(t)\rangle - \gamma_{p_{act}} \langle p(t)\rangle \tag{A9}$$

We further define *recB* mRNA and RecB protein concentration as $c_m(t) = m(t)/V(t)$ and $c_p(t) = p(t)/V(t)$ and differentiate with respect to t:

$$\frac{dc_m(t)}{dt} = \frac{1}{V(t)}\frac{dm(t)}{dt} - \frac{c_m(t)}{V(t)}\frac{dV(t)}{dt} \tag{A10}$$

$$\frac{dc_p(t)}{dt} = \frac{1}{V(t)}\frac{dp(t)}{dt} - \frac{c_p(t)}{V(t)}\frac{dV(t)}{dt} \tag{A11}$$

In order to obtain time-dependent solutions for mRNA and protein concentrations, we divide the rate equations (Eq. (A8) & Eq. (A9)) by cell volume $V(t)$ and replace $dm(t)/dt$ and $dp(t)/dt$ with the expressions from the equations (Eq. (A10) & Eq. (A11)). Based on the experimental evidence [131, 132], we assume the exponential dependence of cell volume from time, $V(t) = V_0 e^{\gamma t}$, where γ is a growth rate. Thus, we obtain the following ordinary differential equations for mRNA and protein concentrations:

$$\frac{d\langle c_m(t)\rangle}{dt} = k_m - (\gamma_m + \gamma)\langle c_m(t)\rangle \tag{A12}$$

$$\frac{d\langle c_p(t)\rangle}{dt} = k_p \langle c_m(t)\rangle - (\gamma_{p_{act}} + \gamma)\langle c_p(t)\rangle \tag{A13}$$

By equating the left-hand sides of the equations to zero, we find the steady-state mRNA and protein concentrations, c_m^{ss} and c_p^{ss}:

$$c_m^{ss} = \frac{k_m}{\gamma_m + \gamma} \qquad c_p^{ss} = \frac{k_m}{(\gamma_m + \gamma)}\frac{k_p}{(\gamma_{p_{act}} + \gamma)} \tag{A14}$$

Next, by assuming the initial conditions for mRNA and protein concentration to be $c_{m(t=0)}$ and $c_{p(t=0)}$, we obtain the time-dependent solutions for mRNA and protein concentrations:

$$\langle c_m(t) \rangle = \left(c_{m0} - \frac{k_m}{\gamma_m + \gamma} \right) e^{-(\gamma_m+\gamma)t} + \frac{k_m}{\gamma_m + \gamma} \tag{A15}$$

$$\begin{aligned} \langle c_p(t) \rangle = & \left(c_{p0} + \frac{c_{m0} k_p}{\gamma_m - \gamma_{p_{act}}} - \frac{k_p k_m}{(\gamma_m - \gamma_{p_{act}})(\gamma_{p_{act}} + \gamma)} \right) e^{-(\gamma_{p_{act}}+\gamma)t} + \\ & + \left(-\frac{c_{m0} k_p}{\gamma_m - \gamma_{p_{act}}} + \frac{k_p k_m}{(\gamma_m - \gamma_{p_{act}})(\gamma_m + \gamma)} \right) e^{-(\gamma_m+\gamma)t} + \frac{k_p k_m}{(\gamma_{p_{act}} + \gamma)(\gamma_m + \gamma)} \end{aligned} \tag{A16}$$

The last expression can be simplified by taking into account that the RecB protein does not have an active degradation mechanism (Fig. 2D). Thus, after imposing the condition of $\gamma_{p_{act}} = 0$, we obtain a time-dependent solution for RecB protein concentration as follows:

$$\langle c_p(t) \rangle = \left(c_{p0} + \frac{c_{m0} k_p}{\gamma_m} - \frac{k_p k_m}{\gamma_m \gamma} \right) e^{-\gamma t} + \left(-\frac{c_{m0} k_p}{\gamma_m} + \frac{k_p k_m}{\gamma_m (\gamma_m + \gamma)} \right) e^{-(\gamma_m+\gamma)t} + \frac{k_p k_m}{\gamma(\gamma_m + \gamma)} \tag{A17}$$

Now we can estimate the mRNA and protein response times from the expressions Eq. (A15) & Eq. (A17), respectively. It is worth noting that although cells are elongated upon ciprofloxacin treatment, they do not slow down the growth (Fig. S5). Thus, the analysis of the obtained equations shows that mRNA solution has one mode, $e^{-(\gamma_m+\gamma)t}$, and hence a characteristic response time-scale can be estimated as $1/(\gamma_m + \gamma) \sim 1.6\ min$ (the results of growth rate measurements, $\gamma = \gamma_{cipro^+} = 0.0165\, min^{-1}$, are shown in Fig. S5c). In contrast, the expression for protein concentration is given by two time-scales: $e^{-\gamma t}$ and $e^{-(\gamma_m+\gamma)t}$, where the slowest one defines the response time as $1/\gamma \sim 61\ min$. This means that a perturbation needs to be applied for at least ~ 1 hour in order to be able to detect changes in protein concentrations. In our experiments, DSBs were induced for two hours to guarantee fulfillment of the time-scale requirement (Fig. 3).

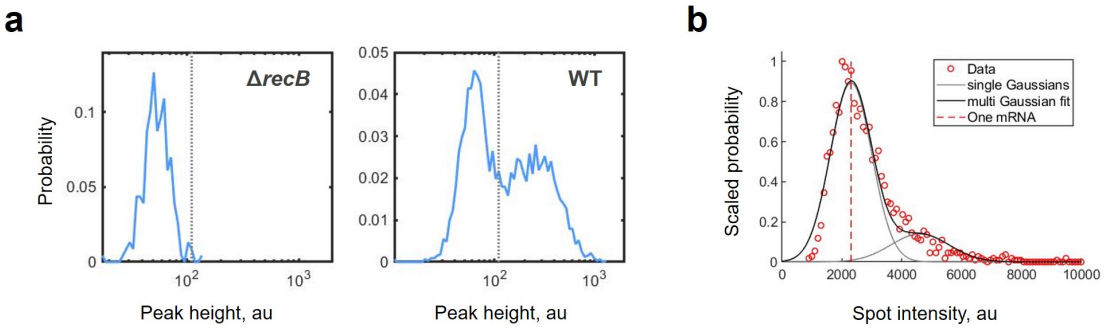

Figure S1: smFISH image analysis performed by Spätzcell. a: Peak height intensity profiles for spots detected in the wild-type and Δ*recB* samples. The dashed line indicates the intensity threshold (99.9%) that separates a specific signal from the background. **b:** Integrated spot intensity histogram for foci detected in the wild-type sample. The data, binned and averaged in groups, are shown in red circles. The data were fitted by the sum of two Gaussian distributions corresponding to one and two mRNA molecule(s) per focus. The grey lines correspond to single Gaussian distributions, while the black solid curve is the sum of these two Gaussians. The red dashed line indicates the intensity equivalent corresponding to the total integrated intensity of FISH probes (in average) bound to a single mRNA (One mRNA).

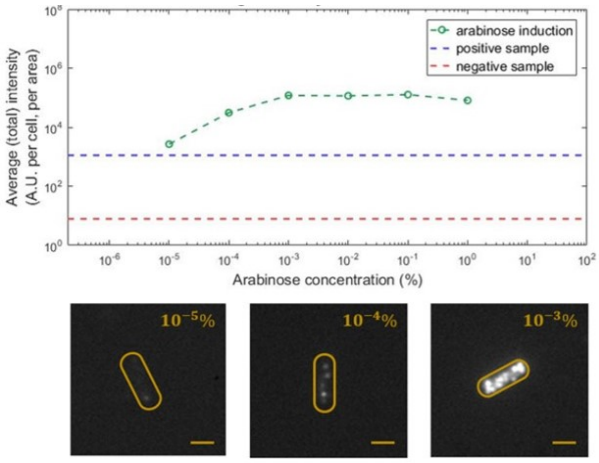

Figure S2: Sensitivity of the smFISH protocol allows for quantification of low-abundant *recB* mRNAs. Top: Fluorescence signal, detected in the samples with *recB* over-expression from an arabinose-inducible plasmid (pIK02), is presented as a function of arabinose concentration (green-dashed line). The intensity of the fluorescence signal in the wild-type and Δ*recB* strains is indicated with blue-dashed and red-dashed lines, respectively. More than 350 cells for each induction condition were taken for the analysis. **Bottom:** Representative fluorescence images of the samples where *recB* expression was induced with 10^{-3}%, 10^{-4}% or 10^{-5}% of arabinose. All fluorescence images are shown in the same intensity range. The yellow line segment represents 1 μm.

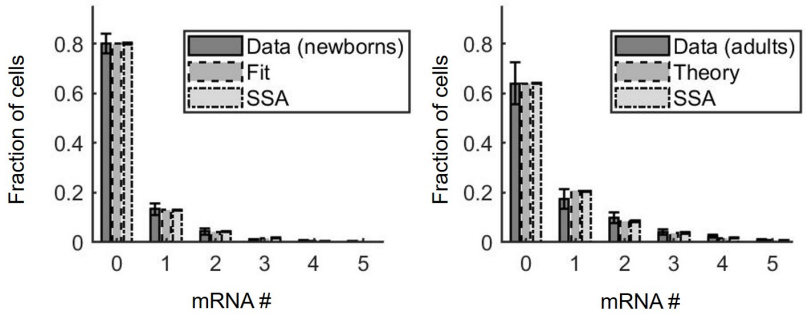

Figure S3: Independent transcription from *recB* gene copies. *recB* mRNA distributions for *newborns* (**left**) and *adults* (**right**). The experimental data from Fig. 1C was conditioned on cell size (<3.0 μm^2 for *newborns*) and (>3.5 μm^2 for *adults*). Total number of cells: 2,180 (*newborns*) and 6,413 (*adults*). The *recB* mRNA statistics for *newborns* is well described by a negative binomial distribution, $NB(r, p)$ (Fig. 1F). The sum of two independent variables, each of which is distributed as $NB(r, p)$, is then described by a negative binomial distribution $NB(2r, p)$. Thus, the *adults* data was compared to the predicted distribution, $NB(2r, p)$, based on (i) the assumption of independent transcription from *recB* gene copies and (ii) the fitting parameters inferred on the *newborns*. The results were verified with Gillespie's simulations (SSA). The Kullback–Leibler divergence between the experimental and simulated distributions for *newborns* and *adults* is $D_{KL} = 0.003$ and $D_{KL} = 0.007$, respectively.

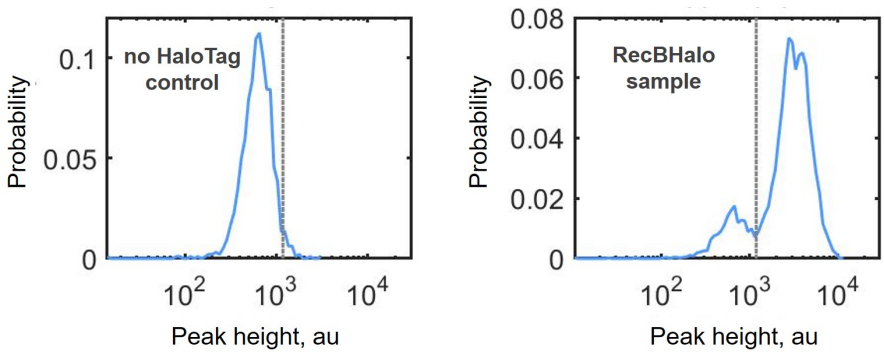

Figure S4: Image analysis of Halo-labelling experiments performed by a modified version of Spätzcell. Peak height intensity profiles for spots detected in the strain with RecB-HaloTag fusion (RecBHalo sample) and its parental strain (no HaloTag control). The dashed line indicates a chosen intensity threshold in order to separate a specific signal from the background.

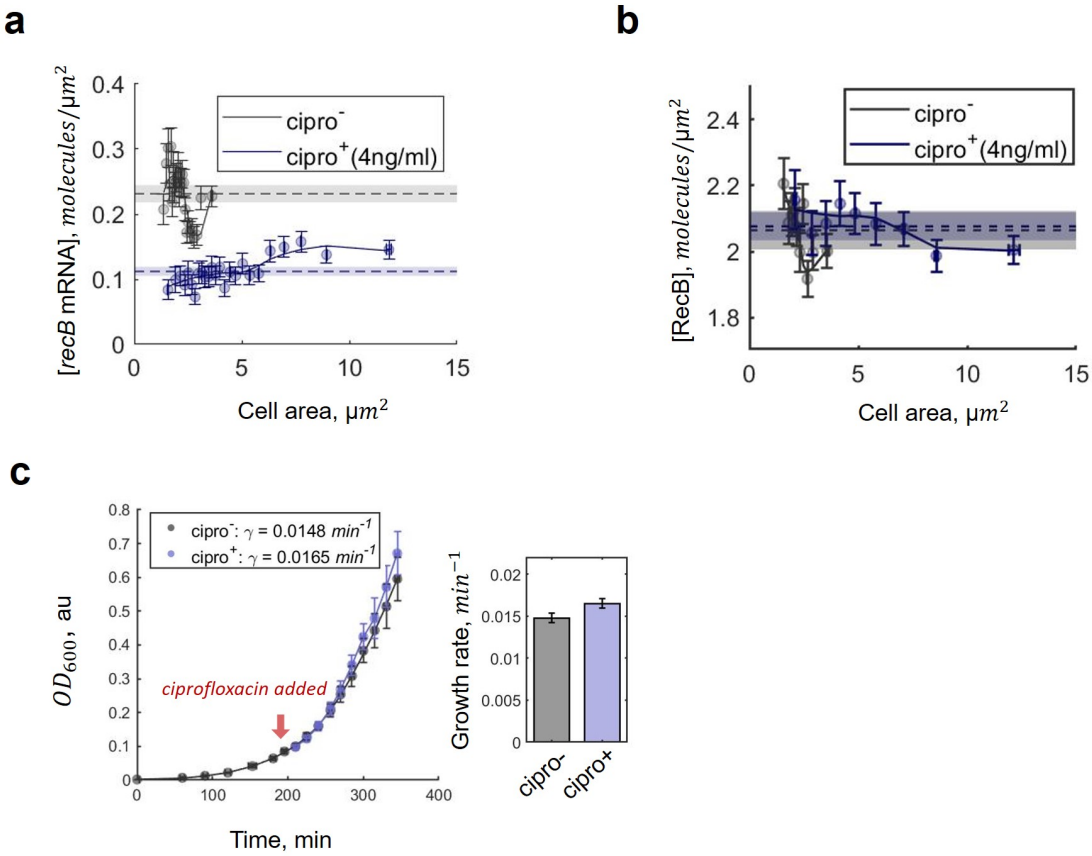

Figure S5: Analysis of *recB* mRNA and RecB protein concentrations and growth rate measurements upon DSB induction. **a:** *recB* mRNA concentration for perturbed (blue) and unperturbed (grey) samples represented as a function of cell area. The data from Fig. 3D were binned by cell size and averaged in each group (mean ± s.e.m.). The solid lines connect the averages while the dashed lines and shaded areas show mean ± s.e.m. calculated across all cells in each sample. **b:** RecB protein concentration for perturbed (blue) and unperturbed (grey) samples shown as a function of cell area. The data from Fig. 3E were binned by cell size and averaged in each group (mean ± s.e.m.). The solid lines connect the averages while the dashed lines and shaded areas show mean ± s.e.m. calculated across all cells in each sample. **c:** Growth curves obtained with optical density measurements (OD_{600}). Cell cultures were grown with 4 ng/ml (blue) or without (grey) ciprofloxacin. The red arrow indicates the time when ciprofloxacin was added to the $cipro^{+}$ sample. The exponential phase of growth ($OD_{600} \sim$ [0.08-0.4]) was used for growth rate estimation. The growth rates, averaged across three replicated experiments, are shown in the legend. Error bars on the bar chart represent 95% confidence intervals.

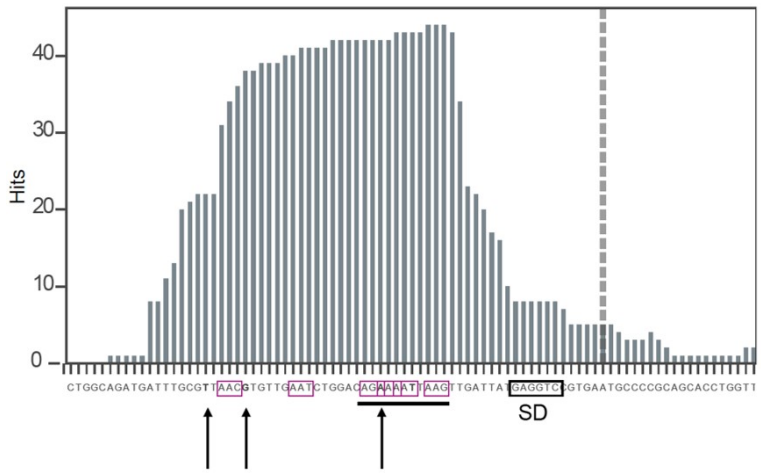

Figure S6: Main Hfq binding peak within *ptrA* gene identified in the Hfq-CRAC experiment in *E. coli* [78]. The Hfq binding peak located within *ptrA-recB-recD* mRNA presented in raw counts (Hits). The vertical dashed line and black box indicate a translation initiation site (ATG) and Shine-Dalgarno sequence (SD), respectively. The black vertical arrows show the nucleotides where direct Hfq cross-linking was detected in the CRAC data. Hfq-binding motifs, A-R(A/G)-N(any nucleotide) [73, 79], are highlighted in the red boxes while the cluster of Hfq binding motifs is underlined.

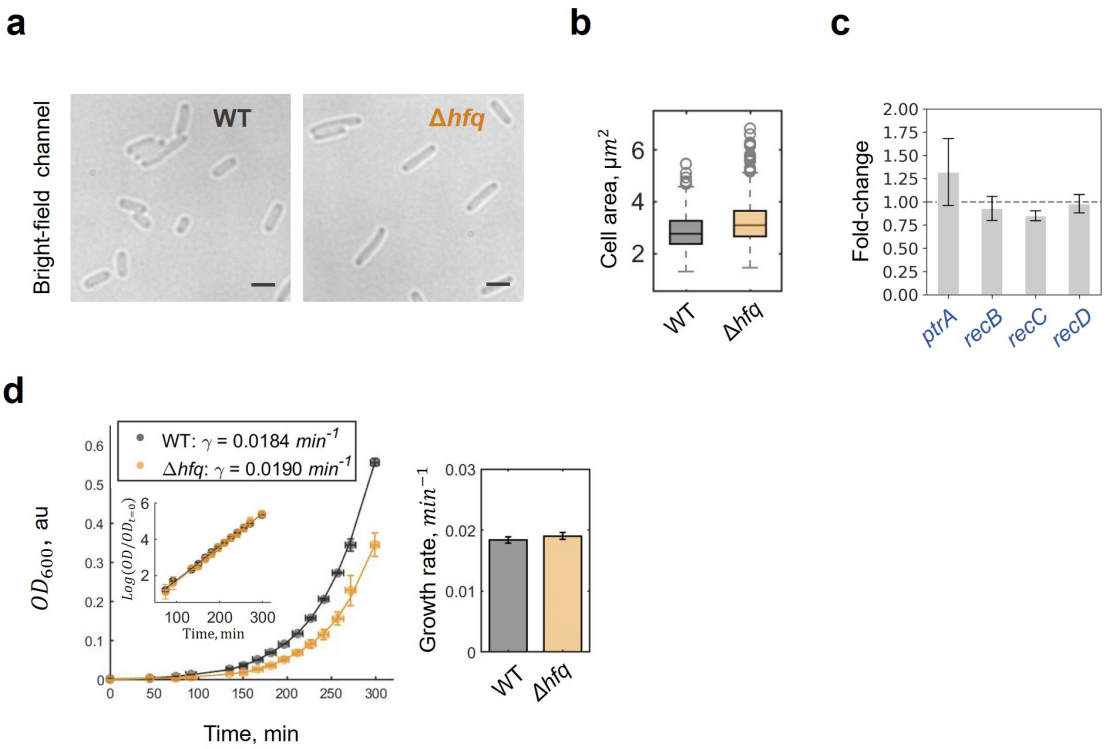

Figure S7: Cell size analysis, RT-qPCR quantification and growth rate measurements in the Δ*hfq* mutant. a: Examples of bright-field images for wild-type and Δ*hfq* strains. Scale bars represent 2 μm. **b:** Box plots represent cell area distributions for wild-type and Δ*hfq* samples. **c:** *ptrA*, *recB*, *recC* and *recD* transcripts quantified by RT-qPCR in the Δ*hfq* mutant and normalized to the corresponding expression levels in wild-type cells. The data represent averages and standard deviations across three replicates for each gene. *rrfD* was used as a reference gene. **d:** Growth curves obtained by optical density measurements (OD_{600}) in Δ*hfq* (orange) and wild type (grey). The strains were grown in the medium supplemented with glucose and amino acids. The exponential phase of growth ($OD_{600} \sim$ [0.08-0.4]), shown in the logarithmic scale in the insert, was used for growth rate calculation. The growth rates, averaged across three replicated experiments, are shown in the legend. Error bars on the bar chart represent 95% confidence intervals.

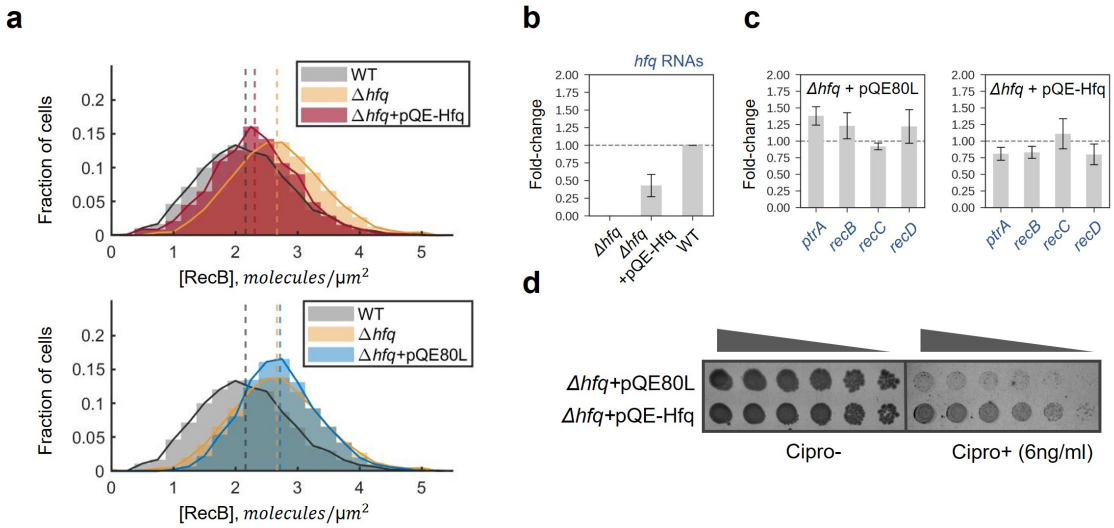

Figure S8: Controls for Hfq complementation experiment. **a:** RecB protein concentration distributions quantified in the Δ*hfq* mutant carrying pQE80L (blue) or pQE-Hfq (red) plasmid. The histograms represent the average across two replicated experiments per each condition. The RecB concentration histograms for wild type and Δ*hfq* mutant are plotted as references in grey and orange, respectively. The dashed lines represent the mean RecB concentration in each condition. Significance was evaluated with two-sample *t*-test (*P* values: 0.045(*) for Δ*hfq* and Δ*hfq*+pQE-Hfq; 0.71(ns) for Δ*hfq* and Δ*hfq*+pQE80L). **b:** *hfq* RNA levels quantified by RT-qPCR in Δ*hfq* cells carrying pQE-Hfq plasmid and normalized to the *hfq* RNA level in wild-type cells. **c:** *ptrA*, *recB*, *recC* and *recD* transcripts quantified by RT-qPCR in Δ*hfq* carrying pQE80L (left) or pQE-Hfq (right) plasmid and normalized to the corresponding expression levels in Δ*hfq* and Δ*hfq*+pQE80L, respectively. **d:** A 5-fold serial dilution assay of Δ*hfq* strain carrying pQE80L or pQE-Hfq plasmid. Cells were plated onto LB plates without or with 6 ng/ml of ciprofloxacin.

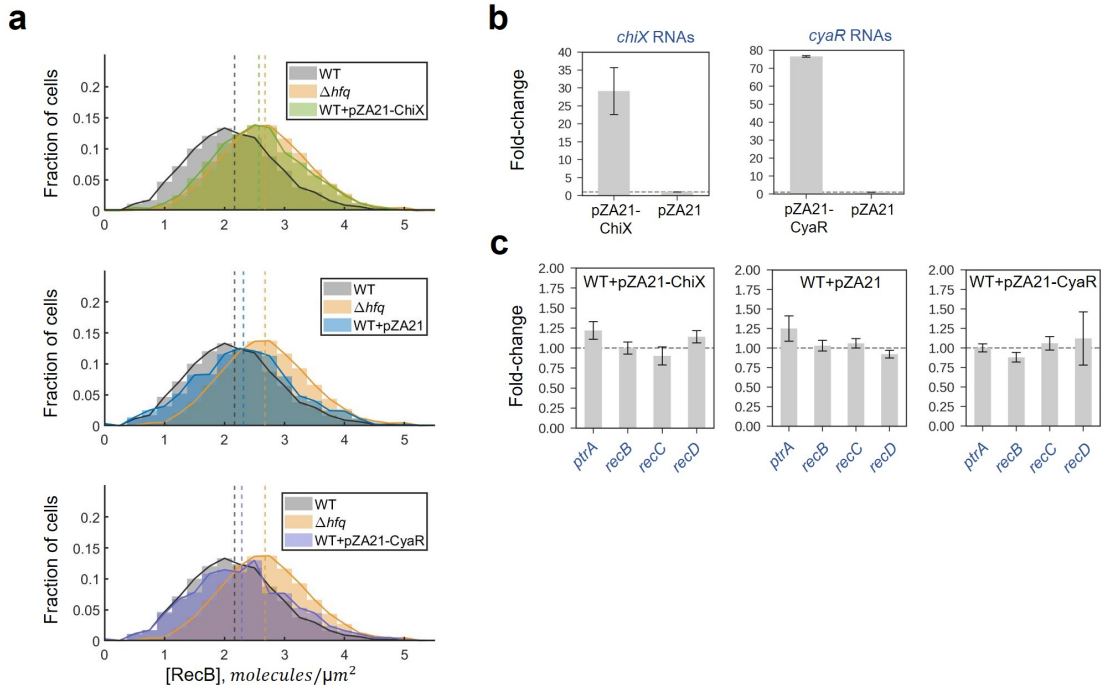

Figure S9: Controls for ChiX over-expression experiment. **a:** RecB protein concentration distributions quantified in wild-type cells carrying pZA21-ChiX (green), pZA21 (blue) or pZA21-CyaR (purple) plasmid. The histograms represent the average across two replicated experiments per each condition. The RecB concentration histograms for wild type and the Δ*hfq* mutant are plotted as references in grey and orange, respectively. The dashed lines represent the mean RecB concentration in each condition. Significance was evaluated with two-sample *t*-test (*P* values: 0.021(*) for WT and WT+pZA21-ChiX; 0.42(ns) for WT and WT+pZA21; 0.37(ns) for WT and WT+pZA21-CyaR). **b:** *chiX* and *cyaR* RNA levels quantified by RT-qPCR in wild-type cells carrying pZA21-ChiX (left) or pZA21-CyaR (right) over-expression plasmid. The results were normalized to the corresponding expression in the cells carrying the backbone plasmid (pZA21). **c:** *ptrA*, *recB*, *recC* and *recD* transcripts quantified by RT-qPCR in wild-type cells carrying pZA21-ChiX (left), pZA21 (middle) or pZA21-CyaR (right) plasmid. RNA levels were normalized to the corresponding expression levels in wild type.

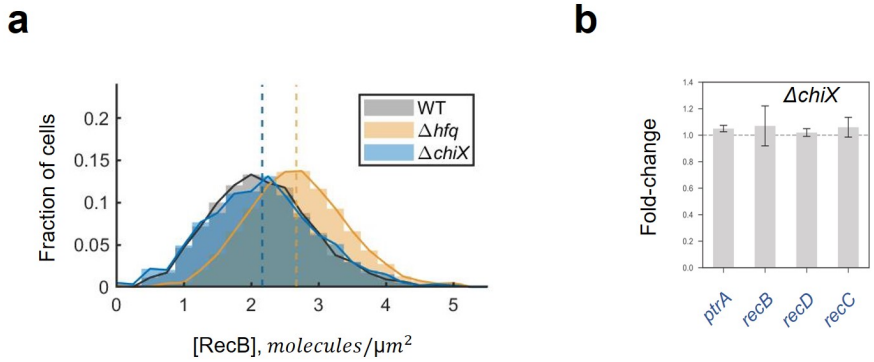

Figure S10: RecB protein and mRNA quantification in Δ*chiX*. a: RecB protein concentration distribution quantified in Δ*chiX* cells. The histogram represents the average across two replicated experiments. The RecB concentration histograms for wild type and the Δ*hfq* mutant are plotted as references in grey and orange, respectively. The dashed lines represent the mean RecB concentration in each condition. Significance was evaluated with two-sample *t*-test (*P* value for WT and Δ*chiX* is 0.96(ns)). **b:** *ptrA*, *recB*, *recD* and *recC* transcripts quantified by RT-qPCR in *chiX* mutants and normalized to the corresponding expression levels in the wild-type sample.

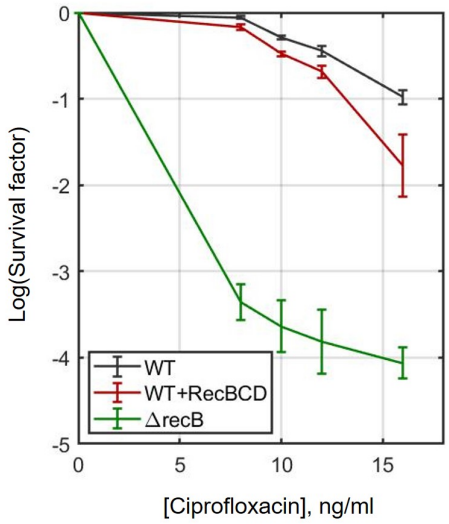

Figure S11: Toxicity of RecBCD overexpression upon DSB induction. Viability assays performed for the wild type, wild type carrying RecBCD over-expression plasmid, pDWS2 [133], and Δ*recB*. Cells were plated onto LB plates supplemented with ampicillin and either without or with 8/10/12/16 *ng/ml* of ciproflaxacin. The average survival factors were calculated for at least three replicated experiments while the error bars indicate standard estimation of the mean.

Table 1: *E. coli* strains used in the study.

Strain	**Genotype**	**Source**
MG1655	$F^-\lambda^-$ *ilvG*$^-$ *rfb-50 rph-1*	Lab stock
BW27783	$F^-\lambda^-$ Δ*(araD-araB)567* Δ*lacZ4787*::rrnB-3 Δ*(araH-araF)570(::FRT)* Δ*araEp-532::FRT* ϕP_{cp8}*araE535 rph-1* Δ*(rhaD-rhaB)568 hsdR514*	Lab stock
DH5α	$F^-\phi$80*lacZ*Δ*M*15 Δ(*lacZYA-argF*)*U*169 *recA1 endA1 hsdR17*(r_K^-, m_K^+) *phoA supE44* λ^-*thi-1 gyrA96 relA1*	Lab stock
MEK65	MG1655 *recB165*::*halotag*	[23]
MEK1326	MG1655 Δ*recB*	[23]
MEK1329	BW27783 Δ*recB*	This work
MEK1902	MG1655 Δ*hfq*	This work
MEK1457	MG1655 *recB165*::*halotag* Δ*hfq*	This work
MEK1888	MG1655 Δ*chiX*	This work
MEK1449	MG1655 *recB165*::*halotag* Δ*chiX*	This work
MEK1938	MG1655 *recB165*::*halotag* 5′UTR-*recB* Δ$(TTAA...TGAT)_{36nt}$	This work

Table 2: Plasmids used in the study.

Plasmids	**Description**	**Source**
pTOF24Δ*recB*	pTOF24-derivative plasmid used for construction of Δ*recB* strain	[108] [1]
pTOF24 *recB*-5′UTR	pTOF24-derivative plasmid used for construction of *recB*-5′UTR strain	This work
pBAD33	Backbone plasmid used for construction of pIK02	[134]
pIK02	pBAD33-derivative plasmid carrying *recB* gene under control of arabinose-inducible promoter, P*araBAD*	This work
pZA21MCS	Backbone plasmid used for construction of pZA21-ChiX	Expressys
pZA21	pZA21MCS expression vector with MCS removed from +1 to terminator	[73]
pZA21-ChiX	pZA21MCS-derivative plasmid carrying *chiX*	This work
pZA21-CyaR	pZA21MCS-derivative plasmid carrying *cyaR*	[78]
pDWS2	pBR322-derivative plasmid carrying native *recC-ptrA-recB-recD* chromosomal region	[133] [2]
pQE80L	Backbone plasmid used for construction of pQE-Hfq	[83] [3]
pQE-Hfq	pQE80L-derivative plasmid carrying *hfq*	[83] [3]

[1] Gift from Prof. David Leach, The University of Edinburgh
[2] Gift from Prof. Gerald Smith, The Fred Hutchinson Cancer Research Center
[3] Gift from Prof. Teppei Morita and Prof. Hiroji Aiba, Suzuka University of Medical Sciences

Table 3: Primers used for strain and plasmid construction.

Primer ID	**Sequence (5′-3′)**
hfq_H1_P1	GAATCGAAAGGTTCAAAGTACAAATAAGCATATAAGGAAAAGAG AGAATGGTGTAGGCTGGAGCTGCTTC
hfq_H2_P2	CTCCCCGTGTAAAAAAACAGCCCGAAACCTTATTCGGTTTCTTCGC TGTCGGTCCATATGAATATCCTCCTTAG
ChiX_H1_P1	TCTTGCCTAAGAGTATTGGCAGGATGGTGAGATTGAGCGACAATC GAGTTGTGTAGGCTGGAGCTGCTTC
ChiX_H2_P2	CACCTGTATGGAGAAGGGAATTTGCCGCAAATGTTGCGCTAAAAA AATGGCGGTCCATATGAATATCCTCCTTAG
chiX_ZA21	ACACCGTCGCTTAAAGTGACGGCATAATAATAAAAAAATGAAATT CCTCTTTGACGGGCCAATAGCGATATTGGCCATTTTTTTGGTACGCG TGCTAGAGGCATC
pZA21MCS_5P_rev	5P-GTGCTCAGTATCTCTATCACTGA
oIK03	AAGCTTGGCTGTTTTGGCGGATGAG
oIK04	ACCGAGCTCGAATTCGCTAGCCCAAAA
oIK10	ACCCGTTTTTTTGGGCTAGCGAATTCGAGCTCGGTCCTGATGAGTG AAAAGAATGAGTG
oIK11	GAAAATCTTCTCTCATCCGCCAAAACAGCCAAGCTTCTCCACAGCT TCCAGTAATTGC

Table 4: The sequence of the gBlock used for the construction of the *recB*-5′UTR strain.

5′-GTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGGTTCAG CAATAGCATCGGCCAGGCGGTCTACCGCACCAGGCAAGGCGTCGTTCTCAACTTCCAGAT AGAAAGCCGTGCGATACGGCGCAGTGCTGGCATTGTGACTACCGCCGTGCATTTTGAGAT ATTCGGCCAGACTGTCAGCCTGCGGGTACTTTTTCGACCCCATCAGACTCATATGTTCAAG GTAATGTGCCAGCCCCTGGTACGCCTCGGGATCTTCCAGCGACCCAACGGGCACCACCAG CGCCGAGAGCGATTTAACTGCCTGCGGATCAGAAACCAGCAAGACCACCATACCGTTATC CAGACGTATAGCCTGATACTGGCGGTTATCTTTATCACTTTTACGGATGGTTTCCTGAATCG GCTGCCATCCCGTTTCTGCCTGACTTAAGGGTGCCCAAAGGGCAACTAACAACAATAATG CTTTGAACCAGGTGCTGCGGGGCATTCACGGACCTCATAAGCTTCGCAAATCATCTGCCA GAATTTAATCTTGTGCTGCACGAGTCAGCCTATGTTTATATAACCATCAGTCCGTGACTGGT GCGCATCATAAAGTAAGCGGATAGATTGCGCAATTTTTATACAGCACTCATGACTGATTAA AGCGAAACAGCGGTAACAGGAAACGTTGCGACTGTTCAACGATGGCCTCCATTGTCTCTG GTGTTAATTGCCGCCAGAGCCTTTGATACCAGATATCATCACCTTCGCCACGCACCATCAT GTTGCCTTCGTAAGCCTGAAGGAATTTCGTACGGGCTTTTTGCAACGTGGAATCGTCATCC AGCATGGCATCGTTTTGCGCGTCATAACAGGTTTTTAGCCACGCGCCGCCACTTTCAGGTA ACACCAGCAATGGCGCGGACATTCCTTCACGATACCCCTCAATCAGTTGTGAGAGGTAAT GCAAAGCCTGTTCGGCGTCGACCGGTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTA AGCGCGGC-3′

Table 5: Oligos used for RT-qPCR quantification.

Oligo ID	**Sequence (5′-3′)**	**Gene**
oIK30	TGGCCTGACGCGTATGTTGT	*recC*
oIK31	TGCCCCACCAGTTCTGCAAT	*recC*
oIK32	CTCTTGCGGGTTACGGACGT	*recD*
oIK33	ATTCAGTCCAGCCACGCCAA	*recD*
oIK34	TCTGGCTTCATCGCTCGCAA	*ptrA*
oIK35	TGGGGCATGGGCTTCCTTTT	*ptrA*
oIK62 (recB1_F)	ACCCGCGCATTGGCTGAGAT	*recB*
oIK63 (recB1_R)	CACCGCTTTCGCTACGCAGC	*recB*
oIK64 (rrfD_1_F)	CGGTGGTCCCACCTGACC	*rrfD*
oIK65 (rrfD_1_R)	CCTACTCTCGCATGGGGAGACC	*rrfD*
oIK99 (hfq_F_1)	CGTCTCGCCCGGTTTCTCAT	*hfq*
oIK100 (hfq_R_1)	GGAAGTATTCTGCGCGCTGC	*hfq*
oIK101 (chiX_F_1)	GTCGCTTAAAGTGACGGCAT	*chiX*
oIK102 (chiX_R_1)	TCGCTATTGGCCCGTCAAAGA	*chiX*
oIK105 (cyaR_F_1)	GCTAGCTGTACCAGGAACCACC	*cyaR*
oIK106 (cyaR_R_1)	GGGAGATTACACAGGCTAAGGAGG	*cyaR*

Table 6: Sequences of *recB* RNA FISH probes labelled with TAMRA dye.

Probe ID	**Sequence (5′-3′)**
RecB-TAMRA_1	ctgtaagggcaagcgcaaag
RecB-TAMRA_2	gcaatcgtaaaggttttgcc
RecB-TAMRA_3	tcgtggatattgctacggat
RecB-TAMRA_4	gttcggctaacaacaaccac
RecB-TAMRA_5	aaaggcattcaggttgagca
RecB-TAMRA_6	aatcagctgctgctcaaaca
RecB-TAMRA_7	aaagacgacctgggctattt
RecB-TAMRA_8	cgccttgcagataacgatta
RecB-TAMRA_9	ctgctgttttaccgtatcaa
RecB-TAMRA_10	accagaagattcgatcagcg
RecB-TAMRA_11	cggttaaacttgcgtcgatc
RecB-TAMRA_12	atcttgtcgatccatttagc
RecB-TAMRA_13	ccggcaactgataactgttt
RecB-TAMRA_14	cttcgtgcgatcttctaaga
RecB-TAMRA_15	ttgatcgatcgcctcaaaca
RecB-TAMRA_16	cagatcgcggatcgacaatg
RecB-TAMRA_17	agccgacttaacatgtcatc
RecB-TAMRA_18	cgccaattagcaacaatgcg
RecB-TAMRA_19	tacgcgccttcatataagtg
RecB-TAMRA_20	tgtctaaagtgtagtgggcg
RecB-TAMRA_21	aagcttattcacgctgttca
RecB-TAMRA_22	tcgcgaaacatgaacgcgtc
RecB-TAMRA_23	aaacggaagggatttccagc
RecB-TAMRA_24	cctgcaacaaccaaagcatt
RecB-TAMRA_25	cagatttgccgataaccatc
RecB-TAMRA_26	gttttcagcaatgttacgcg
RecB-TAMRA_27	ttgtagcagttcgctgatat
RecB-TAMRA_28	tggatcgtgacaatctgcac
RecB-TAMRA_29	tggacgcggaaattggtgat
RecB-TAMRA_30	atcgtgataaaacgcctgct
RecB-TAMRA_31	ttaagatccagaactgcctc
RecB-TAMRA_32	aagcaaacgcagatcttccg
RecB-TAMRA_33	aatgccaaaccgaacgtgtc
RecB-TAMRA_34	aacgcttcaatacaggtgcg
RecB-TAMRA_35	ttggttatcaccagtttgtg
RecB-TAMRA_36	tcagctctgctgtagaaaca
RecB-TAMRA_37	aaaccagagtagctggtgac
RecB-TAMRA_38	tgatgtggtgttaacgtcgg
RecB-TAMRA_39	tcaaccggctgggtaaaatc
RecB-TAMRA_40	ttattgcgggcggaaagttg
RecB-TAMRA_41	gataaaactccatctccacc
RecB-TAMRA_42	aacgtatcaagctgactggc
RecB-TAMRA_43	gccagggtataaagctgata
RecB-TAMRA_44	caatgcgatggcgcagataa
RecB-TAMRA_45	tggtgctcatagtcgtaatc
RecB-TAMRA_46	acagataaataacgccgcca
RecB-TAMRA_47	gatgttctttatcaacgcca
RecB-TAMRA_48	cataccggcaaacatctcat

Table 7: Parameters of a two-stage model of RecB expression.

Parameter	**Meaning**	**Value**	**95% CI**	**Source**
k_m	transcription rate	0.21 min^{-1}	[0.13 0.33] min^{-1}	Estimated
γ_m	mRNA degradation rate	0.62 min^{-1}	[0.48 0.75] min^{-1}	Measured
b	mRNA burst size	0.95 molec	[0.76 1.19] molec	Estimated
k_p	translation rate	0.15 min^{-1}	[0.12 0.18] min^{-1}	Estimated
γ_p	protein removal rate	0.015 min^{-1}	[0.011 0.019] min^{-1}	Measured

References

[1] C Wyman and R Kanaar. "DNA double-strand break repair: all's well that ends well". In: *Annu Rev Genet* 40 (2006), pp. 363–83. DOI: 10.1146/annurev.genet.40.110405.090451.

[2] A Mehta and J Haber. "Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair". In: *Cold Spring Harb Perspect Biol* 6.9 (2014). DOI: 10.1101/cshperspect.a016428.

[3] A Sinha et al. "Broken replication forks trigger heritable DNA breaks in the terminus of a circular chromosome". In: *PLOS Genetics* 14.3 (2018), pp. 1–28. DOI: 10.1371/journal.pgen.1007256.

[4] M Spies et al. "RecBCD enzyme switches lead motor subunits in response to chi-recognition". In: *Cell* 131.4 (2007), pp. 694–705. DOI: 10.1016/j.cell.2007.09.023.

[5] J Wiktor et al. "Direct observation of end resection by RecBCD during double-stranded DNA break repair *in vivo*". In: *Nucleic Acids Res* 46.4 (2018), pp. 1821–1833. DOI: 10.1093/nar/gkx1290.

[6] D Arnold and S Kowalczykowski. "Facilitated Loading of RecA Protein is Essential to Recombination by RecBCD Enzyme". In: *J Biol Chem* 275.16 (2000), pp. 12261–12265. DOI: 10.1074/jbc.275.16.12261.

[7] B Michel and D Leach. "Homologous Recombination-Enzymes and Pathways". In: *EcoSal Plus* 5.1 (2012). DOI: 10.1128/ecosalplus.7.2.7.

[8] L Roman, Eggleston A, and S Kowalczykowski. "Processivity of the DNA helicase activity of *Escherichia coli* recBCD enzyme". In: *J Biol Chem* 267.6 (1992), pp. 4207–14.

[9] P Bianco et al. "Processive translocation and DNA unwinding by individual RecBCD enzyme molecules". In: *Nature* 409 (2001), pp. 374–378. DOI: 10.1038/35053131.

[10] B Liu, R Baskin, and S Kowalczykowski. "DNA unwinding heterogeneity by RecBCD results from static molecules able to equilibrate". In: *Nature* 500 (2013), 482–485. DOI: 10.1038/nature12333.

[11] A Taylor et al. "Control of RecBCD enzyme activity by DNA binding- and Chi hotspot-dependent conformational changes". In: *J Mol Biol* 426.21 (2014), pp. 3479–99. DOI: 10.1016/j.jmb.2014.07.017.

[12] CA Cockram et al. "Quantitative genomic analysis of RecA protein binding during DNA double-strand break repair reveals RecBCD action *in vivo*". In: *PNAS* 25.112 (2015), p. 34. DOI: 10.1073/pnas.1424269112.

[13] D Dermic et al. "RecBCD enzyme overproduction impairs DNA repair and homologous recombination in *Escherichia coli*". In: *Res Microbiol* 156.3 (2005), pp. 304–11. DOI: 10.1016/j.resmic.2004.10.005.

[14] P Emmerson. "Recombination deficient mutants of *Escherichia coli* K12 that map between *thyA* and *argA*". In: *Genetics* 60.1 (1968), pp. 19–30. DOI: 10.1093/genetics/60.1.19.

[15] M White et al. "RecBCD coordinates repair of two ends at a DNA double-strand break, preventing aberrant chromosome amplification". In: *Nucleic Acids Res* (June 2018). DOI: 10.1093/nar/gky463.

[16] S Amundsen, A Taylor, and G Smith. "Chi hotspot control of RecBCD helicase-nuclease by long-range intramolecular signaling". In: *Sci Rep* 10.1 (2020), p. 19415. DOI: 10.1038/s41598-020-73078-0.

[17] K Cheng et al. "A conformational switch in response to Chi converts RecBCD from phage destruction to DNA repair". In: *Nat Struct Mol Biol* 27.1 (2020), pp. 71–77. DOI: 10.1038/s41594-019-0355-2.

[18] S Amundsen and G Smith. "The RecB helicase-nuclease tether mediates Chi hotspot control of RecBCD enzyme". In: *NAR* 47.1 (2019), pp. 197–209. DOI: 10.1093/nar/gky1132.

[19] R Zananiri et al. "Synergy between RecBCD subunits is essential for efficient DNA unwinding". In: *Elife* 8 (2019), e40836. DOI: 10.7554/eLife.40836.

[20] M Dillingham and S Kowalczykowski. "RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks". In: *Microbiol Mol Biol Rev* 72.4 (2008), pp. 642–671. DOI: 10.1128/MMBR.00020-08.

[21] I Hickson et al. "Reconstitution of RecBC DNase activity from purified *Escherichia coli* RecB and RecC proteins". In: *J Biol Chem* 260.2 (1985), pp. 1224–29. DOI: 10.1016/S0021-9258(20)71232-9.

[22] A Taylor and G Smith. "Unwinding and Rewinding of DNA by the RecBC Enzyme". In: *Cell* 22 (1980), pp. 447–57. DOI: 10.1016/0092-8674(80)90355-4.

[23] A Lepore et al. "Quantification of very low-abundant proteins in bacteria using the HaloTag and epi-fluorescence microscopy". In: *Sci Rep* 9.7902 (2019). DOI: 10.1038/s41598-019-44278-0.

[24] D Hawley and W McClure. "Compilation and analysis of *Escherichia coli* promoter DNA sequences". In: *Nucl Acids Res* 11 (1983), pp. 2237–55. DOI: 10.1093/nar/11.8.2237.

[25] F Claverie-Martin, M Diaz-Torres, and S Kushner. "Analysis of the regulatory region of the protease III (*ptrA*) gene of *Escherichia coli* K-12". In: *Gene* 54 (1987), pp. 185–95. DOI: 10.1016/0378-1119(87)90486-0.

[26] G Cromie. "Phylogenetic Ubiquity and Shuffling of the Bacterial RecBCD and AddAB Recombination Complexes". In: *J Bacteriol* 191.16 (2009), pp. 5076–84. DOI: 10.1128/JB.00254-09.

[27] D Oppenheim and C Yanofsky. "Translational coupling during expression of the tryptophan operon of *Escherichia coli*". In: *Genetics* 95.4 (1980), pp. 785–95. DOI: 10.1093/genetics/95.4.785.

[28] D Schümperli et al. "Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon". In: *Cell* 30.3 (1982), pp. 865–71. DOI: 10.1016/0092-8674(82)90291-4.

[29] I Hickson et al. "Molecular amplification and purification of the *E. coli recC* gene product". In: *Nucl Acids Res* 12 (1984), pp. 3807–19. DOI: 10.1093/nar/12.9.3807.

[30] J Wade et al. "Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites". In: *Genes Dev* 19.21 (2005), pp. 2619–30. DOI: 10.1101/gad.1355605.

[31] J Paulsson. "Summing up the noise in gene networks". In: *Nature* 427 (2004), pp. 415–8. DOI: 10.1038/nature02257.

[32] S Uphoff et al. "Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation". In: *Science* 351.6277 (2016), pp. 1094–1097. DOI: 10.1126/science.aac9786.

[33] M Elowitz et al. "Stochastic Gene Expression in a Single Cell". In: *Science* 297.5584 (2002), pp. 1183–1186. DOI: 10.1126/science.1070919.

[34] A Raj and A Oudenaarden. "Nature, Nurture, or Chance: Stochastic Gene Expression and its Consequences". In: *Cell* 135.2 (2008), pp. 216–226. DOI: 10.1016/j.cell.2008.09.050.

[35] N Balaban et al. "Bacterial Persistence as a Phenotypic Switch". In: *Science* 305.5690 (2004), pp. 1622–1625. DOI: 10.1126/science.1099390.

[36] A Harms, E Maisonneuve, and K Gerdes. "Mechanisms of bacterial persistence during stress and antibiotic exposure". In: *Science* 354.6318 (2016). DOI: 10.1126/science.aaf4268.

[37] N Christopoulou and S Granneman. "The role of RNA-binding proteins in mediating adaptive responses in Gram-positive bacteria". In: *FEBS J* (2021). DOI: 10.1111/febs.15810.

[38] T Kambara, K Ramsey, and S Dove. "Pervasive Targeting of Nascent Transcripts by Hfq". In: *Cell Rep* 23.5 (2018), pp. 1543–1552. DOI: 10.1016/j.celrep.2018.03.134.

[39] T Nogueira and M Springer. "Post-transcriptional control by global regulators of gene expression in bacteria". In: *Curr Opin Microbiol* 3.2 (2000), pp. 154–8. DOI: 10.1016/s1369-5274(00)00068-0.

[40] Y Chao and J Vogel. "The role of Hfq in bacterial pathogens". In: *Curr Opin Microbiol* 13.1 (2010), pp. 24–33. DOI: 10.1016/j.mib.2010.01.001.

[41] E Van Assche et al. "RNA-binding proteins involved in post-transcriptional regulation in bacteria". In: *Front Microbiol* 6.141 (2015), pp. 1–12. DOI: 10.3389/fmicb.2015.00141.

[42] E Holmqvist and J Vogel. "RNA-binding proteins in bacteria". In: *Nat Rev Microbiol* 16.10 (2018), pp. 601–15. DOI: 10.1038/s41579-018-0049-5.

[43] T Updegrove, A Zhang, and G Storz. "Hfq: the flexible RNA matchmaker". In: *Curr Opin Microbiol* 30 (2016), pp. 133–138. DOI: 10.1016/j.mib.2016.02.003.

[44] K Kavita, F Mets, and S Gottesman. "New aspects of RNA-based regulation by Hfq and its partner sRNAs". In: *Curr Opin Microbiol* 42 (2018), pp. 53–61. DOI: 10.1016/j.mib.2017.10.014.

[45] J Chen and S Gottesman. "Hfq links translation repression to stress-induced mutagenesis in *E. coli*". In: *Genes Dev* 31.13 (2017), pp. 1382–95. DOI: 10.1101/gad.302547.117.

[46] N Sudo et al. "RNA-binding protein Hfq downregulates locus of enterocyte effacement-encoded regulators independent of small regulatory RNA in enterohemorrhagic *Escherichia coli*". In: *Mol. Microbiol.* 00 (2021), pp. 1–16. DOI: 10.1111/mmi.14799.

[47] E Sonnleitner and U Bläsi. "Regulation of Hfq by the RNA CrcZ in *Pseudomonas aeruginosa* Carbon Catabolite Repression". In: *PLOS Genetics* 10.6 (2014). DOI: 10.1371/journal.pgen.1004440.

[48] M Ellis, R Trussler, and D Haniford. "Hfq binds directly to the ribosome-binding site of IS10 transposase mRNA to inhibit translation". In: *Mol Microbiol* 96.3 (2015), pp. 633–50. DOI: 10.1111/mmi.12961.

[49] S Skinner et al. "Measuring mRNA copy-number in individual *Escherichia coli* cells using single-molecule fluorescent *in situ* hybridization (smFISH)." In: *Nat Protoc* 8.6 (2013), pp. 1100–1113. DOI: 10.1038/nprot.2013.066.

[50] H Chen et al. "Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*". In: *Mol. Syst. Biol.* 11.781 (2015). DOI: 10.15252/msb.20145794.

[51] J Moffitt et al. "Spatial organization shapes the turnover of a bacterial transcriptome". In: *eLife* 5 (2016). DOI: 10.7554/eLife.13065.

[52] M Wang et al. "Measuring transcription at a single gene copy reveals hidden drivers of bacterial individuality". In: *Nat. Microbiol* 4 (2019), 2118–2127. DOI: 10.1038/s41564-019-0553-z.

[53] J Peterson et al. "Effects of DNA replication on mRNA noise". In: *PNAS* 112.52 (2015), pp. 15886–15891. DOI: 10.1073/pnas.1516246112.

[54] J Bryant et al. "Chromosome position effect on gene expression in *Escherichia coli* K12". In: *NAR* 42.18 (2014), pp. 11383–92. DOI: 10.1093/nar/gku828.

[55] J Peccoud and B Ycart. "Markovian Modeling of Gene-Product Synthesis". In: *Theor Popul Biol* 48.2 (1995), pp. 222–234. DOI: 10.1006/tpbi.1995.1027.

[56] A Raj et al. "Stochastic mRNA synthesis in Mammalian Cells". In: *PLOS Biol* 4.10 (2006), e309. DOI: 10.1371/journal.pbio.0040309.

[57] V Shahrezaei and P Swain. "Analytical distributions for stochastic gene expression". In: *PNAS* 105.45 (2008), pp. 17256–17261. DOI: 10.1073/pnas.0803850105.

[58] S Proshkin et al. "Cooperation between translating ribosomes and RNA polymerase in transcription elongation". In: *Science* 328.5977 (2010), pp. 504–8. DOI: 10.1126/science.1184939.

[59] H Bremer and P Dennis. "Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates". In: *EcoSal Plus* 3.1 (2008). DOI: 10.1128/ecosal.5.2.3.

[60] U Vogel and K Jensen. "The RNA chain elongation rate in *Escherichia coli* depends on the growth rate". In: *J Bacteriol.* 176.10 (1994), pp. 2807–13. DOI: 10.1128/jb.176.10.2807-2813.1994.

[61] A Sanchez, S Choubey, and J Kondev. "Stochastic models of transcription: From single molecules to single cells". In: *Methods* 62.1 (2013), pp. 13 –25. DOI: 10.1016/j.ymeth.2013.03.026.

[62] P Swain. "Efficient attenuation of stochasticity in gene expression through post-transcriptional control". In: *J Mol Biol* 344.4 (2004), pp. 965–76. DOI: 10.1016/j.jmb.2004.09.073.

[63] M Eldeeb et al. "A molecular toolbox for studying protein degradation in mammalian cells". In: *J Neurochem* 151.4 (2019), pp. 520–533. DOI: 10.1111/jnc.14838.

[64] J Yewdell et al. "Out with the old, in with the new? Comparing methods for measuring protein degradation". In: *Cell Biol Int* 35.5 (2011), pp. 457–62. DOI: 10.1042/CBI20110055.

[65] M Doherty et al. "Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC". In: *J Proteome Res* 8.1 (2009), pp. 104–12. DOI: 10.1021/pr800641v.

[66] R Merrill et al. "A robust and economical pulse-chase protocol to measure the turnover of HaloTag fusion proteins". In: *J Biol Chem* 294.44 (2019), pp. 16164–16171. DOI: 10.1074/jbc.RA119.010596.

[67] K Yamaguchi et al. "Pulse-chase experiment for the analysis of protein stability in cultured mammalian cells by covalent fluorescent labeling of fusion proteins". In: *Methods Mol Biol* 577 (2009), pp. 121–31. DOI: 10.1007/978-1-60761-232-2_10.

[68] M Thattai and A van Oudenaarden. "Intrinsic noise in gene regulatory networks". In: *PNAS* 98.15 (2001), pp. 8614–9. DOI: 10.1073/pnas.151588598.

[69] Johan Paulsson. "Models of stochastic gene expression". In: *Phys Life Rev* 2.2 (2005), pp. 157 –175. DOI: 10.1016/j.plrev.2005.03.003.

[70] D Gillespie. "Exact stochastic simulation of coupled chemical reactions". In: *J. Phys. Chem.* 81.25 (1977), 2340–2361. DOI: 10.1021/j100540a008.

[71] JM Pedraza and J Paulsson. "Effects of molecular memory and bursting on fluctuations in gene expression". In: *Science* 319.5861 (2008), pp. 339–43. DOI: 10.1126/science.1144331.

[72] D Hooper and G Jacoby. "Topoisomerase Inhibitors: Fluoroquinolone Mechanisms of Action and Resistance". In: *Cold Spring Harb Perspect Med* 6.9 (2016), a025320. DOI: 10.1101/cshperspect.a025320.

[73] J Tree et al. "Identification of Bacteriophage-Encoded Anti-sRNAs in Pathogenic *Escherichia coli*". In: *Mol Cell* 55.2 (2014), pp. 199–213. DOI: 10.1016/j.molcel.2014.05.006.

[74] A Zhang et al. "Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets". In: *J Mol Biol* 425.19 (2013), pp. 3678–97. DOI: 10.1016/j.jmb.2013.01.006.

[75] E Holmqvist et al. "Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo". In: *EMBO J* 35.9 (2016), pp. 991–1011. DOI: 10.15252/embj.201593360.

[76] B Barreto et al. "The Small RNA GcvB Promotes Mutagenic Break Repair by Opposing the Membrane Stress Response". In: *J Bacteriol* 198.24 (2016), pp. 3296–308. DOI: 10.1128/JB.00555-16.

[77] J Pribis et al. "Gamblers: An Antibiotic-Induced Evolvable Cell Subpopulation Differentiated by Reactive-Oxygen-Induced General Stress Response". In: *Mol Cell* 74.4 (2019), pp. 785–800. DOI: 10.1016/j.molcel.2019.02.037.

[78] I Iosub et al. "Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability". In: *eLife* 9 (2020). DOI: 10.7554/eLife.54655.

[79] T Link, P Valentin-Hansen, and R Brennan. "Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA". In: *PNAS* 106.46 (2009), pp. 19292–19297. DOI: 10.1073/pnas.0908744106.

[80] H Tsui, H Leung, and M Winkler. "Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12". In: *Mol Microbiol* 13.1 (1994), pp. 35–49. DOI: 10.1111/j.1365-2958.1994.tb00400.x.

[81] A Muffler et al. "The RNA-binding protein HF-I plays a global regulatory role which is largely, but not exclusively, due to its role in expression of the sigmaS subunit of RNA polymerase in *Escherichia coli*". In: *J Bacteriol* 179.1 (1997), pp. 297–300. DOI: 10.1128/jb.179.1.297-300.1997.

[82] T Møller et al. "Hfq: A Bacterial Sm-like Protein that Mediates RNA-RNA Interaction". In: *Mol Cell* 9.1 (2002), pp. 23–30. DOI: 10.1016/S1097-2765(01)00436-1.

[83] T Morita and H Aiba. "Mechanism and physiological significance of autoregulation of the *Escherichia coli hfq* gene". In: *RNA* 25.2 (2019), pp. 264–76. DOI: 10.1261/rna.068106.118.

[84] R Faigenbaum-Romm et al. "Hierarchy in Hfq Chaperon Occupancy of Small RNA Targets Plays a Major Role in Their Regulation". In: *Cell Rep* 30.9 (2020), pp. 3127–3138. DOI: 10.1016/j.celrep.2020.02.016.

[85] K Moon and S Gottesman. "Competition among Hfq-binding small RNAs in *Escherichia coli*". In: *Mol Microbiol* 82.6 (2011), pp. 1545–62. DOI: 10.1111/j.1365-2958.2011.07907.x.

[86] E Małecka et al. "Structure of bacterial regulatory RNAs determines their performance in competition for the chaperone protein Hfq". In: *Biochemistry* 54.5 (2015), pp. 1157–70. DOI: 10.1021/bi500741d.

[87] P Mandin and S Gottesman. "Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA". In: *EMBO J* 29.18 (2010), pp. 3094–107. DOI: 10.1038/emboj.2010.179.

[88] F Chen et al. "5′UTR sequences influence protein levels in *Escherichia coli* by regulating translation initiation and mRNA stability". In: *Front Microbiol* 13 (2022). DOI: 10.3389/fmicb.2022.1088941.

[89] S Emory, P Bouvet, and J Belasco. "A 5′-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*". In: *Genes Dev* 6 (1992), pp. 135–148. DOI: 10.1101/gad.6.1.135.

[90] H Salvail et al. "Antagonistic functions between the RNA chaperone Hfq and an sRNA regulate sensitivity to the antibiotic colicin". In: *EMBO J.* 32.20 (2013), pp. 2764–78. DOI: 10.1038/emboj.2013.205.

[91] S Melamed et al. "Global mapping of small RNA-target interactions in bacteria". In: *Mol Cell* 63 (5 2016), pp. 884–897. DOI: 10.1016/j.molcel.2016.07.026.

[92] P Mandin and S Gottesman. "A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system". In: *Mol Microbiol* 72.3 (2009), pp. 551–65. DOI: 10.1111/j.1365-2958.2009.06665.

[93] S Modi et al. "Functional characterization of bacterial sRNAs using a network biology approach". In: *PNAS* 108.37 (2011), pp. 15522–15527. DOI: 10.1073/pnas.1104318108.

[94] W Li et al. "Systematically integrated metabonomic-proteomic studies of *Escherichia coli* under ciprofloxacin stress". In: *J Proteomics* 179 (2018), pp. 61–70. DOI: 10.1016/j.jprot.2018.03.002.

[95] E Guisbert et al. "Hfq Modulates the E-Mediated Envelope Stress Response and the 32-Mediated Cytoplasmic Stress Response in *Escherichia coli*". In: *J. Bacteriol.* 189.5 (2007), pp. 1963–1973. DOI: 10.1128/jb.01243-06.

[96] RF Dos Santos et al. "Hfq and RNase R Mediate rRNA Processing and Degradation in a Novel RNA Quality Control Process". In: *mBio.* 11.5 (2020), e02398–20. DOI: 10.1128/mBio.02398-20.

[97] B Vecerek et al. "Interaction of the RNA chaperone Hfq with mRNAs: direct and indirect roles of Hfq in iron metabolism of *Escherichia coli*". In: *Mol Microbiol* 50.3 (2003), pp. 897–909. DOI: 10.1046/j.1365-2958.2003.03727.x.

[98] C Dykstra, D Prasher, and S Kushner. "Physical and biochemical analysis of the cloned *recB* and *recC* genes of *Escherichia coli* K-12". In: *J Bacteriol* 157 (1984), pp. 21–7. DOI: 10.1128/jb.157.1.21-27.1984.

[99] Y Shimoni et al. "Regulation of gene expression by small non-coding RNAs: a quantitative view". In: *Mol Syst Biol* 3.138 (2007). DOI: 10.1038/msb4100181.

[100] A Singh. "Negative feedback through mRNA provides the best control of gene-expression noise". In: *IEEE Trans Nanobioscience* 10.3 (2011), pp. 194–200. DOI: 10.1109/TNB.2011.2168826.

[101] P Bokes, M Hojcka, and A Singh. "MicroRNA Based Feedforward Control of Intrinsic Gene Expression Noise". In: *IEEE* 18.1 (2021), pp. 272–282. DOI: 10.1109/TCBB.2019.2938502.

[102] P Mehta, S Goyal, and NS Wingreen. "A quantitative comparison of sRNA-based and protein-based gene regulation". In: *Mol Syst Biol* 4.221 (2008). DOI: 10.1038/msb.2008.58.

[103] V Siciliano et al. "miRNAs confer phenotypic robustness to gene networks by suppressing biological noise". In: *Nat Commun* 4.2364 (2013). DOI: 10.1038/ncomms3364.

[104] C Kelly et al. "Synthetic negative feedback circuits using engineered small RNAs". In: *NAR* 46.18 (2018), 9875–9889. DOI: 10.1093/nar/gky828.

[105] M Mundt et al. "A System for Gene Expression Noise Control in Yeast". In: *ACS Synth Biol.* 7.11 (2018), pp. 2618–2626. DOI: 10.1021/acssynbio.8b00279.

[106] J Schmiedel et al. "MicroRNA control of protein expression noise". In: *Science* 348.6230 (2015), pp. 128–132. DOI: 10.1126/science.aaa1738.

[107] A Link, D Phillips, and G Church. "Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization". In: *J Bacteriol* 179.20 (1997), pp. 6228–37. DOI: 10.1128/jb.179.20.6228-6237.1997.

[108] E Darmon et al. "*E. coli* SbcCD and RecA Control Chromosomal Rearrangement Induced by an Interrupted Palindrome". In: *Mol Cell* 39.1-2 (2010), pp. 59–70. DOI: 10.1016/j.molcel.2010.06.011.

[109] S Jaramillo-Riveri et al. "Growth-dependent heterogeneity in the DNA damage response in *Escherichia coli*". In: *Mol Syst Biol.* 18.5 (2022). DOI: 10.15252/msb.202110441.

[110] I Iosub et al. "The mRNA derived MalH sRNA contributes to alternative carbon source utilization by tuning maltoporin expression in *E. coli*". In: *RNA Biology* 18.6 (2021), pp. 914–31. DOI: 10.1080/15476286.2020.1827784.

[111] URL: https://git.ecdf.ed.ac.uk/sgrannem.

[112] URL: https://pypi.org/user/g_ronimo/.

[113] URL: https://www.mathworks.com/matlabcentral/fileexchange/20688-kullback-leibler-divergence.

[114] URL: https://github.com/nvictus/Gillespie.

[115] N Friedman, L Cai, and XS Xie. "Linking stochastic dynamics to population distribution: an analytical framework of gene expression". In: *Phys Rev Lett* 97.16 (2006), p. 168302. DOI: 10.1103/PhysRevLett.97.168302.

[116] N Kumar, A Singh, and RV Kulkarni. "Transcriptional Bursting in Gene Expression: Analytical Results for General Stochastic Models". In: *PLoS Comput. Biol.* 11.10 (2015). DOI: 10.1371/journal.pcbi.1004292.

[117] S Smith and R Grima. "Single-cell variability in multicellular organisms". In: *Nat Commun* 9.345 (2018). DOI: 10.1038/s41467-017-02710-x.

[118] NG Van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, 1992.

[119] I Golding et al. "Real-Time Kinetics of Gene Activity in Individual Bacteria". In: *Cell* 123.6 (2005), pp. 1025–36. DOI: 10.1016/j.cell.2005.09.031.

[120] L So et al. "General properties of transcriptional time series in *Escherichia coli*". In: *Nat Genet* 43 (2011), 554–60. DOI: 10.1038/ng.821.

[121] R Perez-Carrasco, C Beentjes, and R Grima. "Effects of cell cycle variability on lineage and population measurements of messenger RNA abundance". In: *J. R. Soc. Interface* 17.168 (2020). DOI: 10.1098/rsif.2020.0360.

[122] A Klindziuk, B Meadowcroft, and A Kolomeisky. "A Mechanochemical Model of Transcriptional Bursting". In: *Biophys J* 118.5 (2020), pp. 1213–20. DOI: 10.1016/j.bpj.2020.01.017.

[123] D Jones and J Elf. "Bursting onto the scene? Exploring stochastic mRNA production in bacteria". In: *Curr Opin Microbiol* 45 (2018), pp. 124–130. DOI: 10.1016/j.mib.2018.04.001.

[124] D Selinger et al. "Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation". In: *Genome Res* 13.2 (2003), pp. 216–23. DOI: 10.1101/gr.912603.

[125] R Mosteller, J Rose, and C Yanofsky. "Transcription Initiation and Degradation of *trp* mRNA". In: *Cold Spring Harb Symp Quant Biol* 35 (1970), pp. 461–6. DOI: 10.1101/SQB.1970.035.01.060.

[126] T Schwartz, E Craig, and D Kennell. "Inactivation and degradation of messenger ribonucleic acid from the lactose operon of *Escherichia coli*". In: *JMB* 54.2 (1970), pp. 299–311. DOI: 10.1016/0022-2836(70)90431-6.

[127] Y Taniguchi et al. "Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells". In: *Science* 329.5991 (2010), pp. 533–538. DOI: 10.1126/science.1188308.

[128] AL Koch and HR Levy. "Protein turnover in growing cultures of *Escherichia coli*". In: *J Biol Chem* 217.2 (1955), pp. 947–57.

[129] J Elf and M Ehrenberg. "Fast evaluation of fluctuations in biochemical networks with the linear noise approximation". In: *Genome Res.* 13.11 (2003), pp. 2475–2484. DOI: 10.1101/gr.1196503.

[130] D Schnoerr, G Sanguinetti, and R Grima. "Approximation and inference methods for stochastic biochemical kinetics — a tutorial review". In: *J. Phys. A: Math. Theor.* 50.9 (2017). DOI: 10.1088/1751-8121/aa54d9.

[131] P Wang et al. "Robust growth of *Escherichia coli*". In: *Curr Biol.* 22.12 (2010), pp. 1099–103. DOI: 10.1016/j.cub.2010.04.045.

[132] M Mir et al. "Optical measurement of cycle-dependent cell growth". In: *PNAS* 108.32 (2011), pp. 13124–9. DOI: 10.1073/pnas.1100506108.

[133] A Ponticelli et al. "Chi-dependent DNA strand cleavage by RecBC enzyme". In: *Cell* 41.1 (1985), pp. 145–51. DOI: 10.1016/0092-8674(85)90069-8.

[134] L Guzman et al. "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter". In: *J Bacteriol* 177.14 (1995), pp. 4121–30. DOI: 10.1128/jb.177.14.4121-4130.1995.