Fine-tuning of RecBCD expression by post-transcriptional regulation is required for optimal DNA repair in *Escherichia coli*

Irina Kalita[1,2], Ira Alexandra Iosub[1,3,4], Sander Granneman[1,3], and Meriem El Karoui[1,2,*]

[1]*Centre for Synthetic and Systems Biology (SynthSys), University of Edinburgh, UK*
[2]*Institute of Cell Biology, University of Edinburgh, UK*
[3]*Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh, UK*
[4]*The Francis Crick Institute, London, UK*
[*]For correspondence: meriem.elkaroui@ed.ac.uk

Abstract

To preserve genome integrity, all living organisms have developed strategies to respond to chromosomal damage. One such response is the repair of DNA double-strand breaks (DSBs), one of the most toxic forms of DNA lesions. In *E. coli*, DSBs are repaired via the homologous recombination pathway, initiated by the RecBCD enzyme. 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 number 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 levels of fluctuations yet cell-to-cell protein variability remains remarkably low. We show that the post-transcriptional regulator Hfq binds to *recBCD* mRNAs and down-regulates RecB protein expression *in vivo*. Furthermore, when Hfq-mediated regulation is perturbed, we observe less effective noise reduction and reduced DNA repair capacity. Taken together, our results suggest a post-transcriptional regulatory mechanism where Hfq fine-tunes RecB expression by inhibiting RecB translation. This fine-tuning of RecB expression contributes to reducing noise in RecB protein expression and protects cells against the toxic consequences of too high RecBCD numbers.

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), which are 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 [124]. In addition to various exogenous sources such as ionizing radiation, mechanical stresses and DNA-damaging agents, DSBs can also occur during the normal cell cycle due to endogenous metabolic reactions and replication events [63]. For instance, replication fork breakage has been shown to result in spontaneous DSBs in 18% of wild-type *Escherichia coli* cells in each generation [100] 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 its 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 [105, 123]. This leads to the formation of a RecA-ssDNA nucleoprotein filament [8], which catalyses homology search and strand exchange [67].

Recent *in vitro* and *in vivo* experiments have highlighted the extraordinary processive DNA degradation activity of RecBCD upon recognition of a dsDNA end [89, 11, 59]. Indeed, RecBCD has been shown to process the chromosome for up to $\sim$100 *kb* at $\sim$1.6 *kb*/*s* in live bacteria [123]. Such a powerful DNA degradation activity is controlled by *chi* recognition [110]. However, because this recognition is probabilistic [21], RecBCD can degrade large fragments of the chromosome before DNA repair is initiated. As a result, overproduction of RecBCD has been shown to sensitise wild-type cells to ultraviolet and gamma irradiations [25]. This strongly supports the idea that an excess

of RecBCD imposes a threat for cells upon DNA damage. In contrast, RecBCD expression is essential upon DSB induction [33, 122] and hence too low levels might impair DNA repair. Thus, the observation that RecBCD can have both positive and negative impacts on 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 ([7, 18, 6, 127, 110], for review see [26, 67]), less is known about the regulation of its expression. RecBCD has been reported to be expressed at very low levels [42, 109] and is present in less than 10 molecules per cell [55]. 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 [40, 20] (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* [22]. As sometimes found in bacterial operons [77, 94], 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 [41, 42] and the *recBCD* genes were shown not to belong to the SOS-response regulon [119]. 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, which is known to often lead to high stochastic fluctuations, could result in large cell-to-cell variability [78, 115, 32]. Although these fluctuations might not play a significant role during normal cell growth, they may become crucial and determine cell fate upon stress conditions [87, 9, 39]. Thus, the low abundance of RecBCD enzyme molecules raises the following questions: (i) to what extent 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 its regulation. 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 [19, 50]. 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 [76, 15, 116, 43]). In the most prevalent case, Hfq facilitates base-paring between small non-coding RNAs (sRNAs) and target mRNAs what often results in mRNA translation and/or stability modulation [114]. There is also increasing evidence that Hfq can control mRNA translation directly [51, 17]. This mechanism is typically based on altering the accessibility of the RBS of a target mRNA by direct binding near this region [106, 104, 31].

Here we show that the expression of RecBCD is regulated post-transcriptionally by Hfq and we demonstrate that Hfq plays a role in reducing fluctuations in RecB copy numbers in *Escherichia coli*. We chose to focus on the RecB subunit as (i) it is an essential component for the complex formation with both helicase and nuclease activity and (ii) fluctuations in RecB numbers are transmitted to the noise of the complex abundance. By quantifying RecB mRNAs and proteins with single-cell resolution, we found that constitutive transcription leads to low and noisy *recB* mRNA abundance. In contrast, fluctuations in RecB protein levels are significantly lower than might be expected given the very low number of proteins present in cells. We show that Hfq negatively regulates RecB protein production and that Hfq plays an important role in reducing RecB protein fluctuations, possibly by inhibiting translation initiation of RecB. We propose that Hfq controls the optimal level of RecBCD molecules and protects bacterial cells against the toxic consequences of too high RecBCD copy numbers.

Results

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

To quantitatively measure RecB expression, we first precisely quantified *recB* mRNAs using single-molecule fluorescent *in situ* hybridisation (smFISH) [101]. We confirmed that our hybridisation 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 control that this result was not due to 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 smFISH 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 in average ∼2.5 *min* across all mRNA species [16, 70].

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 [120, 83, 14]. 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 due to 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 [120].

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 (Supp Info: section 1). The expected probability distribution for mRNA numbers at the steady-state is given by a negative binomial [80, 88, 97]. 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 6, Methods). The latter is remarkably low and close to the lower bound of the reported range of transcription rates in *E. coli*, 0.17-1.7 min^{-1} [86, 13, 118]. Therefore, the low abundance of *recB* mRNAs can be explained by infrequent transcription combined with fast mRNA decay.

Variations in mRNA molecule numbers, due to their low abundance and short lifetimes, can significantly contribute to gene expression noise [91, 78, 107]. To evaluate fluctuations of *recB* transcription, we used 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 a 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 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 individual cells using the Halo self-labelling tag [55] (fig. 2A,B). As we previously reported, the translational RecB-HaloTag fusion does not affect the activity of the RecBCD enzyme [55]. 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 due to cell growth and division, we aimed to estimate RecB protein removal rate. This is usually obtained from either pulse-chase experiments using radio-labelled aminoacids or detection of protein concentration decay after treatment with a translation inhibitor [29, 126, 27]. While the first approach might not be sensitive enough to accurately detect the changes of a low abundant protein, the latter one may affect the measurements on long time-scales because of the physiological impacts of stopping all protein production. We therefore used an alternative approach [66, 125] where we pulse-labelled RecB molecules with the HaloTag ligand, then removed excess of the dye by extensive washing and measured the abundance of the labelled molecules over time. Assuming protein removal to 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]) (fig. 2D). Comparison to the population growth rate in these growth conditions (0.017 min^{-1}) suggests that RecB protein is stable and effectively degraded only due to dilution and molecule partitioning between daughter cells.

RecB protein fluctuations are lower than expected

We then extended the transcription model described above by taking into account protein synthesis and removal. This is known as the *two-stage* or *mRNA-protein* model of gene expression [111, 79, 97]. 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 (Supp Info: section 1). As earlier discussed, we focused on newborn cells and estimated the rate of translation as 0.15 min^{-1} (table 6, Methods). This value fell below the lower bound of the reported translation rates in *E. coli* (0.25-1 min^{-1}) [86, 13]. Therefore, we concluded that *recB* mRNAs are translated at low levels compared to other proteins in a cell.

We next sought to examine whether the constructed 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) [36] for the two-stage model with the estimated parameters (table 6). Strikingly, the simulated distribution had larger fluctuations than the experimentally measured one (fig. 2E). While both protein means were equal due to the first-moment-based parameter estimation (Methods), we found the standard deviations to be different (fig. 2E): 2.2 (simulations) and 1.5 (experiment). In other words, a two-stage model of protein production predicts a higher level of fluctuations than that detected experimentally.

This phenomenon can be also 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 two-stage gene expression model with burst transcription, it has been established that the coefficient of variation is given by the following expression (Supp Info: section 1) [111, 81, 79, 97]:

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

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 a good agreement. This deviation from the theoretical prediction implies a potential regulatory mechanism of RecB expression that actively suppresses variation in protein production in order to maintain protein level within a certain range.

RecB transcription and protein production are decoupled upon DSB induction

Next we addressed the question of whether there is any further evidence of this regulation once bacterial cells are exposed to stress. We hypothesised that RecB mRNA and/or protein levels may exhibit a concentration change upon DNA damage in order to enhance 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 [45]. After a two-hour exposure to ciprofloxacin, we quantified proteins by Halo-labelling and mRNAs by smFISH (fig. 3A). We used a deterministic analysis to show that this induction duration is sufficient to detect changes in mRNA and protein levels (Supp Info: section 2). 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 number normalised by cell area) rather than absolute numbers. It is worth noting that 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. We observed that RecB protein concentration does 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 (fig. 3E: insert). 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 is ∼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, *recB* gene is not included within the SOS regulon [119] and hence not necessarily expected to be upregulated upon DNA damage. However, downregulation of *recB* transcription is an unforeseen observation and may be due to 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 ± 5.3 whereas it is 8.1 ± 2.0 proteins/mRNA in intact cells (fig. 3F). Thus, a $\sim$2.4 increase in translational efficiency was detected upon DSB induction. Taken together, our results suggest a post-transcriptional regulatory mechanism that maintains a constant RecB protein level in response to the fluctuations in mRNA production.

Hfq regulates RecB expression

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 [112, 128, 44] and (ii) its activity has been directly linked to the DNA damage and SOS responses [17, 10, 85].

To assess whether Hfq is involved in post-transcriptional regulation of RecBCD, we analysed available transcriptome-wide datasets wherein Hfq binding sites were identified in *E. coli* using CLASH (cross-linking and analysis of cDNA) [46]. During CLASH, Hfq-RNA complexes are cross-linked *in vivo* and purified under highly stringent conditions, ensuring specific recovery of Hfq RNA targets [46]. Analysis of the Hfq binding profile across *recBCD* transcripts showed that Hfq directly cross-links to the *recBCD* mRNAs (fig. S6a). Hfq binding mainly localised at two sites: one in close proximity to the RBS of the *ptrA* gene, and the second one further downstream in the coding sequence of *recB* gene. Interestingly, multiple trinucleotide Hfq-binding motifs, A-R(A/G)-N(any nucleotide) [112, 58] were found $\sim$8 nucleotides upstream of the RBS of the *ptrA* sequence (fig. S6b) suggesting that Hfq may control translation initiation and/or decay of the operon mRNA. Because of the potential translational coupling between *ptrA* and *recB*, this may directly impact RecB protein production. Alternatively, the other Hfq binding site might also affect *recB* translation. Taken together, these results led us to hypothesise that Hfq may control RecB expression *in vivo*.

To test this hypothesis, we quantified RecB protein and mRNA levels in *E. coli* strains lacking Hfq. *hfq* mutants have been reported to have an increased average cell size and various growth defects [113, 74, 75]. Indeed, we confirmed increased cell size of *hfq* mutants compared to the wild-type cells (fig. S7a,b). Consistent with the previous study [113], we observed a longer lag phase for the *hfq* mutants but no difference in exponential growth rates in our conditions (fig. 4C, fig. S7d).

We quantified *recB* mRNA molecules in the Δ*hfq* strain using smFISH and analysed the mRNA concentration distributions (fig. 4A,B). Although average *recB* mRNA concentration is slightly decreased in *hfq* mutants: 0.21 $mol/\mu m^2$ ($CI_{95\%}$: [0.08, 0.30]) compared to 0.28 $mol/\mu m^2$ ($CI_{95\%}$: [0.23, 0.32]) in the wild-type cells, this difference is not significant (fig. 4B: insert). 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.

In contrast to the transcriptional level, we observed an increase of RecB protein concentration in the Δ*hfq* cells (fig. 4D,E). *hfq* mutants showed more fluorescent foci relative to the wild-type cells (fig. 4D), and even after normalisation by cell area (to take into account the larger cell size in Δ*hfq* cells), RecB concentration distribution was shifted towards higher values (fig. 4E). Quantification across five replicated experiments showed a significant 1.3-fold increase in RecB protein concentration in the Δ*hfq* strain (fig. 4E: insert): 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. Altogether, these results suggest that Hfq post-transcriptionally controls RecB expression by inhibiting its translation. Based on the localisation of the Hfq-binding uncovered in the CRAC data, we further propose a mechanism of Hfq-dependent modulation of translational initiation of RecB (fig. 5A) (similarly to the models of Hfq regulation suggested recently [31, 17, 106]).

To test whether this increase in RecB concentration has an impact on the capacity of cells to perform DNA repair, we exposed wild type and Δ*hfq* strains to sub-lethal levels of ciprofloxacin and measured their viability in a serial dilution assay (fig. 4F). The wild-type strain is fully viable upon exposure to 4 ng/ml of ciprofloxacin, whilst the viability of the Δ*recB* mutant is drastically reduced. Importantly, the *hfq* mutants were more sensitive to ciprofloxacin than the wild-type strain suggesting that the increase in RecB protein levels in the absence of Hfq may negatively impact the capacity of cells to repair DSBs.

RecB distribution is partially recovered by Hfq complementation

To confirm that Hfq directly regulates RecB expression, we complemented the *hfq* mutant with a multi-copy plasmid that expresses Hfq, pQE-Hfq [72]. After complementation, the RecB protein concentration distribution was shifted towards lower values (fig. 5B,D) whilst no change of RecB expression was detected in the Δ*hfq* cells harboring the backbone plasmid, pQE80L (fig. S8). Interestingly, RecB protein level was not fully recovered to the wild-type level (fig. 5B). This is consistent with partial recovery of *hfq* mRNA level measured

by RT-qPCR ($\sim$ a third of the *hfq* mRNA level in wild type, fig. S8). Additionally, we did not detect significant changes in the *recB* mRNA levels when expressing Hfq from the plasmid (fig. S8). Therefore, lower level of Hfq expression is expected to lead to less effective control of RecB translation and hence an intermediate level of RecB protein concentration (fig. 5A).

Sequestering Hfq leads to an increase in RecB production

To further ascertain the role of Hfq in controlling RecB expression whilst avoiding potential additional effects due to the pleiotropic phenotype of *hfq* mutants, we used another strategy to sequester Hfq proteins away from its *recB* mRNA targets. Hfq is known to bind $\sim$100 sRNAs [44, 112], some of which can reach high expression levels and/or have higher affinity for Hfq and thus compete for Hfq binding [34, 71]. We reasoned that because *recB* mRNA is present in very low abundance, it could be out-competed for Hfq binding by over-expression of a small RNA that efficiently binds Hfq, such as ChiX [62, 71]. To test this, 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 intact cells, this overproduction is very likely to sequester a large number of Hfq molecules [31, 61, 106].

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,D). Indeed, RecB concentration in the cells where ChiX was over-expressed almost fully overlapped with the obtained distributions in *hfq* mutants. This strongly suggests that fewer Hfq molecules are available for RecB regulation when ChiX is overproduced (fig. 5C). A similar experiment with a plasmid that did not carry ChiX did not lead to any change in RecB protein concentration thus showing that ChiX over-expression specifically impairs RecB post-transcriptional regulation (fig. S9).

In order to exclude the possibility of a direct ChiX involvement in RecB regulation, we characterised RecB expression in *chiX* mutants. No significant change relative to wild-type cells was detected in RecB mRNA or protein levels (fig. S10). Thus, we concluded that ChiX over-production indirectly affects RecB translation. Consistently with the suggested mechanism of RecB regulation, titration of Hfq proteins from a shared pool by ChiX sRNAs leads to the increased availability of *ptrA-recB* RBS for ribosomes and consequently results in a more efficient RecB translation (fig. 5C: insert).

We further tested whether ChiX over-expression has an effect on the efficiency of DNA repair with a serial dilution assay where DSBs were induced with 8-10 ng/ml of ciprofloxacin (fig. 5E: upper panel). While ChiX overproduction does not affect cell size as much as the *hfq* deletion, it sensitises cells to DNA DSB damage compared to wild type. That implies that the increased RecB protein production decreases the ability to faithfully perform DNA repair. Strikingly, we observed a similar decrease in DNA repair efficiency when RecBCD is over-expressed from a plasmid encoding the entire *recBCD* locus (fig. 5E: lower panel). This observation supports the idea that increased number of RecBCD molecules is toxic when cells are exposed to DNA damage.

Finally, 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 *hfq* mutants and under ChiX over-expression. To this end, we computed the ratio between analytically predicted noise (CV^2_{Theory}) and experimentally measured (CV^2_{Exp}). As no significant change was detected in *recB* mRNA levels among the samples (fig. 4B, fig. S9), we assumed that mRNA contribution is equal for all three conditions and estimated theoretical values of protein noise via the average protein abundance as $1/\langle p \rangle$ (eq. (1)). We observed a less effective noise reduction in Δhfq and ChiX^{++} in comparison to wild type (fig. 5F) suggesting that post-transcriptional regulation by Hfq decreases fluctuations in RecB protein numbers.

Discussion

Despite its essential role in DSB repair, over-expression of RecBCD decreases cell viability upon DNA damage induction (fig. 5E, [25]). Such a dual impact on cell fitness suggests the existence of an underlying regulation that tightly controls RecBCD copy numbers within an optimal range. Using state-of-the-art single-molecule mRNA and protein quantification we show that, despite a large cell-to-cell variability in *recB* mRNA abundance, RecB protein levels show lower fluctuations than expected and remain remarkably constant upon DNA damage (fig.1-3). We show that this is due, at least in part, to post-transcriptional downregulation of RecB expression by Hfq *in vivo* (fig.4-5). To our knowledge, this provides the first

experimental evidence of an RNA-binding protein involved in noise suppression in bacterial cells (fig. 5F). Finally, we demonstrate a reduced capacity of cells to repair DNA when the Hfq-mediated regulation is perturbed (fig. 4F, fig. 5E). We thus speculate that the inhibition of RecB translation by Hfq protects bacteria against the toxic consequences of too high RecBCD molecule numbers.

To further investigate the uncovered post-transcriptional regulation mechanism, we asked whether a small RNA (sRNA) could be involved in Hfq-mediated regulation of RecB expression. We first tested Hfq mutants containing point mutations in the rim, distal and proximal surfaces of the protein, that were shown to impact the interaction of Hfq with different RNA classes [128]. A moderate but non-significant effect on RecB protein distributions was detected in all three cases (fig. S11). While these results provide an additional confirmation of Hfq involvement in the regulation of RecB translation, they do not strongly support a sRNA-dependent mechanism. This is consistent with recent studies demonstrating examples of direct regulation by Hfq in a sRNA-independent manner [106, 17, 104, 31, 90]. We nevertheless looked for potential candidates of sRNAs, using the available Hfq RIL-seq and CLASH datasets [65, 46]. Four sRNAs (ArcZ, CyaR, ChiX, GcvB) were identified as potentially interacting with the *ptrA-recB-recD* mRNA. However, in preliminary experiments we did not detect any significant changes in RecB expression when ArcZ, CyaR and GcvB were overproduced or deleted. Strictly speaking, we cannot completely rule out the possibility of another small RNA being involved 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.

The Hfq-mediated control of RecB expression may be just a hint of a larger bacterial post-transcriptional stress-response program to DNA damage. Indeed, post-transcriptional regulation has already been shown to impact DNA repair and genome maintenance pathways [60, 69, 10, 85]. 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 [17]. In addition, a modest (but not significant) decrease ($\sim$30%) in Hfq protein abundance has been seen in a proteomic study in *E. coli* upon DSB induction with ciprofloxacin [56]. However, 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). If Hfq indeed directly represses RecB translation (fig. 5A), a reduction in Hfq protein abundance is expected to result in more efficient RecB synthesis. Taken together, this suggests that Hfq-mediated regulation might play a larger role in DNA repair pathways that has been previously considered. As many metabolic pathways are regulated through Hfq, 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.

We have shown that RecBCD overproduction leads to less efficient repair when DSBs were induced using ciprofloxacin (fig. 5E). This observation is in keeping with previously reported results where DNA damage was induced with UV and gamma irradiation [25]. We note that another study did not confirm the toxic effect of RecBCD overproduction upon DNA damage [28] 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 relatively high level of ciprofloxacin exposure (10 ng/ml with the minimum inhibitory concentration being at 16 ng/ml in our conditions). Additionally, the necessity to limit RecBCD production might also explain the decreased viability of cells upon ChiX over-expression after DNA damage. In fact, we observe a striking similarity (in sensitivity to DNA damage) between the direct overproduction of RecBCD through plasmid expression and ChiX overproduction (which leads to increased number of RecB molecules). This suggests that inhibition of RecB translation by Hfq controls the upper bound of RecBCD expression level and is necessary for optimal DNA repair (fig. 5E).

Additionally to the control of the average number of proteins per cells, 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. 5F). Post-transcriptional regulation has been proposed theoretically as an effective strategy for reducing protein noise [107, 98, 99, 12]. In addition to theoretical studies, suppressed fluctuations of protein expression were experimentally detected for hundreds of microRNA-regulated genes in mouse embryonic stem cells [92]. Interestingly, those genes were among the ones expressed at very low level. Further investigation of a coarse-grained stochastic model (applicable to both eukaryotic and prokaryotic cells) of a gene regulated post-transcriptionally emphasised the noise reduction in low-copy regime because of the buffering

mRNA fluctuations with post-transcriptional regulators present in large abundance (relative to the mRNA targets) [92, 64]. This is exactly the case for the regulation of *recB* translation by Hfq which is a very abundant protein and the buffering of mRNA fluctuations is likely to be robust to any changes in *recB* mRNA levels upon stress conditions. These theoretical results supported by the experimental observations of miRNA-regulated genes in mammalian cells and our own results could hint to a universal functional 'niche' for post-transcriptional regulators across a wide range of organisms. Finally, 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.

Methods

Strains and plasmids

All strains and plasmids, used in the study, are listed in the table 1 and table 2, respectively. The *Escherichia coli* MG1655 and its derivatives were used for all conducted experiments except the one with arabinose-inducible expression where we used *Escherichia coli* BW27783 background strain. The strains with RecB-HaloTag fusion, MEK65 [55], was used in Halo-labelling experiments while the corresponding parental strain, MG1655, was used in smFISH experiments.

MEK1329 was built with plasmid-mediated gene replacement (PMGR) method [57] with a pTOF24 derivative plasmid, pTOFΔ*recB* [23]. The deletion of *recB* gene was verified with a phenotypical assay, polymerase chain reaction (PCR) and Sanger sequencing of the locus of interest.

MEK1902 and MEK1457 were constructed with lambda red cloning technique [24] using hfq_H1_P1 and hfq_H2_P2 primers. MEK1888 and MEK1449 were constructed with lambda red using ChiX_H1_P1 and ChiX_H2_P2 primers. *hfq* and *chiX* deletions were verified by PCR.

pIK02 was constructed to allow arabinose-inducible expression of *recB*. In brief, a fragment of the backbone plasmid was amplified with the primers oIK03 and oIK04 from pBAD33 vector and treated with DpnI (NEB) enzyme. *recB* gene was amplified from *E. coli* chromosome using the oligos oIK10 and oIK11. The purified PCR product was inserted into the backbone plasmid using Gibson assembly and the output was transformed into DH5α cells. 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 [46] using chiX_ZA21 and pZA21MCS_5P_rev primers.

All oligonucleotide sequences, used in the study, are listed in the table 3.

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 (fig.1-2) or 1X MEM Essential and MEM Non-Essential Amino Acids (Gibco®) (fig.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 in the final concentration of 500 $\mu g/ml$. In the DSB induction experiments, a sub-lethal concentration of ciprofloxacin 4 ng/ml was used (minimal inhibitory concentration (MIC) is 16 ng/ml). Ampicillin (50 $\mu g/ml$), kanamycin (50 $\mu g/ml$) or chloramphenicol (30 $\mu g/ml$) were added to cultures when appropriate.

Cell cultures from frozen -80°C stocks were grown in a medium overnight (14-16 hours) at 37°C. The overnight cultures were diluted (1:300) into 15 ml medium and grown at 37°C to the mid-exponential phase (optical density OD_{600} = 0.2-0.3). Unless otherwise noted, 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 ($t = 0$). 800 μl of cells were harvested with 1 min intervals and fixed directly in 200 μl of 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* hybridisation (smFISH) experiments were carried out according to the established protocol [101]. Bacterial cells were grown as described above, fixed in formaldehyde solution, permeabilised in ethanol and hybridised with a *recB* specific set of TAMRA-labelled RNA FISH probes (LGC Biosearch Technologies). Probe sequences are listed in the table 5. Unbound RNA FISH probes were removed with multiple washes, and then samples were visualised as described below.

Halo-labelling

For single-molecule RecB protein quantification we followed the Halo-labelling protocol, described here [55]. *E. coli* strains with RecB-HaloTag fusion, grown as described above, were labelled with JaneliaFluor®549 (JF549) dye (Promega) for one hour, washed with aspirator (4-5 washes), fixed in 2.5% formaldehyde and imaged on agarose pads (the details of imaging procedure are given below). In each experiment a parental strain

that does not have HaloTag fusion ('no HaloTag') was subjected to the protocol in parallel with the primary samples.

Microscopy setup 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 an 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 visualised with an exposure time of 2 s and EM gain 200. For each XY position on an agarose pad, a Z-stack of 6 images centered around the focal plane (total range of 1 μm with 0.2 μm step) was obtained in both channels. A set of multiple XY positions were acquired for each slide to visualise $\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 [48]. 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 [48]) 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 utilised to detect foci in fluorescence smFISH images [1, 101]. 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, specific signal was separated from false positive spots (fig. S1a, [101]). 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, [101]), and all identified spots were normalised by the one-mRNA equivalent.

To detect fluorescent single-molecule foci in the Halo-labelling experiments, the first step of Spätzcell analysis was modified (similarly to the modification for fluorescent protein detection implemented here [120]). Specifically to 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 to separate specific signal from false positive spots (fig. S4). Due to 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 number 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 $vol \times OD_{600} = 2.5\ ml$ and flash-frozen in liquid nitrogen. In the Halo-labelling experiments, samples for RNA extraction were grown at 37°C without the dye and collected right before the fixation step. Cells were lysed and total RNA was extracted with the Guanidium thiocyanate phenol protocol [47]. 10 μg of the extracted RNA was treated with 1 μl of TURBO^{TM} DNase (Thermo Fisher Scientific) and 0.1 μl of SUPERase · In^{TM} RNase Inhibitor (Thermo Fisher Scientific) in a 10 μl reaction for 1 hour at 37°C and purified by Ethanol precipitation.

Primers for real-time qPCR experiments were designed in the OligoPerfect Primer Designer (Thermo Fisher Scientific) and listed in the table 4.

RT-qPCR quantification was carried out on 20 *ng* of RNA with the Luna® Universal One-Step RT-qPCR Kit (NEB). All qPCRs were performed in technical triplicates alongside with no template control (NTC) and no reverse transcriptase control (NRTC).

The qPCR reactions were carried out on a LightCycler®96 (Roche) with a 40 cycles of two-step amplification (95°C, 10s, 4.4°C/s; 60°C, 30s, 2.2°C/s) followed by a melting curve (65°C, 60s, 2.2°C/s; 97°C, continuous acquisition). Initial steps of the data analysis were performed in the LightCycler®96 application software (v1.1.0.1320, Roche) with the built-in protocols: (i) *Absolute Quantification (Fit Point)* for determination of threshold cycles (Ct) and (ii) *High Resolution Melting* for evaluation of primer specificity. 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 measurements

Bacterial cells, seeded by inoculating from fresh colonies, were grown in the appropriate medium at 37°C overnight. The cultures were diluted (1:1000) into 20 *ml* of the medium and grown in a shaking incubator at 37°C. Optical density measurements were taken overday at an absorbance wavelength of 600 *nm* (OD_{600}) with the interval of 15-20 *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 (`fitlm`) on the OD_{600} measurements and normalised to the $OD_{600(t=0)}$.

In serial dilution assays, cell cultures were grown in Luria Broth at 37°C for ∼15 *hrs*. The overnight cultures were normalised to the optical density of OD_{600} = 1.0, spotted in 5- or 10-fold serial dilutions (started from OD_{600} = 1.0) onto LB plates supplemented with 0-10 ng/ml of ciprofloxacin and incubated at 37°C overnight.

Data analysis of CRAC data

The CLASH data (NCBI Gene Expression Omnibus (GEO), accession number GSE123050) were analysed in the study [46] with the entire pipeline available here [2, 3]. The SGR file, generated with `CRAC_pipeline_PE.py`, was used to plot the genome browser track in fig. S6.

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:

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

The probability density function of the fitted distribution was found with the maximum likelihood estimation (MLE) method in MATLAB (`fitdist`). The fitted parameters, r and p, with the 95% confidence intervals are given as $r = 0.335$ ($CI_{95\%}$: [0.267, 0.403]) and $p = 0.513$ ($CI_{95\%}$: [0.456, 0.570]).

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) and eq. (2)):

$$b = \frac{1-p}{p} = 0.95\, molec \tag{3}$$

$$k_m = \gamma_m r = 0.21\, min^{-1} \tag{4}$$

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 (here we used protein removal rate, $\gamma_p = 0.015\ min^{-1}$, and mean mRNA number in *newborns*, $\langle m \rangle = 0.323$):

$$k_p = \frac{\gamma_p \langle p \rangle}{\langle m \rangle} = 0.15\, min^{-1} \tag{5}$$

All parameters of the model are listed in table 6.

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 [4]. 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`. The difference between experimental protein concentration distributions was tested with a two-sample Kolmogorov-Smirnov test (`kstest2`, MATLAB). Empirical cumulative distribution functions (CDFs) were plotted with

`ecdf` (MATLAB). To visualise data with box plots, MATLAB function `boxplot` was utilised.

In the DSB induction experiments, the average translational efficiency (TE) was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m (fig. 3F). The uncertainty on translational efficiency values (error bars) was calculated as $\Delta TE = TE\sqrt{(\Delta c_p/c_p)^2 + (\Delta c_m/c_m)^2}$, where Δc_m and Δc_p are standard deviations of mRNA and protein concentrations across three 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 (direct method) [36], executed in MATLAB environment [5]. Parameters, used in simulations, are listed in the table 6. 10,000 of replicate simulations have been performed for a given condition and probabilistic steady-state mRNA and protein distributions were recovered for population snapshot data.

Data and code availability

Microscopy data that support this study are in the process of depositing to the open access Edinburgh DataShare repository (University of Edinburgh) and will be accessible via a doi. The analysis software used for spot detection in smFISH experiments is available from https://code.google.com/archive/p/spatzcells/. The modified version of the analysis used for spot detection in Halo-labelling experiments and bright-field segmentation pipeline are available from gitlab.com/ikalita/spotdetection_spatzcells and gitlab.com/ikalita/cellsegmentation_mek respectively.

The following published datasets were used in this study:

- CRAC/CLASH data (NCBI Gene Expression Omnibus ID GSE123050) [46]: GSE123050
- RIL-seq data (ArrayExpress ID E-MTAB-3910) [65]: E-MTAB-3910

The entire pipeline for the analysis of the CRAC/CLASH and RIL-seq 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;

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, Hannah Taylor, Lorna McLaren, Benura Azeroglu, and James Holehouse for their technical support and assistance with cloning, image analysis, and modelling. 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.

Main Figures

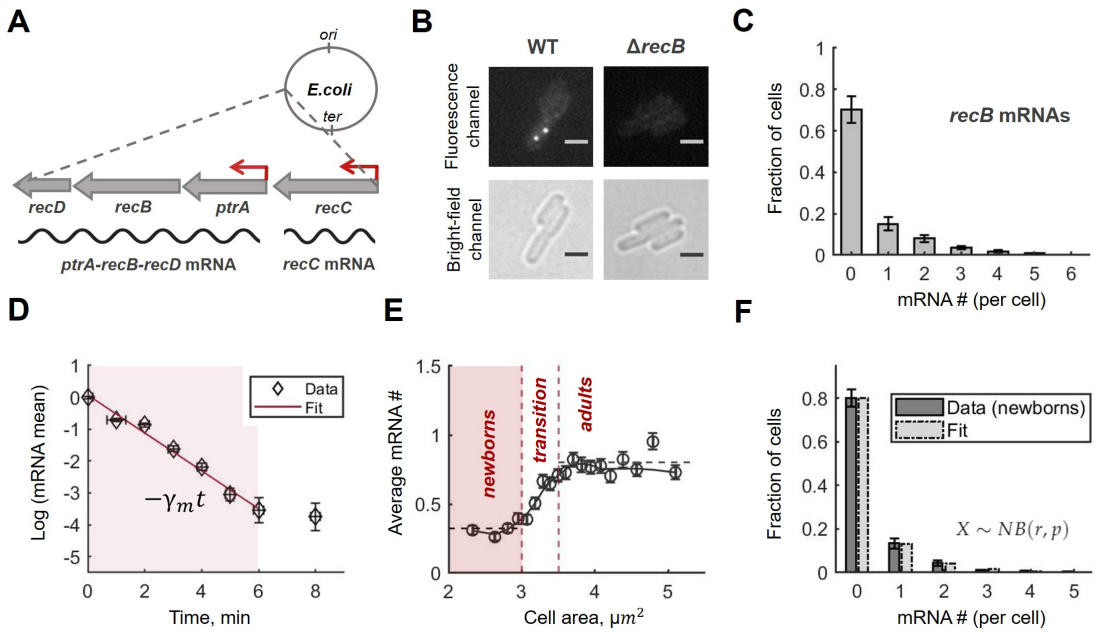

Figure 1: *recB* mRNAs are lowly abundant, short-lived and constitutively expressed.
A: Schematic description of the *recBCD* locus, its location on the *E. coli* chromosome and the corresponding mRNAs. **B:** Examples of fluorescence and bright-field images of *recB* mRNA FISH experiments in wild-type (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 normalised 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 6). **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 neighboring groups. Based on the mean mRNA numbers, all cells were separated into three subpopulations: *newborns*, *cells in transition* and *adults*. **F:** The 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 $D_{KL} = 0.002$.

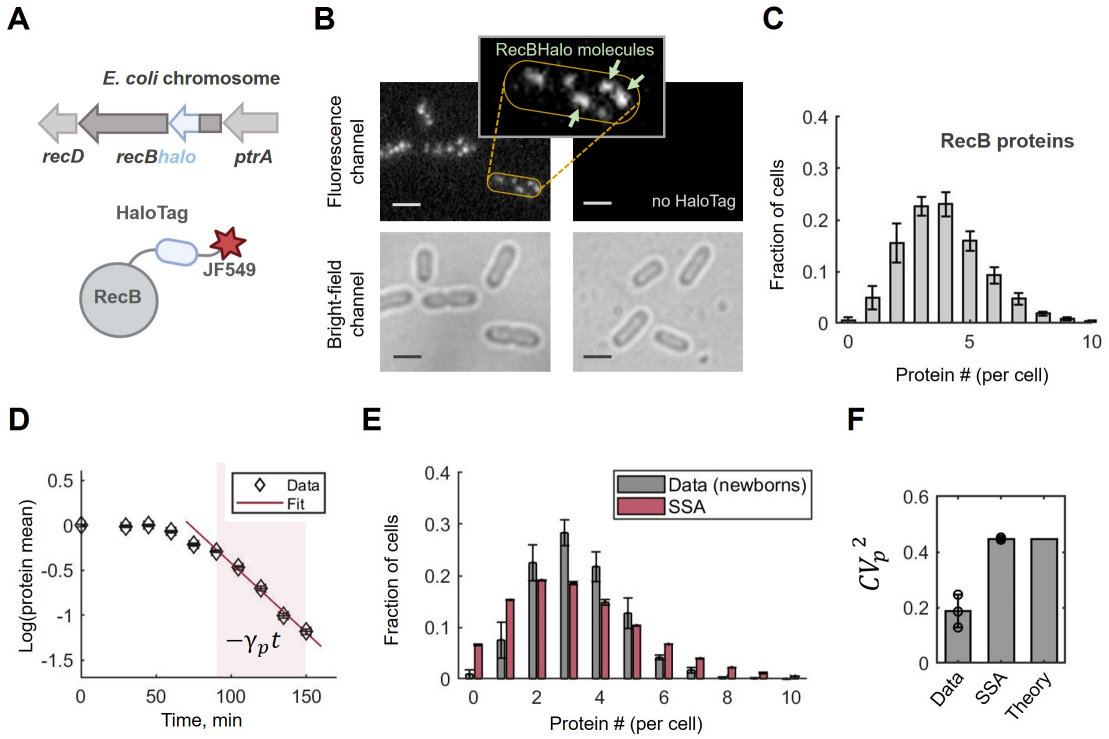

Figure 2: RecB proteins are low abundant, long-lived and show an evidence of noise reduction.
A: Schematic of RecB-HaloTag fusion at the endogenous *E. coli* chromosomal locus of the *recB* gene. RecB-HaloTag is conjugated to the Janelia Fluor 549 dye (JF549). **B:** Examples of fluorescence and bright-field 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 ∼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 normalised 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 6). **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 the table 6. The Kullback–Leibler divergence between the distributions is $D_{KL} = 0.15$. **F:** Comparison of the coefficients of variation (squared), $CV_p^2 = \sigma_p^2/\langle p \rangle^2$, 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).

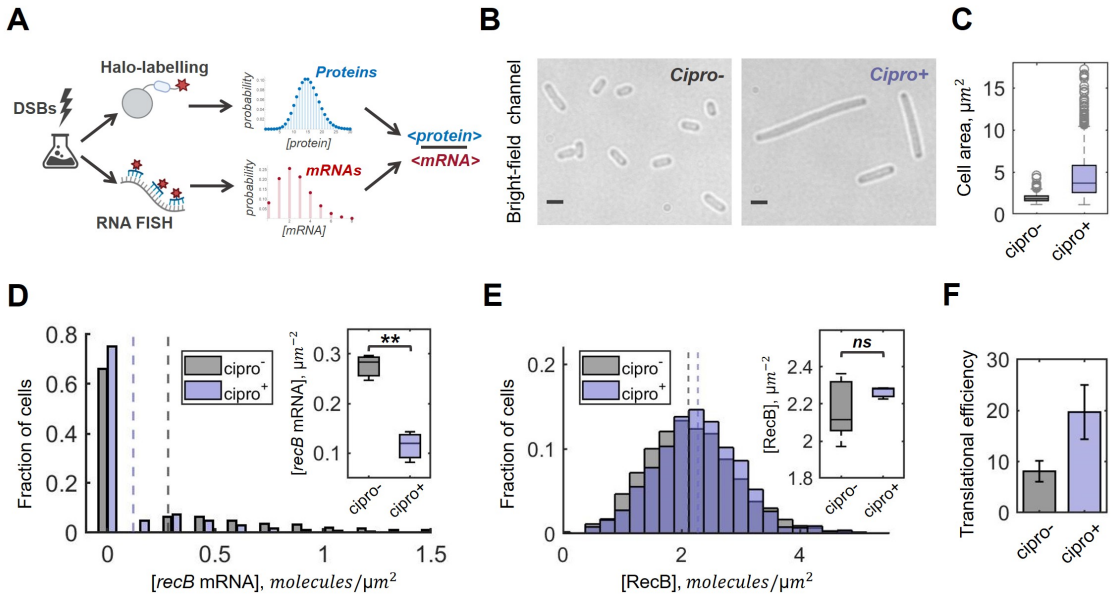

Figure 3: Translational efficiency of RecB is increased upon double-strand break (DSB) induction.
A: Schematic of the experimental workflow. DSBs were induced with 4 ng/ml of ciprofloxacin for two hours, and protein (mRNA) quantification was performed with Halo-labelling (smFISH). Mean protein and mRNA 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 (P value = 0.0023) between the average *recB* mRNA concentrations verified by 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 ($Cipro^+$). A two-sample Kolmogorov-Smirnov test was performed, and empirical cumulative distribution functions are shown in fig. S5. **Insert:** Box plot shows non-significant difference (P value = 0.36) between the average RecB concentrations verified by two-sample t-test. **F:** The average number of proteins produced per one mRNA in intact and damaged conditions (*translational efficiency*). Average translational efficiency TE (for each condition) was calculated as the ratio between the mean protein concentration c_p and the mean mRNA concentration c_m. The TE values for $Cipro^-$ and $Cipro^+$ are (8.1 ± 2.0) and (19.7 ± 5.3), respectively.

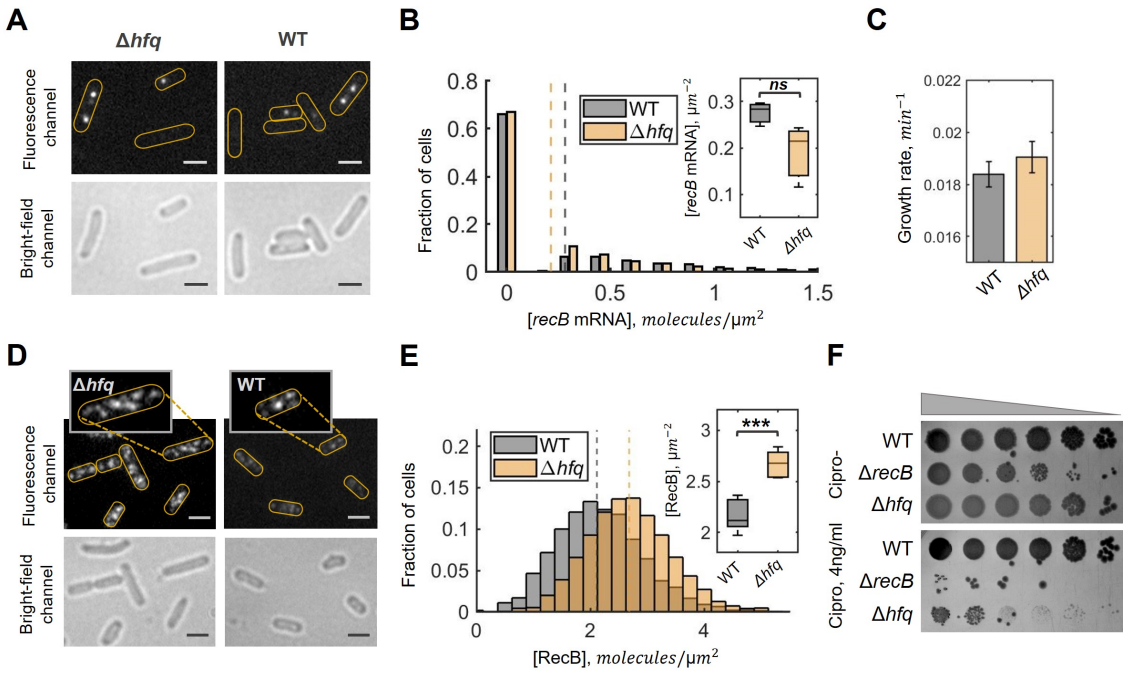

Figure 4: RecB expression is regulated by Hfq protein *in vivo*.
A: Examples of fluorescence and bright-field images of *recB* mRNA FISH experiments in Δ*hfq* and wild-type cells. Yellow outlines indicate rough positions of bacterial cells in the fluorescence channel. Scale bars represent 2 μm. **B:** *recB* mRNA concentration distributions quantified with smFISH in *hfq* mutants (orange) and wild-type cells (grey). 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*). **Insert:** Box plot shows non-significant difference (P value = 0.11) between the average *recB* mRNA concentrations in both strains verified by two-sample *t*-test. **C:** Growth rates for Δ*hfq* (orange) and wild type (grey) strains obtained from optical density measurements at an absorbance wavelength of 600 nm (OD_{600}). 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]) was used for growth rate estimation (fig. S7). Three replicated experiments have been performed for each strain; error bars represent 95% confidence intervals. **D:** 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 2μm. Both fluorescence images are shown in the same intensity range while different background modulation was applied in the zoom-in figures (for better spots visualisation). **E:** RecB concentration distributions quantified with Halo-labelling in wild-type cells and *hfq* mutants. 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*). A two-sample Kolmogorov-Smirnov test was performed, and empirical cumulative distribution functions are shown in fig. S7. **Insert:** Box plot shows significant difference (P value = 6.83×10^{-4}) between the average RecB concentrations in wild-type and Δ*hfq* cells verified by two-sample *t*-test. **F:** A 10-fold serial dilution assay for wild-type, Δ*recB* and Δ*hfq* strains. A plate was supplemented with no (top) or 4 ng/ml (bottom) of ciprofloxacin.

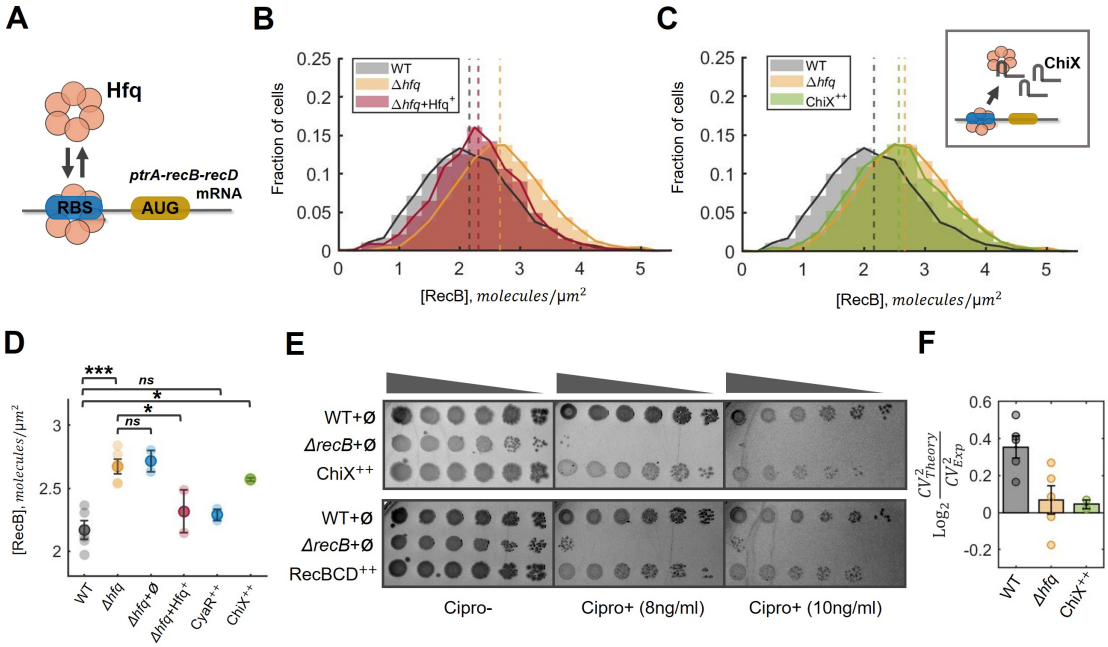

Figure 5: The proposed post-transcriptional regulation mechanism of RecB expression.
A: A hypothetical model of Hfq downregulating RecB translation by blocking the ribosome binding site of the *ptrA-recB-recD* mRNA. **B:** Partial complementation of RecB expression to the wild-type level demonstrated 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 *hfq* mutants from fig. 4E 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 *hfq* mutants from fig. 4E are shown for relative comparison. Dashed lines represent the average RecB concentration in each condition. **D:** Mean RecB concentration measured in wild-type cells (WT), *hfq* mutants (Δ*hfq*), Δ*hfq* cells harboring backbone plasmid pQE80L (Δ*hfq* + ∅) or the plasmid expressing Hfq protein (Δ*hfq*+Hfq^{+}), wild-type strain carrying a plasmid expressing CyaR (CyaR^{++}) or ChiX (ChiX^{++}). Error bars represent s.e.m. calculated from replicated experiments; mean RecB concentration in each replicated experiment is shown with dim dots. Significance was evaluated with two-sample *t*-test, *P* values are indicated with asterisks. **E:** 5-fold serial dilution assays for wild type, Δ*recB* and wild type carrying ChiX expression plasmid (ChiX^{++}) (top) or RecBCD expression plasmid (RecBCD^{++}) (bottom). Plates were supplemented with no (left), 8 ng/ml (middle) or 10 ng/ml (right) of ciprofloxacin. **F:** The ratio between theoretically predicted (CV^2_{Theory}) and experimentally measured (CV^2_{Exp}) coefficient of variation (squared) presented in logarithmic scale. The expected values were computed via the average protein copy number as $1/\langle p \rangle$ for WT, Δ*hfq* and ChiX^{++} samples. Error bars represent s.e.m. calculated from the replicated experiments.

Supplementary Information. Models

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 utilised to study fluctuations underlying protein production in prokaryotic and eukaryotic cells [111, 79, 35, 97, 54, 102]. 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 [117]. We focus on two species, mRNAs and proteins, and denote its molecule numbers in a cell as m and p, respectively.

Considering gene expression in bacterial cells allows to make several further assumptions. First, transcription of bacterial genes was demonstrated to occur in bursts [37, 103] and quite a few transcriptional models describing gene bursting have been proposed in the literature [79, 82, 52, 49]. 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 [80, 88, 91]. Secondly, based on a broadly accepted view of the exponential RNA degradation in bacteria [96, 73, 95], 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 due to cell growth and division [108, 53]. 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:

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

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) and eq. (A3)):

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

Due to 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) [117, 30, 78].

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

Finally, the mRNA and protein coefficients of variation (squared), defined as variance over squared mean, are given as follows [111, 81, 79, 97]:

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

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

The analytical probability distribution of protein numbers under assumption of a long-lived protein relative to short-lived mRNA was derived here [97].

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 $k_m V(t)$ [93, 117]. 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) and eq. (A9)) by cell volume $V(t)$ and replace $dm(t)/dt$ and $dp(t)/dt$ with the expressions from the equations (eq. (A10) and eq. (A11)). Based on the experimental evidence [121, 68], 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:

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

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

$$+ \left(-\frac{c_{m0} k_p}{\gamma_m - \gamma_{p_{act}}} + \frac{k_p k_m}{(\gamma_m - \gamma_{p_{act}})(\gamma_m + \gamma)} \right) e^{-(\gamma_m + \gamma)t} + \frac{k_p k_m}{(\gamma_{p_{act}} + \gamma)(\gamma_m + \gamma)}$$

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) and 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 what guaranteed the fulfillment of the time-scale requirement (fig. 3).

Supplementary Figures

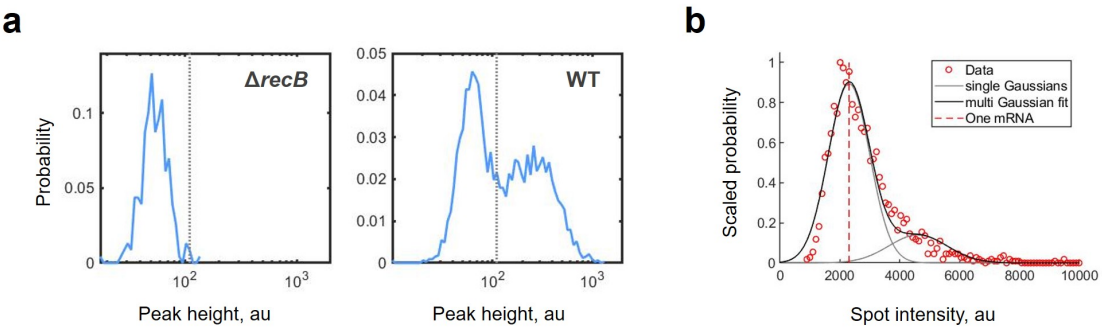

Figure S1: smFISH image analysis performed by Spätzcell. a: Peak height intensity profiles for spots detected in the wild-type and Δ*recB* samples. The dashed line indicates the intensity threshold (99.9%) that separates a specific signal from the background. **b:** Integrated spot intensity histogram 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).

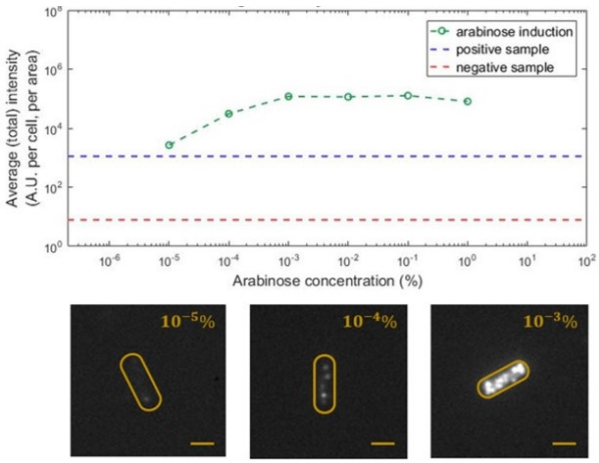

Figure S2: Sensitivity of the smFISH protocol allows for quantification of low-abundant *recB* mRNAs. Top: Fluorescence signal, detected in the samples with *recB* over-expression from an arabinose-inducible plasmid (pIK02), is presented as a function of arabinose concentration (green-dashed line). The intensity of the fluorescence signal in the wild-type and Δ*recB* strains is indicated with blue-dashed and red-dashed lines, respectively. **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.

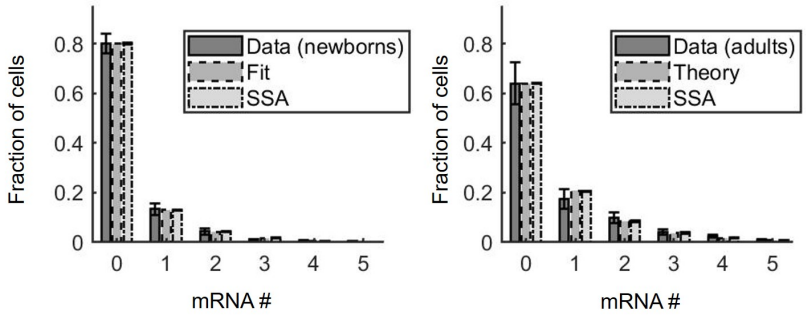

Figure S3: Independent transcription from *recB* gene copies. *recB* mRNA distributions for *newborns* (**left**) and *adults* (**right**). The experimental data from fig. 1C was conditioned on cell size (<3.0 μm^2 for *newborns*) and (>3.5 μm^2 for *adults*). Total number of cells: 2,180 (*newborns*) 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.

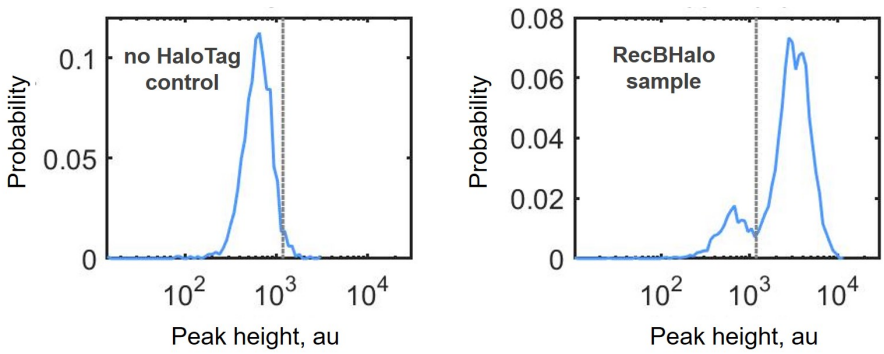

Figure S4: Image analysis of Halo-labelling experiments performed by a modified version of Spätzcell. Peak height intensity profiles for spots detected in the strain with RecB-HaloTag fusion (RecBHalo sample) and its parental strain (no HaloTag control). The dashed line indicates a chosen intensity threshold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

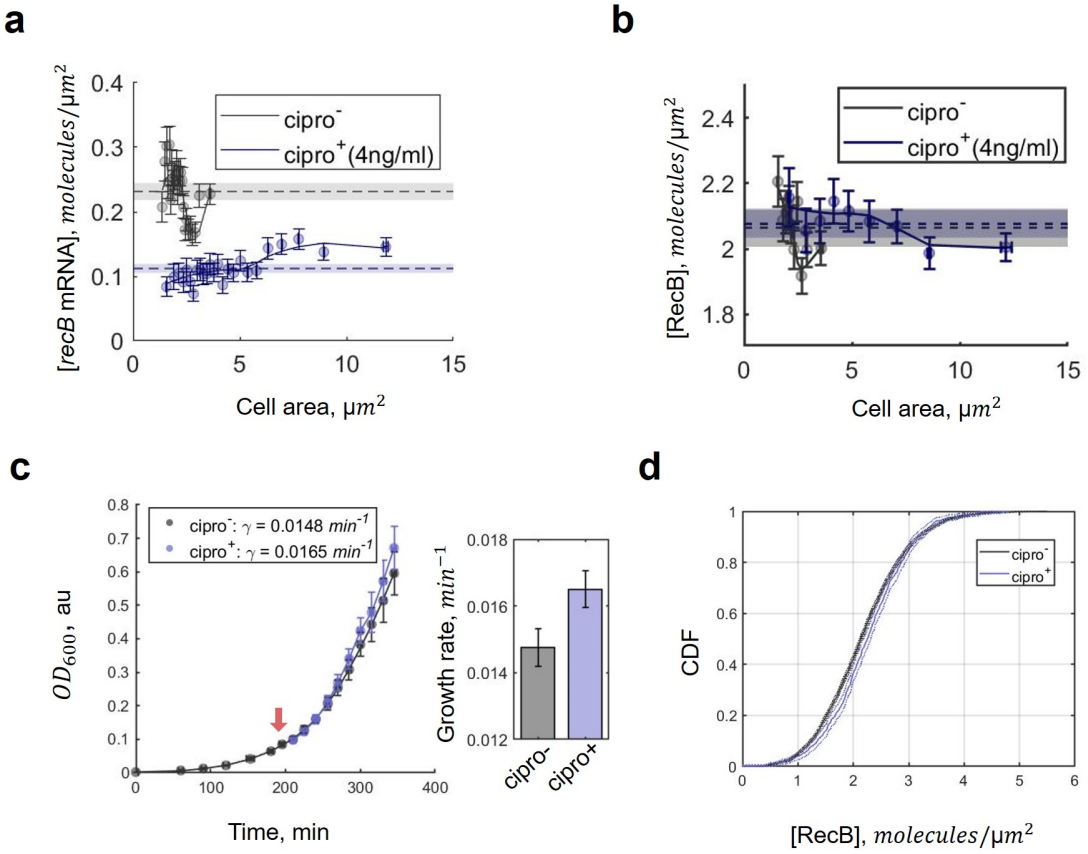

Figure S5: Analysis of *recB* mRNA and RecB protein concentrations and growth rate measurements upon DSB induction. **a:** *recB* mRNA concentration for perturbed (blue) and unperturbed (grey) samples represented as a function of cell area. The data from fig. 3D were binned by cell size and averaged 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. **d:** Empirical cumulative distribution functions (CDFs) for RecB protein concentration in the perturbed (blue) and unperturbed (grey) samples (the data from fig. 3E were used). A two-sample Kolmogorov-Smirnov test was applied to calculate P value ($3 \cdot 10^{-4}$).

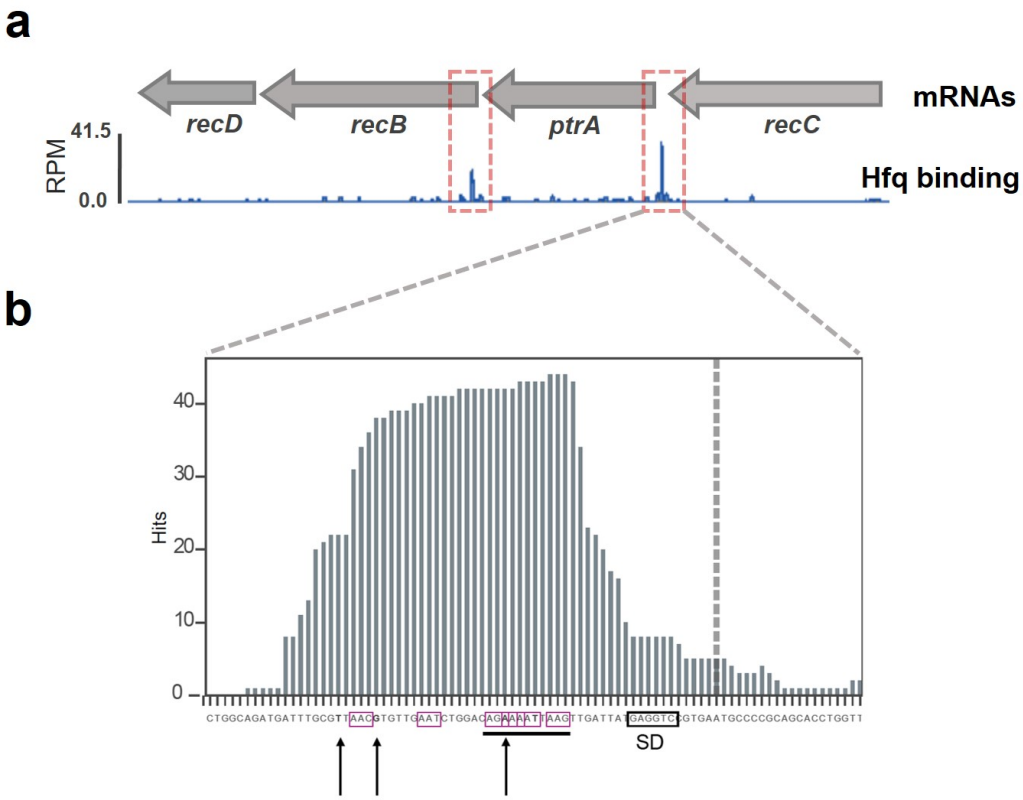

Figure S6: Hfq binding within *recC-ptrA-recB-recD* transcripts identified in the Hfq-CRAC experiment in *E. coli* [46]. a: Genome browser track showing Hfq binding to *recC-ptrA-recB-recD* mRNAs. The coverage is normalised to reads per million (RPM). The major peaks of interest are highlighted by red dashed boxes. **b:** The main binding peak magnified and 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) [112, 58], are highlighted in the red boxes while the cluster of Hfq binding motifs is underlined.

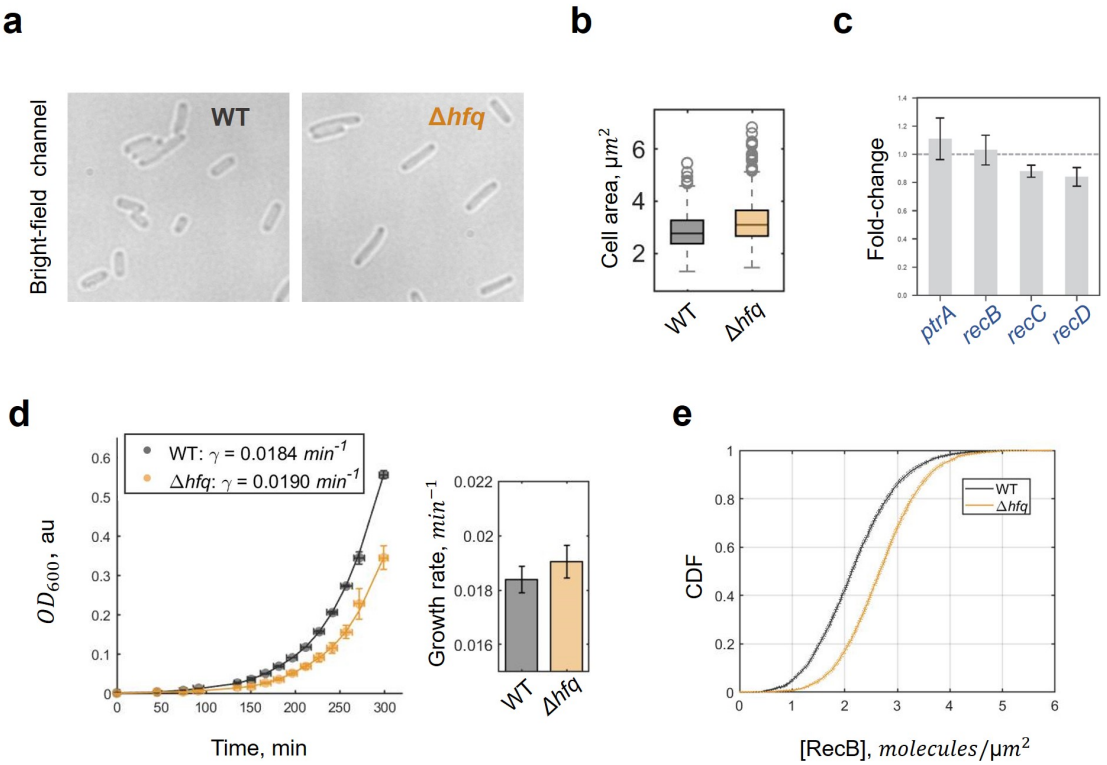

Figure S7: Cell size analysis, RT-qPCR quantification and growth rate measurements in *hfq* mutants. **a:** Examples of bright-field images for wild-type and Δ*hfq* strains. **b:** Box plots represent cell area distributions for wild-type and Δ*hfq* samples. **c:** *ptrA*, *recB*, *recC* and *recD* transcripts quantified by RT-qPCR in *hfq* mutants and normalised 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]) 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. **e:** Empirical cumulative distribution functions (CDFs) for RecB protein concentration in the wild-type (grey) and Δ*hfq* (orange) samples (the data from fig. 4E were used). A two-sample Kolmogorov-Smirnov test was applied to calculate P value ($5 \cdot 10^{-151}$).

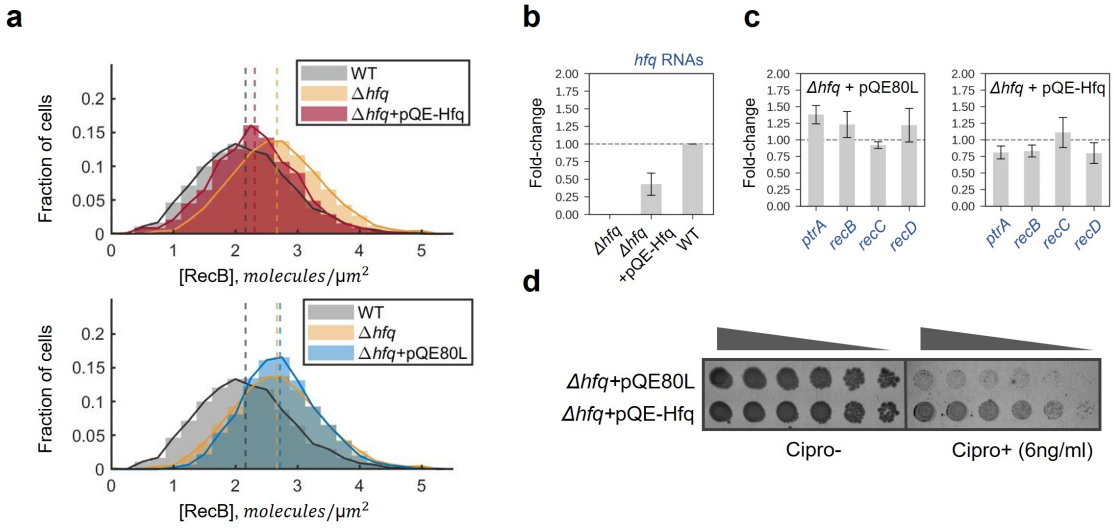

Figure S8: Controls for Hfq complementation experiment. a: RecB protein concentration distributions quantified in *hfq* mutants carrying pQE80L (blue) or pQE-Hfq (red) plasmid. The histograms represent the average across two replicated experiments per each condition. The RecB concentration histograms for wild type and *hfq* mutants 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 normalised 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 normalised 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.

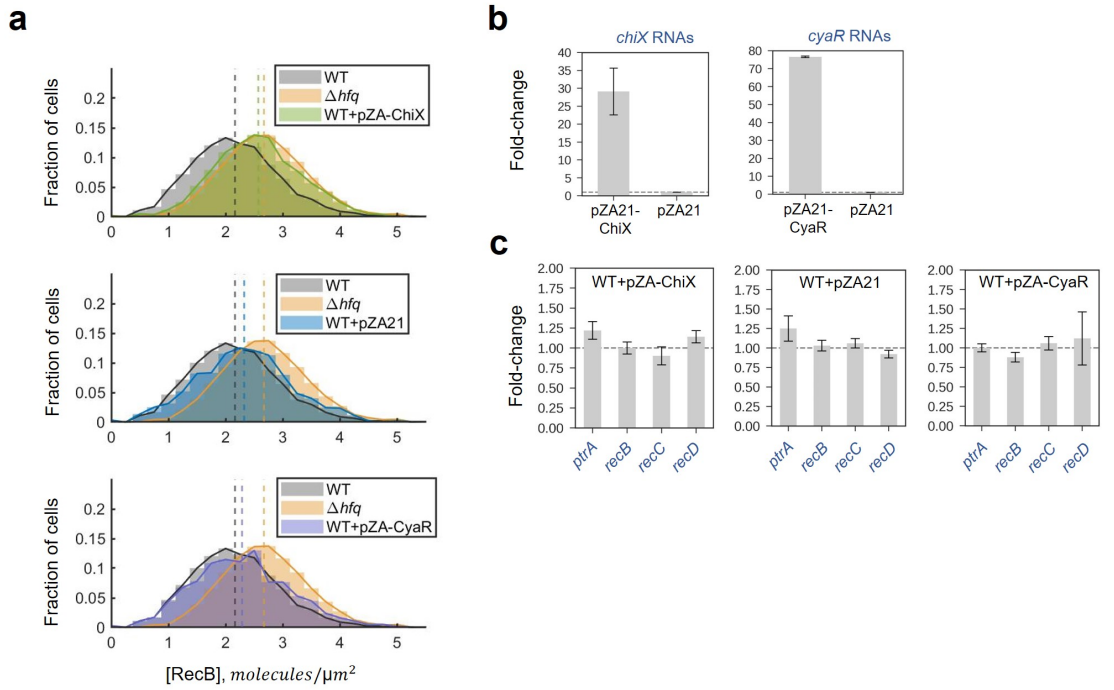

Figure S9: Controls for ChiX over-expression experiment. a: RecB protein concentration distributions quantified in wild-type cells carrying pZA-ChiX (green), pZA21 (blue) or pZA-CyaR (purple) plasmid. The histograms represent the average across two replicated experiments per each condition. The RecB concentration histograms for wild type and *hfq* mutants 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+pZA-ChiX; 0.42(ns) for WT and WT+pZA21; 0.37(ns) for WT and WT+pZA-CyaR). **b:** *chiX* and *cyaR* RNA levels quantified by RT-qPCR in wild-type cells carrying pZA-ChiX (left) or pZA-CyaR (right) over-expression plasmid. The results were normalised 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 pZA-ChiX (left), pZA21 (middle) or pZA-CyaR (right) plasmid. RNA levels were normalised to the corresponding expression levels in wild type.

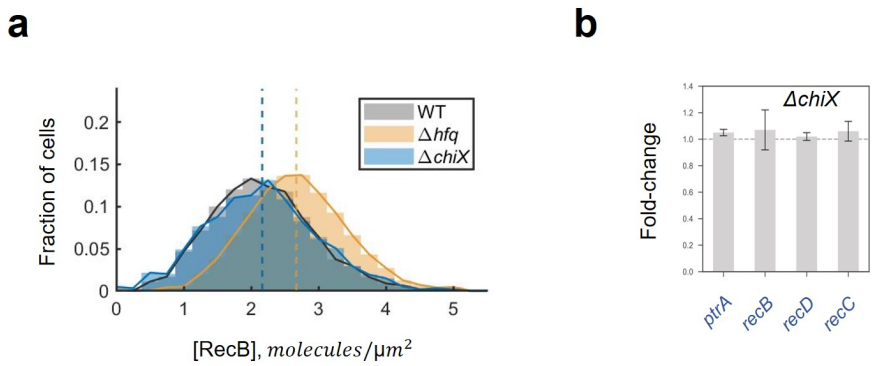

Figure S10: RecB protein and mRNA quantification in Δ*chiX*. a: RecB protein concentration distribution quantified in Δ*chiX* cells. The histogram represents the average across two replicated experiments. The RecB concentration histograms for wild type and *hfq* mutants are plotted as references in grey 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 normalised to the corresponding expression levels in the wild-type sample.

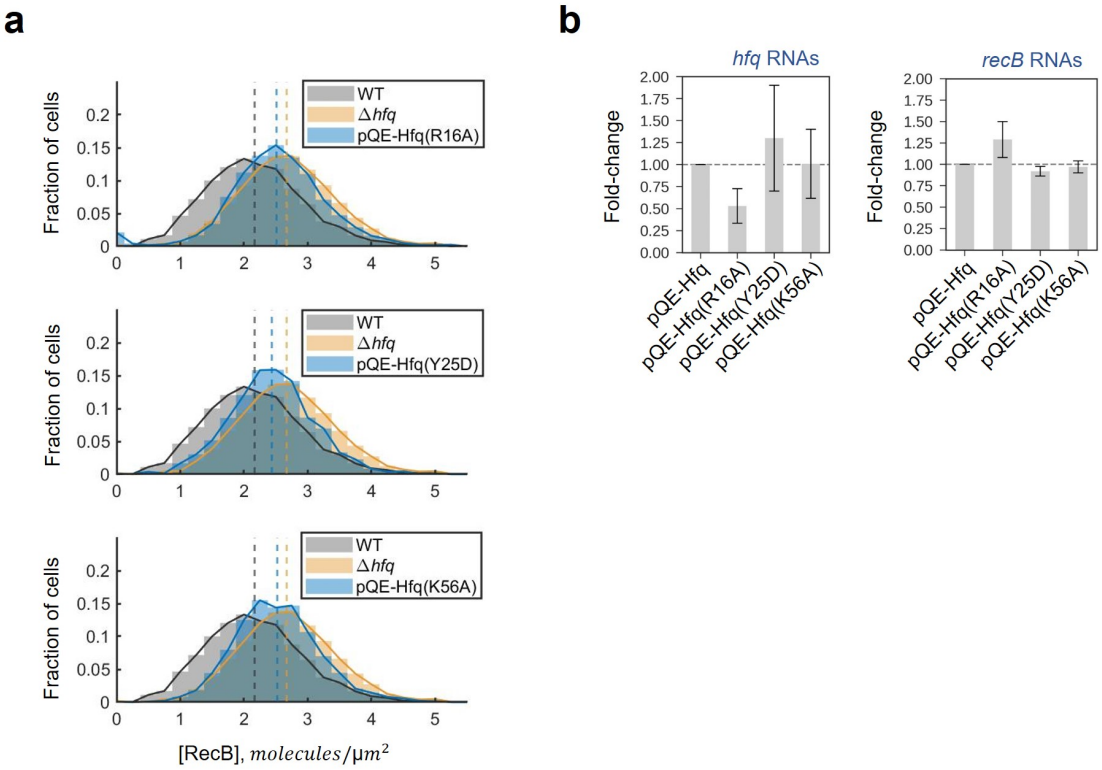

Figure S11: Testing Hfq with point mutations in its binding sites: proximal, distal and rim. a: RecB protein concentration distributions quantified in *hfq* mutants carrying a pQE80L-derivative plasmid with a mutated Hfq protein. Point mutations were made in one of the Hfq binding sites: rim (R16A), distal (Y25D) or proximal (K56A) [128]. The histograms represent the average across two replicated experiments for each mutation. The RecB concentration histograms for wild type and *hfq* mutants 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: 0.19(ns) for Δ*hfq* and Δ*hfq*+pQE-Hfq(R16A); 0.07(ns) for Δ*hfq* and Δ*hfq*+pQE-Hfq(Y25D); 0.30(ns) for Δ*hfq* and Δ*hfq*+pQE-Hfq(K56A)). **b:** *hfq* and *recB* RNA levels quantified by RT-qPCR in Δ*hfq* cells carrying a pQE80L-derivative plasmid with a mutated Hfq protein. The results were normalised to the corresponding expression in Δ*hfq* cells carrying the backbone plasmid, pQE-Hfq.

Supplementary Tables

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*	[55]
MEK1326	MG1655 Δ*recB*	[55]
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

Table 2: Plasmids used in the study.

Plasmids	**Description**	**Source**
pTOF24Δ*recB*	pTOF24-derivative plasmid used for construction of Δ*recB* strain (MEK1329)	[23] [1]
pBAD33	Backbone plasmid used for construction of pIK02	[38]
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	[112]
pZA21-ChiX	pZA21MCS-derivative plasmid carrying *chiX*	This work
pZA21-CyaR	pZA21MCS-derivative plasmid carrying *cyaR*	[46]
pDWS2	pBR322-derivative plasmid carrying native *recC-ptrA-recB-recD* chromosomal region	[84] [2]
pQE80L	Backbone plasmid used for construction of pQE-Hfq and pQE-Hfq(R16A/Y25D/K56A)	[72] [3]
pQE-Hfq	pQE80L-derivative plasmid carrying *hfq-His*$_6$	[72] [3]
pQE-Hfq(R16A)	pQE80L-derivative plasmid carrying *hfq(R16A)-His*$_6$	[72] [3]
pQE-Hfq(Y25D)	pQE80L-derivative plasmid carrying *hfq(Y25D)-His*$_6$	[72] [3]
pQE-Hfq(K56A)	pQE80L-derivative plasmid carrying *hfq(K56A)-His*$_6$	[72] [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 Teppei Morita and 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: 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 5: 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 6: 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] URL: https://code.google.com/archive/p/spatzcells/.

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

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

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

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

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] K Datsenko and B Wanner. "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products". In: *PNAS* 97.12 (2000), pp. 6640–6645. DOI: 10.1073/pnas.120163297.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

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

[33] 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.

[34] 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.

[35] 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.

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

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] S Jaramillo-Riveri et al. "Growth-dependent heterogeneity in the DNA damage response in *Escherichia coli*". In: *bioRxiv* (2021). DOI: 10.1101/2021.05.06.442821.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

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

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] 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.

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

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

[69] 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.

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

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] 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.

[76] 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.

[77] 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.

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

[79] 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.

[80] 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.

[81] 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.

[82] 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.

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

[84] 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.

[85] 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.

[86] 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.

[87] 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.

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

[89] 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.

[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] 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.

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

[93] 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.

[94] 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.

[95] 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.

[96] 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.

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

[98] 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.

[99] 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.

[100] 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.

[101] 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.

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

[103] 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.

[104] 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.

[105] 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.

[106] 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.

[107] 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.

[108] 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.

[109] 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.

[110] 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.

[111] 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.

[112] 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.

[113] 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.

[114] 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.

[115] 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.

[116] 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.

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

[118] 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.

[119] 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.

[120] 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.

[121] 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.

[122] 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.

[123] 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.

[124] 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.

[125] 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.

[126] 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.

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

[128] 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.