Effect of ribosome shielding on mRNA stability

Effect of ribosome shielding on mRNA stability

Carlus Deneke, Reinhard Lipowsky and Angelo Valleriani

Department of Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces,
D-14424 Potsdam, Germany

E-mail: angelo.valleriani@mpikg.mpg.de

Abstract

Based on the experimental evidence that translating ribosomes stabilize the mRNAs, we introduce and study a theoretical model for the dynamic shielding of mRNA by ribosomes. We present an improved fitting of published decay assay data in *E. coli* and show that only one third of the decay patterns are exponential. Our new transcriptome-wide estimate of the average lifetimes and mRNA half-lives shows that these timescales are considerably shorter than previous estimates. We also explain why there is a negative correlation between mRNA length and average lifetime when the mRNAs are subdivided in classes sharing the same degradation parameters. As a by-product, our model indicates that co-transcriptional translation in *E. coli* may be less common than previously believed.

1. Introduction

In the process of gene expression, mRNAs are constantly transcribed from DNA and translated by ribosomes into proteins. Both the amount and the type of proteins expressed in a cell are crucial for its survival and its capacity to respond to changes in the environment. Regulation of transcription is one of the basic mechanisms that controls the expression level of the proteins [1]. It has been found, however, that several post-transcriptional mechanisms play a substantial role in the regulation of gene expression [2]. At the level of translation, regulation occurs by modulating the rate of protein synthesis, mostly through the control of translation initiation [3–5]. In particular, in prokaryotic cells, mRNA degradation is a powerful mechanism able to strongly and quickly reduce the number of mRNA transcripts and thus control the amount of proteins produced in the cell.

Some dynamical aspects of translation have been previously described in the framework of a mean-field model [6, 7], which revealed that both the type of degradation and the timescale associated with it influence the rate of protein synthesis. A number of previous studies [8–11] have

also included mRNA degradation as a factor influencing protein synthesis. However, in all of these studies, the general approach was to assume an exponential lifetime distribution of the mRNA—thereby ignoring the complexity of degradation.

Experimentally, the amount of non-degraded mRNAs is measured in mRNA stability assays at different time points after the inhibition of further transcription [12–14]. An alternative procedure is based on the measurement of the decreasing number of pre-existing, unlabeled mRNA molecules after the introduction of metabolically labeled nucleotides [15, 16]. The resulting *experimental decay patterns* are conventionally analyzed by fitting the data to an exponential decay curve. In those cases where the data led to a poor fitting result, the data were either discarded or another fitting function—such as a double exponential—was assumed [14, 17]. Even though such a more complex function might improve the quality of the fitting procedure, it is quite unsatisfactory because it does not provide additional insight into the underlying mechanism of degradation.

In all organisms, specific ribonucleases are responsible for the degradation of the mRNAs. In particular, in *E. coli* the enzyme *RNase E* and its homologues are believed to initiate the main pathway of mRNA degradation [18]. The action of *RNase E* consists of cleaving the bond of two adjacent nucleotides of an mRNA. Once this has been accomplished, a series of further steps including exonucleolytic digestion and

endonucleolytic cleavage follows. Since the subsequent steps are fast [18], it is the initial cleaving step that determines the stability of mRNAs.

The activity of the ribonucleases is, however, not independent of the translational state of the mRNA. It was shown experimentally that when the mRNA is well covered by ribosomes, mRNA stability is dramatically enhanced [19–21] because the coverage by the ribosomes shields the mRNA from the action of the endonucleolytic enzymes such as *RNase E*. Moreover, ribosomal coverage of mRNA is itself a time-dependent process. It takes indeed a certain amount of time until the whole mRNA is covered by ribosomes and a stable polysome is built [6, 22, 23]. During this period of time, which may be affected by the degree of co-transcriptional translation, the stability of the mRNA must change with time.

This time-dependent built-up of polysomes and its effect on the degradation mechanism were not considered so far as a factor capable of influencing the decay curves in mRNA stability assays such as those presented in [12, 14]. In this paper, we develop a theoretical model of mRNA shielding by ribosomes and apply the model to fit experimental data about mRNA decay in *E. coli*. Our model is relevant for genes that are regulated primarily via the endonucleolytic initiation of mRNA degradation, such as those described in [18]. The essential feature of the model is to assume that there are two basic degradation rates. One rate is proportional to the fraction of the mRNA covered or shielded by the translating ribosomes and one rate is proportional to the remaining, unshielded fraction of the mRNA.

2. Models and methods

In this section, we start by recalling some recent developments in the theoretical description of mRNA degradation and then introduce our model of ribosomal shielding. Finally, we briefly explain the method of simulating translating ribosomes as an exclusion process.

2.1. Review of modeling background

In [24, 25], we have developed a general theoretical description that includes both the stochasticity of the transcription process and the random nature of degradation. As described in [25], our theory assumes that each species of mRNA is generated with a constant rate up to time point $t = 0$, at which transcription is interrupted. It was further assumed that the system had reached a steady state before this interruption, which implies that each mRNA species can be characterized by an average number $N_r(t = 0)$ of molecules. While this average value depends solely on the rate of transcription and the average lifetime of the mRNA, the decay pattern that follows the interruption of transcription is highly sensitive to the nature of the degradation process [25]. From a single-molecule point of view, the degradation process determines the distribution Φ_U of the random lifetime U of each mRNA. After a time interval Δt, the relative number of mRNA, denoted by $\Lambda(\Delta t)$, is then obtained via

$$\Lambda(\Delta t) \equiv \frac{N_r(\Delta t)}{N_r(0)} = \frac{1}{\langle U \rangle} \int_{\Delta t}^{\infty} dt \, (1 - \Phi_U(t)), \qquad (1)$$

where $\Phi_U(t)$ denotes the probability that the lifetime of an mRNA is shorter than t. The average lifetime $\langle U \rangle$ is defined as

$$\langle U \rangle \equiv \int_0^{\infty} dt (1 - \Phi_U(t)), \qquad (2)$$

whereas the half-life $t_{1/2}$ follows from

$$\Lambda(t_{1/2}) \equiv \tfrac{1}{2}, \qquad (3)$$

and gives the time, after which the amount of mRNA is decreased by half relative to the steady state amount at $t = 0$.

In [25], we have shown that a specific degradation mechanism implies a certain form of the lifetime probability function Φ_U. Any such mechanism can then be validated by comparing equation (1) with experimental decay data. The simplest conceivable mechanism assumes that mRNA degradation can be described by first-order kinetics. This naïve assumption leads to a lifetime probability function that has an exponential form with a constant degradation rate ω_{deg} as given by

$$\Phi_U(t) = 1 - \exp(-\omega_{\mathrm{deg}}t). \qquad (4)$$

In general, the degradation rate constant ω_{deg} is proportional to the concentration of the degrading enzymes, but it may also contain sequence-dependent effects such as the biochemical affinity to the enzymes. Therefore, the rate ω_{deg} may be different for each mRNA species and, if equation (4) correctly described the degradation process, ω_{deg} could be determined by comparing equation (1) to experimental decay patterns. The theory presented here abandons the assumption of first-order kinetics, which leads to equation (4), and derives the lifetime distribution $\Phi_U(t)$ from explicit assumption about an age-dependent degradation rate.

2.2. Ribosome shielding and age-dependent degradation rate

A process known to influence the stability of mRNA is the size of the polysome [19–21, 26, 27]. In the process of translation, ribosomes bind to the initially untranslated (or naked) mRNA and start translation, thereby translocating the mRNA one codon at a time. The first ribosome translating a given mRNA molecule will eventually reach the stop codon after a certain time t_L. If we exclude fluctuations, then the time t_L is determined by the ratio of the sequence length L of the mRNA and the average elongation rate v of the ribosomes, i.e., $t_L = L/v$. Thus, only after this time lap, the mRNA can be seen as fully covered by ribosomes. Hence, when the age of the mRNAs is smaller than t_L, the mRNA segment downstream of the leading ribosome is not covered by ribosomes [6, 7]. Note that in bacteria it is possible that ribosomes bind to the mRNA prior to the completion of transcription [28]. In these cases, the time span between the end of transcription and the full coverage would be smaller than t_L. In section 3 we will further discuss this possibility.

The transient ribosomal coverage of an mRNA becomes important when degradation is primarily determined by endonucleolytic attack, such as in many bacteria [18], because the ribosomes would make it more difficult for the degrading molecules to bind to the mRNA [21]. In these cases, the degradation rate is thus not constant during the lifetime of

an mRNA but changes according to its age until the ribosomal coverage has reached the steady state, see figure 1.

To model this phenomenon, we assign *(i)* a degradation rate ω_1 to the more vulnerable, unshielded segment of mRNA downstream of the leading ribosome and *(ii)* a second rate ω_2 to the partially covered segment of mRNA upstream of the leading ribosome. The effective degradation rate of the mRNA depends therefore on the relative size of each segment and evolves from the initial value ω_1 to a final value ω_2. This leads to an effective, age-dependent degradation rate

$$\omega_{\text{deg}}(a) = \begin{cases} \omega_1\,(1 - a/t_L) + \omega_2\,a/t_L & \text{for } a < t_L \\ \omega_2 & \text{for } a \geqslant t_L, \end{cases} \quad (5)$$

with *a* being the age of the mRNA. Non-constant failure rates are widely known in stochastic modeling and are referred to as *hazard rates* in the literature [29]. They are related to the lifetime probability distribution via

$$\Phi_U(t) = 1 - \exp\left(-\int_0^t \mathrm{d}\tau\,\omega_{\text{deg}}(\tau)\right), \quad (6)$$

which, once inserted in equation (1), provides the decay pattern of the mRNA as a consequence of the postulated mechanism of the age-dependent degradation rate [25].

2.3. Computer simulations with growing polysomes and mRNA degradation

Both for the mRNA segment that is covered and shielded by ribosomes and the mRNA segment that is still naked and untranslated, degradation is mediated by the same endonucleolytic enzyme. It is therefore natural to expect that the two rates ω_1 and ω_2 are interrelated and to assume that the shielding effect is proportional to the ribosomal coverage ρ, thus implying

$$\omega_2 = \omega_1(1 - \rho). \quad (7)$$

Experimental results about the polysome size in different species [4, 5, 30] indicate that the coverage of the mRNA by ribosomes is relatively low, leading to a density ρ that is around 30% of the maximal coverage.

There are, furthermore, some features that expression (5) for the degradation rate does not take explicitly into account. First, it neglects the dynamics of the translation process and ignores the interaction between ribosomes, which is responsible for ribosomal traffic jams. Secondly, both the size of the degrading enzymes and the varying distances between the ribosomes, which determine the available cleavage area, are not properly considered.

To elucidate the role of these effects, we have performed simulations of totally asymmetric exclusion processes with extended objects [22, 31]. In these processes, the motion of ribosomes along an mRNA is described as a stochastic stepping along a homogeneous chain of codons (stepping on non-homogeneous chains is presented in [32]) and is a widely used tool to study the dynamics of translation [21, 22, 24, 33, 34]. The ribosomes are represented as extended objects occupying 10 codons each and their interaction is governed by a hard core repulsion, i.e., two ribosomes cannot overlap. At all times of the simulation, the mRNA can be degraded with a rate
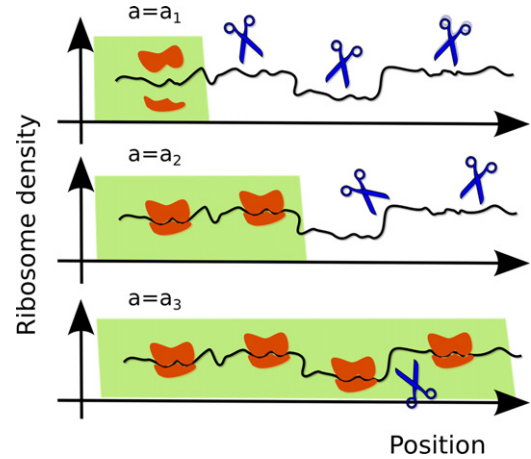

Figure 1. Age-dependent shielding of mRNA by ribosomes. (top row) The nascent chain is initially not protected by ribosomes and it is therefore more vulnerable to endonucleases, represented here as scissors. Once the process of translation has started, the ribosomes move along the mRNA and a larger fraction of the mRNA is (partially) shielded. Only after a transient time t_L the entire chain is covered by ribosomes and hence less vulnerable to endonucleolytic degradation. The portion of the mRNA covered by the ribosomes (green stripe in the cartoon) is attacked by the degrading enzymes with a rate ω_2 per unit length, whereas the portion not yet covered by the ribosomes is degraded with a rate ω_1 per unit length. The colored portion increases with the age *a* of the mRNA and the effective degradation rate interpolates between ω_1 and ω_2 according to equation (5).

ω_{deg}, which is proportional to the total available cleavage area taking into consideration both the available space between the ribosomes and the segments of mRNA that are not yet covered by ribosomes. However, in our simulations we have assumed that only stretches of mRNA of a minimal length of 20 codons do contribute to the cleavage area. This minimal length accounts for the size of the degrading enzymes and the steric repulsion between enzymes and ribosomes.

The simulation of degradation proceeds as follows. At each instant of the simulation, either a ribosome moves or the mRNA is degraded. Which decision is taken—stepping or degradation—depends on the random occurrence of events associated with the corresponding rates. The simulations are performed in continuous time using the Gillespie algorithm [35]. The stepping rate of each ribosome that is not hindered by another downstream ribosome is given by an elongation rate determined by the average velocity of ribosomes in *E. coli*, which is fixed to 10 codons per second. Furthermore, new ribosomes initiate translation with a rate $\omega_{\text{on}} = 0.33\,\text{s}^{-1}$, thus ensuring the experimentally found low-density ribosome traffic. The terminating ribosome, instead, moves smoothly out of the filament with a rate equal to the elongation rate. Moreover, the unshielded mRNA fragments that are at least 20 codons long are summed together to determine the total potential cleavage area and the rate of degradation is then proportional to it. In a next step, one computes which of these events is to take place (initiation, termination, elongation by one of the non-hindered ribosomes or degradation) by assigning to each of these events a probability equal to its own rate divided by the sum of all rates. Since the coverage of the
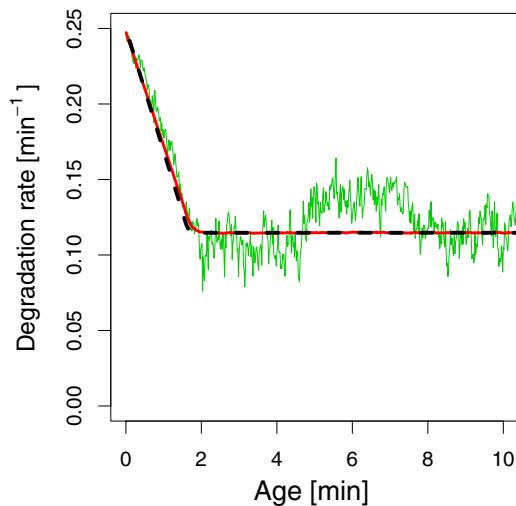
Figure 2. Stochastic simulation of the age-dependent degradation rate. During the lifetime of an mRNA its degradation rate is not constant due to the dynamic coverage by ribosomes. At each instant of a simulation, the positions and the distances between the ribosomes are different and hence the shielding of the mRNA changes with time in a random fashion. Consequently, in a single realization of the simulation the instantaneous degradation rate fluctuates (green), but it tends to decrease as the mRNA ages because a larger portion of mRNA becomes shielded. An average over 10 000 realizations shows that after an average transient time t_L, the coverage by ribosomes stabilizes and the effective degradation rate attains a constant value (red). We compared these simulations with the time-dependent degradation rate according to equation (5) with $\omega_1 = 0.25$ min^{-1} and $\omega_2 = 0.115$ min^{-1} (dashed black line). The excellent agreement allows us to conclude that the analytical mean field theory is suitable to describe the average dynamic shielding effect as obtained from the more detailed simulations. For these simulations, we considered an mRNA with a length of 1000 codons, average ribosome velocity $v = 10$ codons s^{-1} and base degradation rate $\omega_1 = 0.25$ min^{-1}. The minimum distance between ribosomes to allow cleavage was set to 20 codons. The initiation rate was set to $\omega_{on} = 0.33$ codons per second.

mRNA by ribosomes changes dynamically, these simulations give the age-dependent degradation rate. Note that each realization of the simulation has a random duration which represents the lifetime of a single mRNA.

The simulations show that the rates change randomly over time due to the fluctuations of the distances between the ribosomes (an example of a long-lived mRNA is given in figure 2). After averaging over a large number of realizations, we find a behavior of the degradation rate such as in equation (5), i.e., it decreases linearly until it approaches a constant value at a time t_L (see figure 2). This test confirms that the assumptions made in equation (5) are justified. Indeed, equation (5) reproduces the correct average behavior of the stochastic simulation that describe the dynamic shielding effect of ribosomes in more detail. Therefore, we will use the mean-field model described by equation (5) to systematically compare our theoretical results to experimental data.

2.4. Data analysis and fitting procedure

We have compared our proposed theoretical mechanism to experimental data from [14], where the stability of mRNAs
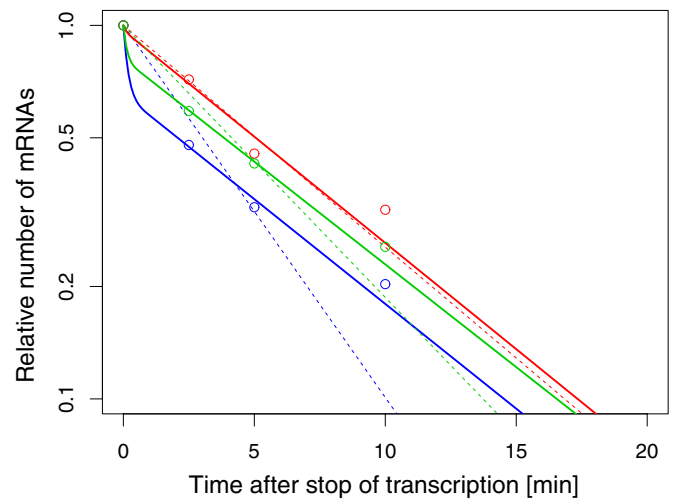

Figure 3. Experimental mRNA decay data (on logarithmic scale) after the inhibition of transcription adapted from [14] (circles). The solid curves obtained from a fit to (1) by using (5) and (6) yield an improved description of the decay data compared to the standard exponential fit (dashed lines). The curves show the decay of *atoS* (blue), *fabB* (green) and *ykgE* (red) mRNA. The mRNAs are of different length L (609, 407, 240 codons, respectively) and consequently they remain in the transient period for different time intervals $t_L = L/v$. The effect on the stability of the mRNA deriving from partial shielding is more pronounced for longer mRNAs. Indeed, with the fitted values of the rates $\omega_1 = 6.6$ min^{-1} and $\omega_2 = 0.13$ min^{-1} (*atoS*, blue line), $\omega_1 = 8.4$ min^{-1} and $\omega_2 = 0.13$ min^{-1} (*fabB*, green line) as well as $\omega_1 = 6.3$ min^{-1} and $\omega_2 = 0.13$ min^{-1} (*ykgE*, red line), we obtain the following average lifetimes, $\langle U \rangle = 0.44$ min, $\langle U \rangle = 0.59$ min and $\langle U \rangle = 2.30$ min, which decrease with the length of the mRNA.

was assessed on a genome-wide scale for *E. coli*. There, cells of the strain MG1655 were grown at 37 ° C in LB medium and further transcription was inhibited via the addition of the antibiotic *rifampicin*. The abundance of each mRNA species was recorded at different time points after the inhibition of transcription by means of the fluorescence intensity in a microarray setup.

In order to compare the theoretical decay curves arising from our model, we have performed a nonlinear regression analysis. First, we have divided the intensity values at $t = 2.5$, 5, 10 and 20 min by the value at $t = 0$ to obtain the relative amount of mRNA for each gene. We have furthermore considered only those genes, for which we believe that the antibiotic had a fast effect on transcription. This led us to discard part of the data as not reliable. Indeed, for some genes, transcription seems not to be fully inhibited since the mRNA levels are rising. In addition, it was pointed out in [25] that each decay pattern should exhibit a positive curvature when the initial condition is a steady state expression level. Therefore, we restricted our analysis to data that met these criteria within a limited margin of error. Nevertheless, the definition of the initial time $t = 0$ in the experiments is not obvious because the drug certainly needs some time before it can block the whole transcription. We have thus only considered nearly *bona-fide* data where the slope between any two adjacent experimental data points does not increase by more than 10% and the curvature between three subsequent data points is
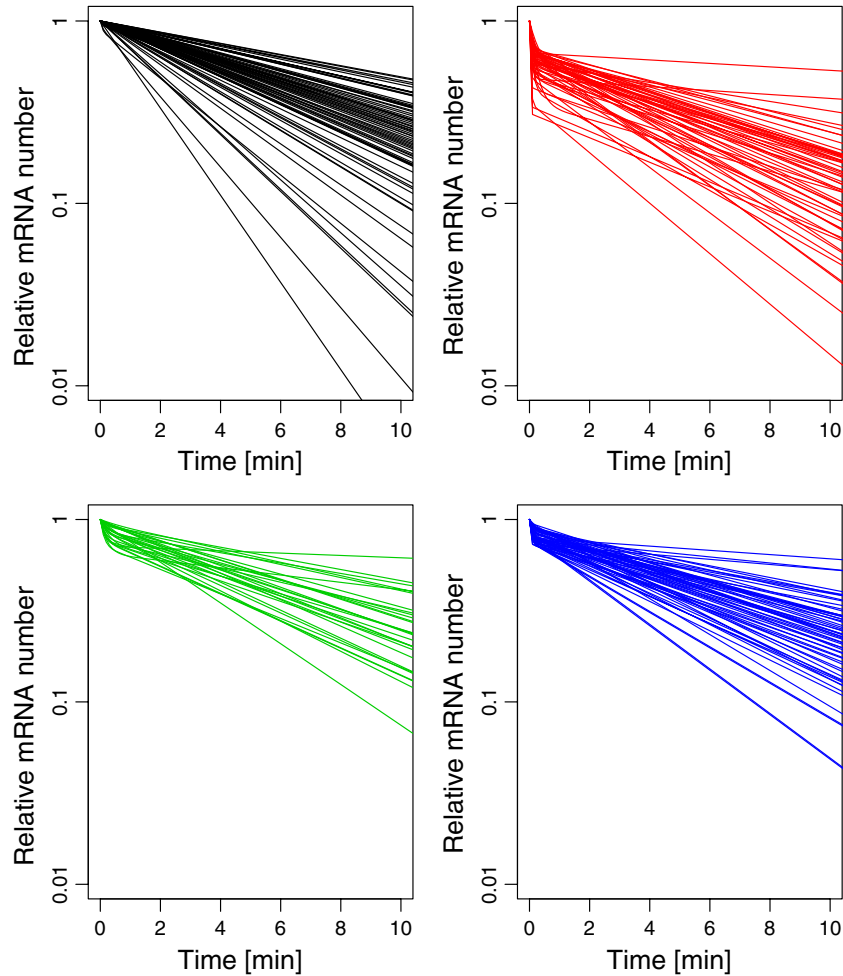
Figure 4. Categories of decay patterns found after fitting the experimental data of *E. coli* from [14] to the ribosome shielding model as described by equation (5). In all cases, the relative amount of mRNAs is plotted semi-logarithmically versus the time after the interruption of transcription. For the current analysis, we considered only experimental decay patterns that show a nearly *bona-fide* behavior (see the text), leading to a total of 220 mRNA species. We could distinguish four different classes of decay patterns. On the one hand, for 74 mRNA species one sees an exponential decay (black, top left). All other patterns deviate from an exponential decay and display a stronger decay for shorter times followed by a more moderate, exponential decay at larger *t*. In particular, the red patterns decay extremely fast for a short period followed by a transition to a moderate decay (top right). The 57 mRNAs of this class have a high value of the unshielded degradation rate ω_1 and are thus degraded very efficiently before reaching a stationary ribosome coverage at time t_L. Conversely, the 60 mRNAs depicted by the blue patterns (bottom, right) have a more moderate decay at short times, but also display a transition to a slower, exponential decay at time t_L. The mRNAs pertaining to the former categories are short to intermediate, whereas long mRNAs (bottom, left) have a transient to stationary ribosome coverage which is particularly long. The 29 mRNAs of this group are longer than 400 codons. Note that the subdivision of the decay patterns in the four classes is by no means unambiguous but gives an overview over different types of patterns. Three characteristic examples are discussed in detail in figure 3.

convex within an error of up to 25%. Furthermore, in [14] the data values have been shifted to account for a spurious background signal. This procedure occasionally led to negative data values, which we have set to zero in order to avoid dealing with negative mRNA numbers.

For the subsequent analysis, we have assumed an average elongation rate $v = 10$ codons per second, which is a typical value for *E. coli* [36, 37]. Hence, for each gene, we have determined the values of the rates ω_1 and ω_2 that minimize the mean square deviation between the experimental decay data and the theoretical curve calculated in equation (1) by using (5) and (6). For the sake of comparison, we have performed the same analysis with a constant degradation rate by using equation (4).

3. Results and discussion

A schematic view of our shielding model is illustrated in figure 1. The timescale required to build a full polysome on a newly synthesized mRNA is t_L. Before this time, the mRNA is only partially shielded by the ribosomes and its degradation is a time-dependent combination of two degradation rates. This model allows us to perform a fit of the experimental data using equations (1), (5) and (6) according to the methods explained in section 2.

In figure 3, we show three examples of the decay patterns from *E. coli* [14] and the corresponding fit with our model illustrated in figure 1. For comparison, we have repeated the analysis with an exponential function as given by equation (4).
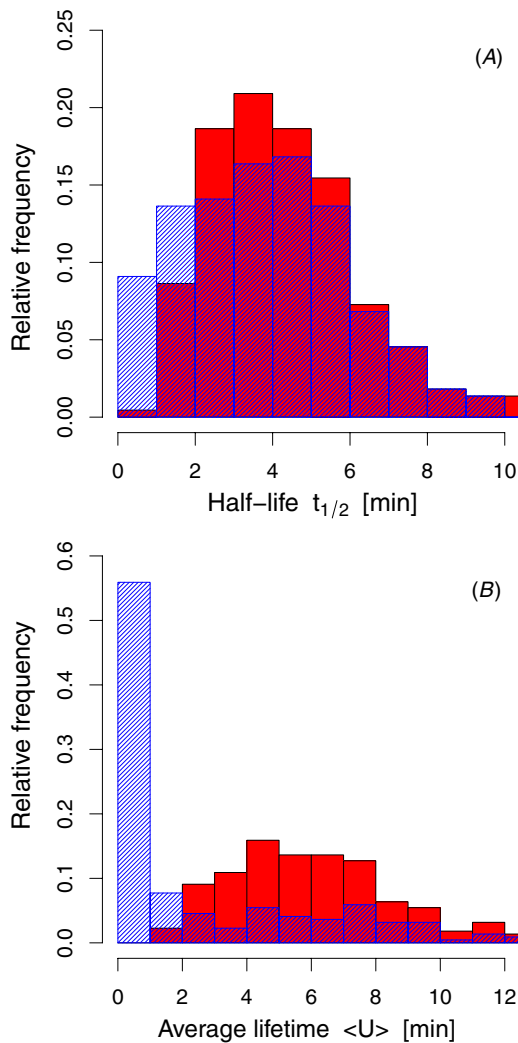
Figure 5. Distribution of half-lives $t_{1/2}$ and average lifetimes $\langle U \rangle$. (A) Comparison of the distribution of the half-lives $t_{1/2}$ when fitting the data with a purely exponential model (red) and the ribosome shielding model (blue). In the latter, the distribution is shifted to smaller times resulting in a reduced average half-life. (B) Comparison of the distribution of the average lifetimes $\langle U \rangle$ for the two models. The ribosome shielding model leads to much smaller average lifetimes because many mRNAs are degraded shortly after their synthesis.

Clearly, the fit based on ribosome shielding, equations (5) and (6) (solid lines), is much more accurate than the one with constant rate (dashed lines). The three examples represent mRNAs of different length, for which the time t_L for reaching a steady state ribosome coverage is different as well. Only for times larger than t_L, the decay follows an exponential curve with a constant rate ω_2.

In figure 4, we display the variety of decay patterns as obtained through the fitting process based on equation (5). From the 220 mRNAs analyzed here, only 74 decay exponentially (data for *E. coli* from [14]). Conversely, in most cases it was found that a steep decay at short times crosses over to a slower decay at later times, in qualitative agreement with the decay pattern expected from our shielding model. Fitting with equation (5) leads to a considerable reduction of the fitting error. Indeed, compared to a model based on a constant

decay rate we find that the mean square deviation of the entire data set under consideration decreases about threefold (not shown). Since the shape of the decay pattern is more accurately described by equation (5), one can also achieve a much more accurate estimate of the half-life $t_{1/2}$ defined in equation (3). Figure 5(A) reveals that there is a systematic decrease of the half-life $t_{1/2}$ if we apply equation (5) in comparison to a simple exponential fit. Indeed, the average half-life over the entire analyzed data set reduces from 4.40 to 3.99 min. This observation becomes intuitive when we consider figure 3 again. In two out of the three cases, the fitted decay curves produced by the shielding model result in shorter half-lives, whereas the fit with a constant rate systematically overestimates them.

Since most decay patterns are not exponential, the half-life of an mRNA alone is not sufficient to characterize these patterns. For an exponential decay law, the average lifetime $\langle U \rangle$, defined in equation (2), and the half-life $t_{1/2}$ are related via $t_{1/2} = \langle U \rangle \cdot \log(2)$. For a generic lifetime distribution, this relationship, however, does not hold. The analysis of the experimental data reveals that the shielding model yields much smaller average lifetimes $\langle U \rangle$ than a model with a constant degradation rate, as shown in figure 5(B). Indeed, the average lifetime over the whole data set reads 2.6 min, whereas the estimate for the simple exponential model is as large as 6.3 min. It is interesting to compare these small lifetimes to the timescale of ribosomal loading. It is usually thought that each mRNA contributes to the synthesis of proteins. However, only mRNAs that live longer than the timescale t_L can produce proteins. The lifetime distribution computed at t_L, $\Phi_U(t_L)$, determines the fraction of mRNAs of a given gene that are degraded until time t_L. Thereby, the explicit computation showed that for many mRNA species a considerable amount of mRNAs produces no protein at all. Hence, although an exponential fit might be able to predict a reasonably accurate half-life $t_{1/2}$, it gives a poor estimate for the average lifetime of an mRNA thus leading to a wrong estimate of the rate of protein synthesis.

One important result of the fit between our theory and the experimental data is that the degradation rate ω_1 associated with the unshielded mRNA segments is always larger than the rate ω_2 associated with the shielded mRNA segments. In fact, for the majority of the genes our fit delivers values where ω_1 differs from ω_2 by at least one order of magnitude (see figure 6). To fully appreciate this point, one should keep in mind that the fitting procedure is made without any bias concerning the values of these two rates. Whenever the fitting with the data gives two similar or identical rates, i.e., $\omega_1 = \omega_2$, one can infer that our hypothesis concerning the variation of mRNA stability during translation plays no role in the process of degradation. This would be the case if the ribosomes followed closely the RNA polymerase during the process of transcription. As shown in figure 4, only 74 of 220 mRNAs analyzed there show an exponential decrease of the mRNA amount. This means that only in a minority of the cases, the two rates ω_1 and ω_2 are identical and no effect of transient shielding by ribosomes is observable. Conversely, for the majority of the mRNAs the rate of the unshielded stretch of mRNA ω_1 is much larger than the rate of the shielded stretch, ω_2. From figure 6, it is also
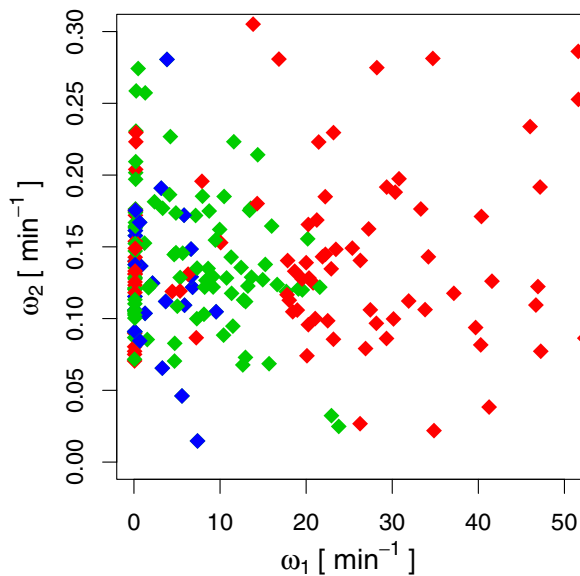
Figure 6. Gene-to-gene comparison between ω_1 and ω_2. Each diamond represents a pair of rates of a given mRNA from our present analysis. The mRNAs denoted by red diamonds are shorter than 250 codons, those with green diamonds range between 250 and 500 codons and the blue ones represent mRNAs longer than 500 codons. The value of ω_2 is widely scattered, whereas there is a hierarchy for the value of ω_1. Longer mRNAs tend to have a smaller rate ω_1. The diamonds that lie close to the *y*-axis represent the mRNAs that decay exponentially, i.e., $\omega_1 = \omega_2$.

evident that longer mRNAs tend to have smaller values of the rate ω_1. There is therefore a competition between two effects both contributing to shorten the lifetime of an mRNA. On the one hand, the longer the mRNA the longer the time during which it is unprotected. On the other hand, the larger the ω_1, the less stable is the mRNA during the transient unprotected phase. In different experiments and organisms, it was found that the length of the coding region of an mRNA and its stability are sometimes correlated [38–40, 12, 6]. Our analysis of the data from [14] reveals, however, no strong correlation between length and stability (see figure 7). This might not be too surprising because each mRNA species is regulated individually thus giving rise to the broad range of lifetimes shown in figure 5. However, if we subdivide the set of mRNA species into classes sharing the same degradation rates ω_1 and ω_2, then—within each class—we observe a clear negative correlation between the length of the coding region and mRNA average lifetime. The three mRNAs displayed in figure 3, for example, vary in length from 609 codons for the gene *atoS*, to 407 codons for the gene *fabB*, and 240 codons for the gene *ykgE*. These three examples exhibit similar degradation rates (see table 1), but the average lifetimes decrease with the mRNA length. Since for longer mRNAs the steady state is reached at a later time point t_L, the initial unprotected phase is also longer. Hence, in this case, longer mRNAs exhibit a shorter average lifetime. In figure 7, we show the length of the mRNA versus the average lifetime $\langle U \rangle$ and half-life t_h, respectively. While for the complete set of genes there is no clear pattern or correlation, among each subset of highlighted genes there is a clear negative correlation. As mentioned
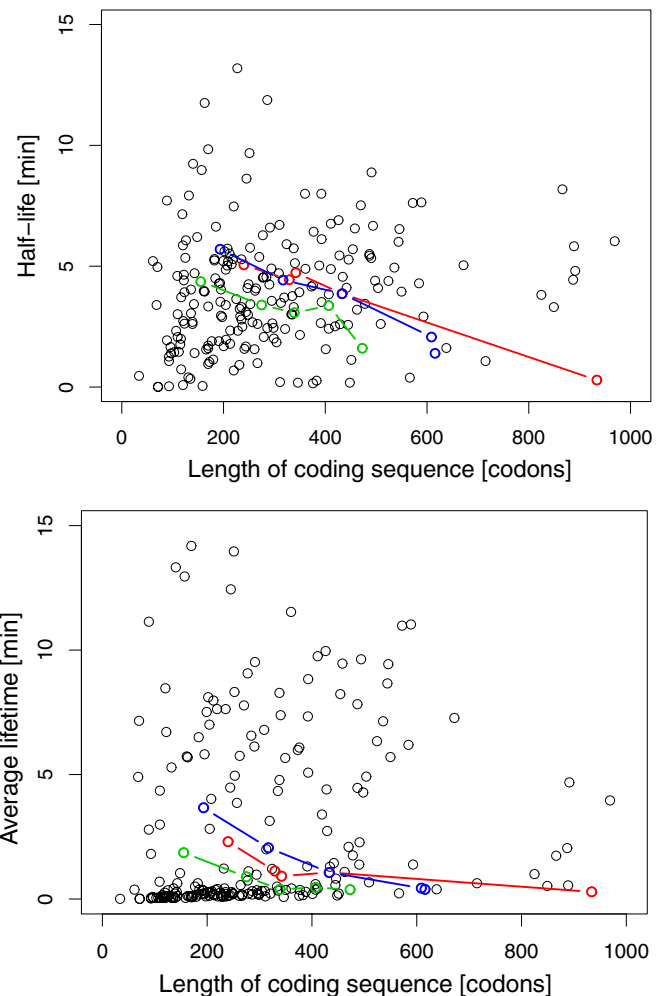

Figure 7. Half-life (top) and mean lifetime (bottom) versus length of the coding sequence for the data analyzed in figure 4. The plot reveals no strong correlation between mRNA stability and mRNA length. However, when we highlight mRNAs with similar rates ω_1 and ω_2, a negative correlation is clearly visible. This phenomenon can be easily explained within the ribosome shielding model. Longer mRNAs are unprotected for a longer period and are therefore more susceptible for degradation.

Table 1. Fitting parameters for the three mRNA species shown in figure 3. The rates ω_1 and ω_2 enter the age-dependent degradation rate defined in equation (5). The length of the mRNA determines the transient time t_L.

Gene name	Length (codons)	ω_1 (min^{-1})	ω_2 (min^{-1})	$\langle U \rangle$ (min)
atoS	609	6.6	0.13	0.44
fabB	407	8.4	0.13	0.59
ykgE	240	6.3	0.13	2.30

earlier, our interpretation is that the mRNAs within any given subset share the same parameters because they are degraded according to common mechanisms and have similar affinities to the degrading enzymes.

We have furthermore investigated if the different decay patterns are related to known gene functions. In figure 8, we assess whether genes of any of the decay category shown in figure 4 are either over- or underrepresented in 18 different
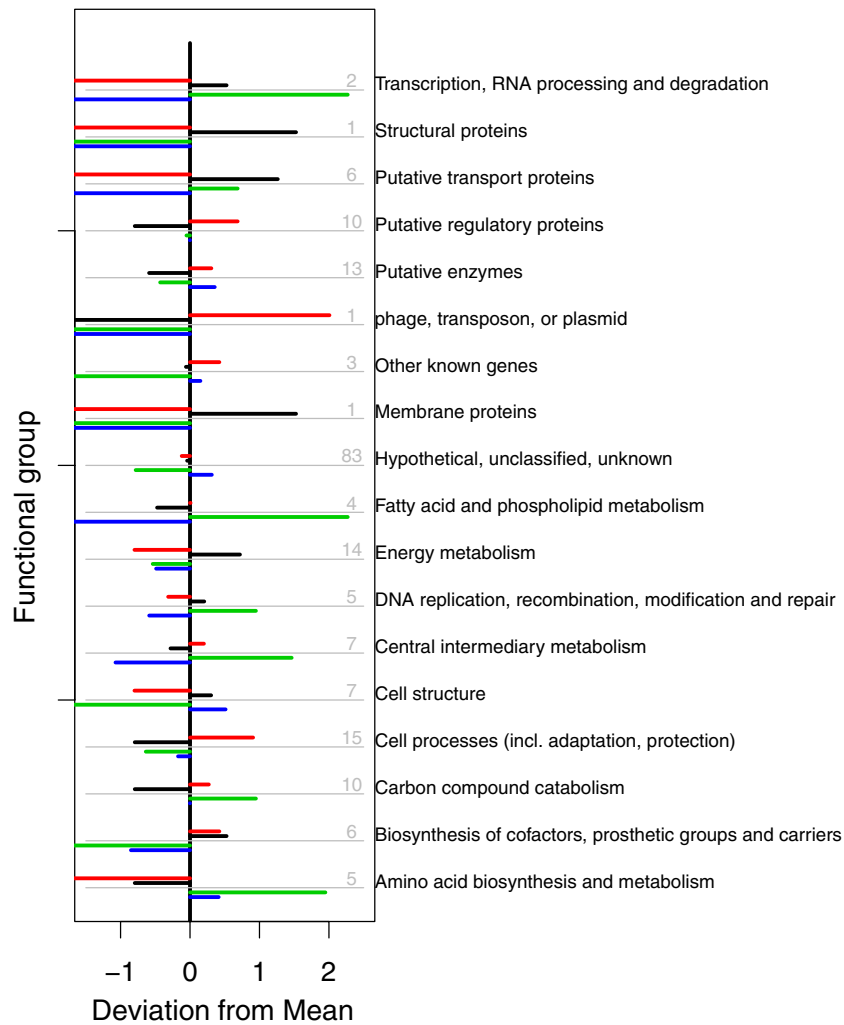
Figure 8. Distribution of the decay patterns in functional groups. The gray-shaded numbers give the numbers of mRNAs in each group. The colored bars represent mRNAs belonging to each of the four mRNA decay categories identified in figure 4. A positive (negative) deviation denotes a disproportional high (low) share of mRNAs of a given decay category in that particular functional group (see the main text). Due to the small sample size, however, these results have a poor statistical significance and are thus only indicative of possible trends.

functional groups. Therefore, we computed the deviation from the mean, defined as the logarithm of the relative frequency of mRNAs of a particular decay category in any of the functional groups. A large positive or negative deviation is observed mostly when only very few mRNAs belong to a particular functional group. The association to any gene function is therefore most likely not significant from a statistical point of view.

In applying our theory to the experimental data, we have made the assumption that the effect of the drug *rifampicin* leads to a fast global interruption of transcription. We have taken into account obvious deviations from this assumption by eliminating from our analyses those decay patterns that do not satisfy our mathematical *bona-fide* criteria determined in [25]. Nevertheless, it is reasonable to believe that transcription stops rather gradually over timescale short compared to the time of the first measurement. Our theory can be readily extended to cases where the stop of transcription is delayed (see [41, appendix C.2]). Practically, this can be achieved by rescaling the time after stop of transcription accordingly. The general

effect of such a delay would be of retarding the decay of the amount of mRNA with the consequence that the values of the degradation rate ω_1 as deduced here may be considered as lower bounds for its real values.

4. Conclusions

In this study, we have developed a theoretical model for endonucleolytic mRNA degradation that was based on the assumption that ribosomes shield the mRNA against endonucleolytic degrading processes. We have applied our model to experimental data of *E. coli* from [14] and we have shown that our description greatly improves the accuracy of the data analysis. Furthermore, we identified different classes of decay patterns where the purely exponential decay appears only in a minority of the cases.

We have assumed that during the transient time t_L the mRNA is only partially covered by ribosomes. The unbiased fit of the model with the data shows that only for about a third of

the genes there seems to be no effect of ribosome shielding. For these genes, one can assume that co-transcriptional translation might render the initial ribosome loading phase negligible. However, for all other genes the data reveal that the degradation rate for the covered region of the mRNA must be smaller than the degradation rate for the uncovered region of the mRNA. This finding leads us to conclude that co-transcriptional translation might be much less common than believed so far.

Our theory produces a more accurate estimate of the average lifetimes because it is able to describe the real behavior of the decay pattern. The values of the average lifetimes thus computed are definitively smaller than those estimated according to the standard exponential fit. This implies that a large fraction of mRNAs is degraded before being able to produce any protein. This surprising and counter-intuitive result must be taken into account when assessing the actual rate of protein synthesis and shows that post-transcriptional control may be of similar significance as transcriptional control for gene regulation in *E. coli* [2]. Moreover, this conclusion confirms the hypothesis in [7] that in *E. coli* longer mRNAs tend to produce less proteins than shorter ones.

References

[1] Alberts B (ed) 2008 *Molecular Biology of the Cell* 5th edn (New York: Taylor and Francis)

[2] Vogel C and Marcotte E M 2012 Insights into the regulation of protein abundance from proteomic and transcriptomic analyses *Nature Rev. Genet.* **13** 227–32

[3] Ingolia N T, Lareau L F and Weissman J S 2011 Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes *Cell* **147** 789–802

[4] Eugene O *et al* 2011 Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in vivo* *Cell* **147** 1295–308

[5] Gene-Wei L, Eugene O and Weissman J S 2012 The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria *Nature* **484** 538–41

[6] Valleriani A, Ignatova Z, Nagar A and Lipowsky R 2010 Turnover of messenger RNA: polysome statistics beyond the steady state *Europhys. Lett.* **89** 58003

[7] Valleriani A, Zhang G, Nagar A, Ignatova Z and Lipowsky R 2011 Length-dependent translation of messenger RNA by ribosomes *Phys. Rev.* E **83** 042903

[8] Hargrove J L and Schmidt F H 1989 The role of mRNA and protein stability in gene expression *FASEB J.* **3** 2360–70

[9] Hargrove J L, Hulsey M G, Schmidt F H and Beale E G 1990 A computer program for modeling the kinetics of gene expression *Biotechniques* **8** 654–9

[10] Ozbudak E M, Thattai M, Kurtser I, Grossman A D and van Oudenaarden A 2002 Regulation of noise in the expression of a single gene *Nature Genet.* **31** 69–73

[11] McAdams H H and Arkin A 1997 Stochastic mechanisms in gene expression *Proc. Natl Acad. Sci. USA* **94** 814–9

[12] Bernstein J A, Khodursky A B, Lin P H, Lin-Chao S and Cohen S N 2002 Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays *Proc. Natl Acad. Sci. USA* **99** 9697–702

[13] Wang Y, Liu C L, Storey J D, Tibshirani R J, Herschlag D and Brown P O 2002 Precision and functional specificity in mRNA decay *Proc. Natl Acad. Sci. USA* **99** 5860–5

[14] Selinger D W, Saxena R M, Cheung K J, Church G M and Rosenow C 2003 Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation *Genome Res.* **13** 216–23

[15] Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W and Selbach M 2011 Global quantification of mammalian gene expression control *Nature* **473** 337–42

[16] Miller C *et al* 2011 Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast *Mol. Syst. Biol.* **7** 458

[17] Steglich C, Lindell D, Futschik M, Rector T, Steen R and Chisholm S W 2010 Short RNA half-lives in the slow-growing marine cyanobacterium Prochlorococcus *Genome Biol.* **11** R54

[18] Belasco J G 2010 All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay *Nature Rev. Mol. Cell Biol.* **11** 467–78

[19] Braun F, Derout J L and Regnier P 1998 Ribosomes inhibit an RNase E cleavage which induces the decay of the *rpsO* mRNA of *Escherichia coli EMBO J.* **17** 4790–7

[20] Deana A and Belasco J G 2005 Lost in translation: the influence of ribosomes on bacterial mRNA decay *Genes Dev.* **19** 2526–33

[21] Pedersen M, Nissen S, Mitarai N, Lo Svenningsen S, Sneppen K and Pedersen S 2011 The functional half-life of an mRNA depends on the ribosome spacing in an early coding region *J. Mol. Biol.* **407** 35–44

[22] Nagar A, Valleriani A and Lipowsky R 2011 Translation by ribosomes with mRNA degradation: exclusion processes on aging tracks *J. Stat. Phys.* **145** 1385–404

[23] Mitarai N, Sneppen K and Pedersen S 2008 Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization *J. Mol. Biol.* **382** 236–45

[24] Deneke C, Rudorf S and Valleriani A 2012 Transient phenomena in gene expression after induction of transcription *PLoS ONE* **7** e35044

[25] Deneke C, Lipowsky R and Valleriani A 2013 Complex degradation processes lead to non-exponential decay patterns and age-dependent decay rates of messenger RNA *PLoS ONE* **8** e55442

[26] Morse D E and Yanofsky C 1969 Polarity and the degradation of mRNA *Nature* **224** 329–31

[27] Belasco J G and Higgins C F 1988 Mechanisms of mRNA decay in bacteria: a perspective *Gene* **72** 15–23

[28] Stent G S 1966 Genetic transcription *Proc. R. Soc. Lond.* **164** 181–97

[29] Taylor H M and Karlin S 1999 *An Introduction to Stochastic Modeling* 3rd edn (San Diego, CA: Academic)

[30] Arava Y, Wang Y, Storey J D, Liu C L, Brown P O and Herschlag D 2003 Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae Proc. Natl Acad. Sci. USA* **100** 3889–94

[31] Shaw L B, Zia R K P and Lee K H 2003 Totally asymmetric exclusion process with extended objects: a model for protein synthesis *Phys. Rev.* E **68** 021910

[32] Dong J J, Schmittmann B and Zia R K P 2007 Inhomogeneous exclusion processes with extended objects: the effect of defect locations *Phys. Rev.* E **76** 051113

[33] Romano M C, Thiel M, Stansfield I and Grebogi C 2009 Queueing phase transition: theory of translation *Phys. Rev. Lett.* **102** 198104

[34] Reuveni S, Meilijson I, Kupiec M, Ruppin E and Tuller T 2011 Genome-scale analysis of translation elongation with a ribosome flow model *PLoS Comput. Biol.* **7** e1002127

[35] Gillespie D T 1976 General method for numerically simulating stochastic time evolution of coupled chemical-reactions *J. Comput. Phys.* **22** 403–34

[36] Neidhardt F C (ed) 1987 *Escherichia Coli and Salmonella Typhimurium* (Washington, DC: American Society for Microbiology)

[37] Sørensen M A, Kadsurland C G and Pedersen S 1989 Codon usage determines translation rate in *Escherichia coli J. Mol. Biol.* **207** 365–77

[38] Feng L and Niu D K 2007 Relationship between mRNA stability and length: an old question with a new twist *Biochem. Genet.* **45** 131–7

[39] Andersson A F, Lundgren M, Eriksson S, Rosenlund M, Bernander R and Nilsson P 2006 Global analysis of mRNA stability in the archaeon Sulfolobus *Genome Biol.* **7** R99

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

[41] Deneke C 2012 Theory of mRNA degradation *PhD Thesis* University of Potsdam