In silico analysis of division times of Escherichia coli populations as a function of the partitioning scheme of non-functional proteins

Abhishekh Gupta, Jason Lloyd-Price and Andre S. Ribeiro*
Laboratory of Biosystem Dynamics, Department of Signal Processing, Tampere University of Technology, Tampere, Finland

Received 6 August 2014
Revised 23 September 2014
Accepted 30 September 2014

Abstract. Recent evidence suggests that cells employ functionally asymmetric partitioning schemes in division to cope with aging. We explore various schemes in silico, with a stochastic model of Escherichia coli that includes gene expression, non-functional proteins generation, aggregation and polar retention, and molecule partitioning in division. The model is implemented in SGNS2, which allows stochastic, multi-delayed reactions within hierarchical, transient, interlinked compartments. After setting parameter values of non-functional proteins' generation and effects that reproduce realistic intracellular and population dynamics, we investigate how the spatial organization of non-functional proteins affects mean division times of cell populations in lineages and, thus, mean cell numbers over time. We find that division times decrease for increasingly asymmetric partitioning. Also, increasing the clustering of non-functional proteins decreases division times. Increasing the bias in polar segregation further decreases division times, particularly if the bias favors the older pole and aggregates' polar retention is robust. Finally, we show that the non-energy consuming retention of inherited non-functional proteins at the older pole via nucleoid occlusion is a source of functional asymmetries and, thus, is advantageous. Our results suggest that the mechanisms of intracellular organization of non-functional proteins, including clustering and polar retention, affect the vitality of E. coli populations.

Keywords: Non-functional proteins, partitioning schemes, functional asymmetries, division times, In silico models

1. Introduction

The mechanisms employed by Escherichia coli in protein production and maintenance [60] are not free from a number of errors, e.g., during folding and activation, particularly when cells are stressed, or following mutations or overexpression [40]. Consequently, there are several non-functional proteins present in a cell at any given moment. How many is unknown but these numbers should depend on the environment and differ between proteins. The excessive accumulation of non-functional proteins has been linked to decreasing
growth rates, not only in *E. coli* [14, 18, 36, 44] but in other organisms as well [15, 25, 50]. Not surprisingly, *E. coli* has evolved a complex machinery to enhance proteins’ functionality. For example, chaperones both catalyze the proper folding of proteins, which prevents aggregation as it assists in reaching stable structures, as well as participate in the rescue of misfolded proteins [6, 60]. Under optimal growth conditions, at least 10% to 20% of newly formed polypeptides appear to be associated to chaperones, such as GroEL and DnaK [10, 16, 53].

When these rescue mechanisms fail, some misfolded proteins can still be targeted for destruction by the protease cell machinery [21, 59]. Evidence suggests that approximately 20% of newly synthesized polypeptides are degraded [20, 38]. It is reasonable to assume that this number roughly corresponds to the percentage of polypeptides that are non-functional following their production. However, this is likely to be an overestimation as a certain fraction of proteins is likely to be degraded at all times, even in the absence of non-functional proteins, to ensure the existence of raw material for the production of novel proteins [20]. Evidence suggests that as much as 10% of the proteins can be degraded per hour in non-growing bacteria [19, 37, 45, 61].

When the above strategies fail, bacteria still have additional strategies to cope with non-functional proteins [54]. For example, misfolded polypeptides of some proteins form aggregates [22, 32, 58, 62]. Relevantly, the process of formation of these aggregates exhibits some similarities to events in eukaryotic cells that have been linked to the emergence of diseases such as Huntington’s, Alzheimer’s, and Parkinson’s [5, 11, 24, 41, 56].

Once formed, some of these aggregates can be gathered into one or two inclusion bodies, which tend to locate at the cell poles, particularly in cells that are under heat or oxidative stress [22, 32]. While these inclusion bodies are stable structures, there is evidence that the process of inclusion of damaged proteins is not necessarily irreversible [6]. Interestingly, the process of aggregate formation appears to be protein-specific [51] and under regulation (e.g., in *E. coli* grpE280 mutant aggregates are small-sized [32]).

The partitioning in division of a specific non-functional polypeptide is likely related to its long-term spatial distribution. For example, if all non-functional polypeptides of a given protein are gathered into a single inclusion body, only one of the daughter cells will inherit them. On the other hand, if the non-functional polypeptides do not aggregate, one could expect them to be less retained at the poles and thus, to be more randomly distributed. As such, it would be more likely that such polypeptides are partitioned by the daughter cells in an unbiased fashion. It is also possible that the partitioning scheme will depend on the location of the functional form of the proteins, provided that the loss of function does not affect the long-term spatio-temporal distribution. The process of partitioning in division of non-functional polypeptides is thus likely to be under regulation, either by direct or indirect mechanisms, and to differ between proteins. This is supported by recent evidence that some protein aggregates, due to polar retention and cell division, accumulate primarily at the older pole of the cells [23, 34].

Here, we explore this issue *in silico*. For that, we developed a stochastic model that includes cell division and the partitioning of intracellular molecules, along with a delayed stochastic model of gene expression that includes the generation of both functional and non-functional proteins. First, we show that, using parameter values extracted from measurements, the model is capable of emulating the reduced division times as a function of the increasing number of non-functional proteins in the cells, over several generations [52]. Next, we investigate how different partitioning schemes of non-functional proteins in division affect cell division times.

2. Results & discussion

2.1. Tuning the effects of non-functional proteins on the growth rate of model cell populations

After setting the parameter values (Table 1) of the model (see Methods), we first studied how the growth rate of the population is affected by β, the number of non-functional proteins necessary for the cell growth rate to be halved, and by n, the exponent of f(Q), the function regulating the degree to which the growth rate is reduced by the presence of non-functional proteins (Equation (8)). Here, n determines how fast the cell growth rate transitions from high to low with increasing numbers of non-functional proteins in the cells. For simplicity, we assume that the environment poses no growth constraints (neither on the rate of growth nor on the total number of cells). At moment zero, each population contains one cell.
By inspection, we verified that decreasing $\beta$ increases the mean division time, as expected. To exemplify this, in Fig. 1 (left) we show the number of cells over time for three values of $\beta$. Fig. 1 (right) shows the corresponding distributions of division times. In these simulations, $n$ was arbitrarily set to 1.

Though the mother cell of each lineage is initiated without non-functional proteins, these will accumulate over time as they are produced. Meanwhile, cell division should act as a decay rate (by 'dilution') on the number of non-functional proteins in individual cells. Thus, we expect the mean division time to differ with the cell generation, until reaching quasi-equilibrium.

The dependence of the division time on the presence of non-functional proteins should differ with $\beta$. Higher $\beta$ should increase the number of generations required for the effects of aging on division times to become visible (and to reach quasi-equilibrium).

This can be seen in example Fig. 2, where we plot the division time of individual cells of each generation, for each value of $\beta$, up to 6 generations. Note that, for $\beta = 10$, the effects of aging are visible already in the first generation (much longer division time than for the other values of $\beta$) and are maximized approximately from the fourth generation onwards. Meanwhile, for $\beta$ equal to 35 and 50, the effects of aging (i.e. the increase of the division time with the generation) appear more gradually, and are (approximately) maximized only in generations 5 and 6, respectively.

From here onwards, we arbitrarily set $\beta$ to 35 as this value (and similar ones) allows the accumulation of non-functional proteins to affect cell division times in a manner that is consistent with observations (see e.g. [34]), whereas much lower values disrupt too strongly cellular well-functioning and much higher values do not affect cell growth significantly.

Next, we study the effects of $n$ on the cell growth rate. Different non-functional polypeptides are expected to have different effects on the cell growth rate, to differ in their degree of aggregation, etc. Also, previous observations [34, 52, 62] suggest that, starting from a single healthy cell, only after a few generations do the effects of the accumulation of non-functional proteins on division times become visible. This can be captured by the model by tuning $n$. From Fig. 3, as $n$ increases, cell division rate increases. This is due to the fact that, for high values of $\beta$, $n$ increases, more non-functional proteins need to accumulate in a cell to tangibly affect its division time. Nevertheless, it is worth noting that no qualitative changes are observed as $n$ is increased. As such, for simplicity, from here onwards we set $n$ to 1.

### Table 1
Parameters used in the model, unless stated otherwise

| Parameter | Description | Value | Source |
|-----------|-------------|-------|--------|
| $t_{div}$ | Value set to the mean division time of E. coli DH5/PRO | 3600 s | [29] |
| $k_{cc}$ | Closed complex formation | $1/574$ s$^{-1}$ | [29] |
| $k_{oc}$ | Open complex formation | $1/391$ s$^{-1}$ | [29] |
| $k_{tr}$ | Translation rate | $1/3$ s$^{-1}$ | [43] |
| $\tau_p$ | Mean of time delay for protein production | 420 s | [43] |
| $\sigma_p$ | Std. dev. of time delay for protein production | 300 s | [43] |
| $k_{drbs}$ | Degradation rate of RNA | $0.011$ s$^{-1}$ | [43] |
| $k_{dp}$ | Degradation rate of proteins | $3.9 \times 10^{-5}$ s$^{-1}$ | [43] |
| $k_{D}$ | Rate at which functional proteins become non-functional | $1.9 \times 10^{-5}$ s$^{-1}$ | [9] |

![Fig. 1. (Left) Mean number of cells over time for different values of $\beta$. As $\beta$ is lowered, the number of cells drastically decreases. Also shown (right) are the corresponding distributions of division times.](image-url)
Fig. 2. Mean division time of cells as a function of their generation and of $\beta$, the number of non-functional proteins necessary for the cell growth rate to be reduced to half the maximum rate.

Fig. 3. Effect of changing parameter $n$ that models the effect of non-functional proteins on the growth rate.

Using these values of $\beta$ and $n$, the model is able to emulate (see example Fig. 4) empirical distributions of division times in near-optimal growth conditions (see e.g. [52]). Next, we use the model to study the effects of different partitioning schemes of non-functional proteins on the population dynamics.

2.2. Effects of different partitioning schemes

We investigate the effects of different partitioning schemes of non-functional proteins in division. In particular, we consider six schemes [26] (schematic representations in Fig. 5, Top): (i) perfect partitioning, where individual non-functional proteins are equally partitioned between the daughter cells, (ii) pair formation, where non-functional proteins form pairs, which are split evenly into the two daughter cells, while the remaining individuals are independently partitioned, (iii) random size partitioning, where the non-functional proteins are assumed to be uniformly distributed in the mother cell, and the daughter cells differ significantly in size (iv) preferential partitioning, where individual non-functional proteins are partitioned according to a biased binomial distribution, (v) cluster formation, where non-functional proteins form clusters of variable size which are then independently partitioned, and (vi) all or nothing partitioning, where all non-functional proteins are inherited by only one of the daughter cells.

Using the parameter values in Table 1, for each partitioning scheme, we simulated 1 cell lineage for $6 \times t_{div}$ (i.e. $6$ times the expected division time). This allows, on average, 6 cell generations to be observed. Aside from these parameters, we have...
Fig. 5. (Top) Different partitioning schemes. (i) Perfect (ii) Pair formation (iii) Random size (iv) Preferential (v) Clustered (vi) All or Nothing. Also shown (bottom) are the non-functional protein distributions resulting from each of the partitioning schemes.

used the following additional set of parameter values (necessary to implement some of the partitioning schemes). For pair formation, we set the fraction of non-functional proteins which form pairs to 0.25. In random size partitioning, the position of the division point of the daughter cells was drawn from a Beta distribution with $\alpha = \beta = 2$. In preferential partitioning, the bias was set to 0.25 (i.e., the cell inheriting the older pole retains, on average, 75% of the non-functional proteins produced by its mother cell). Finally, in the
cluster formation scheme, we set the degree of clustering, $k_{cl}$, to 0.03.

As stated in the methods section, inherited non-functional proteins are not subject to these schemes, only novel ones are. This assumes perfect retention of inherited non-functional proteins at the older pole of the cell, which, while not being exact, is a close approximation to observations [23].

For each partitioning scheme, we obtained the number of non-functional proteins in each cell at the end of the simulation time. Distributions are shown in Fig. 5 (Bottom). Table 2 shows the means and standard deviations of these distributions. As expected, the standard deviation increases from a perfectly symmetric to a perfectly asymmetric partitioning scheme. Meanwhile, as protein production rates (functional and non-functional) are identical in all models, mean numbers of non-functional protein are necessarily also identical.

To investigate if the differences in the standard deviation generate different mean cell division times in cells with different partitioning schemes, we performed 1000 simulations of the model for each partitioning scheme. Each simulation was performed from one initial cell, for 15 times the expected division time (i.e., $15 \times t_{div}$). For each scheme, we obtained the mean increase in the number of cells in the last 30 minutes of the simulation, to eliminate the effects of the transient observed in Fig. 2. Also, we obtained the division times of all cells that divided during the simulation time, from which we extracted mean and standard deviation. Results are shown in Table 3.

From Table 3, increasing the degree of asymmetry in the partitioning of novel non-functional proteins in division decreases the mean division time, while not significantly affecting the variance (except in the all-or-nothing case, where a significant decrease is observed).

### Table 2

Mean and standard deviation of the distributions of the number of non-functional proteins in individual cells (one lineage, after 360 min, or approximately 6 generations), for different partitioning schemes

| Partitioning Scheme | Mean no. of non-functional proteins | Standard deviation of the no. of non-functional proteins |
|---------------------|------------------------------------|--------------------------------------------------------|
| Perfect             | 15.4                               | 7.1                                                    |
| Pair                | 15.4                               | 7.8                                                    |
| Random Size         | 15.4                               | 10.4                                                   |
| Preferential        | 15.3                               | 10.9                                                   |
| Clustered           | 15.4                               | 12.1                                                   |
| All or Nothing      | 15.4                               | 20.4                                                   |

### Table 3

Number of cells at the end of the simulation time and mean and standard deviation of the division time of individual cells in the various partitioning schemes

| Partitioning Scheme | Mean increase in no. of cells at the last 30 mins of simulation time | Mean division time (m) | Std of division time (m) |
|---------------------|-----------------------------------------------------------------------|------------------------|--------------------------|
| Perfect             | 104.7                                                                 | 114.2                  | 42.6                     |
| Pair                | 107.0                                                                 | 113.7                  | 42.6                     |
| Random Size         | 147.7                                                                 | 105.7                  | 41.9                     |
| Preferential        | 163.5                                                                 | 103.8                  | 41.1                     |
| Clustered           | 189.9                                                                 | 100.6                  | 41.5                     |
| All or Nothing      | 402.9                                                                 | 85.8                   | 36.6                     |

### Table 4

Number of cells at the end of the simulation time and mean and standard deviation of the division time of individual cells in lineages subject to the clustering partitioning scheme, for different degrees of clustering ($k_{cl}$)

| $k_{cl}$ | Mean increase in no. of cells at the last 30 mins of simulation time | Mean division time (m) | Std of division time (m) |
|----------|-----------------------------------------------------------------------|------------------------|--------------------------|
| 0.01     | 337.5                                                                 | 109.2                  | 38.7                     |
| 0.03     | 189.9                                                                 | 107.0                  | 41.5                     |
| 0.05     | 155.6                                                                 | 105.1                  | 42.0                     |
| 0.07     | 140.1                                                                 | 107.4                  | 42.1                     |
| 0.09     | 131.7                                                                 | 108.8                  | 42.2                     |

2.3. Effects of the degree of clustering and of preferential partitioning

We next investigate two partitioning schemes in more detail. Clustered and Preferential, since these are employed by *E. coli*. For example, proteins such as Tsr form clusters at the poles [64]. Further, in some, the proteins preferentially locate at the older cell pole (e.g. IbpA [62]).

First, to study how the clustering rate of non-functional proteins affects division times, we ran simulations setting the clustering parameter, $k_{cl}$, to values between 0.01 and 0.09. Note that higher $k_{cl}$ produces a weaker ‘degree of clustering’. E.g., for $k_{cl} = 0.01$, there are very few clusters, with each necessarily including many non-functional proteins, while for $k_{cl} = 0.09$ the number of clusters is higher, each with fewer non-functional proteins.

From 1000 independent lineages simulated for $15 \times t_{div}$ s (each starting from a single cell), we extracted the mean increase in the number of cells in the last 30 minutes of the simulations and the cells’ division times. Results are shown in Table 4. Visibly, both the mean and the standard deviation of the division time
increase with increasing $k_{2}$, i.e., decreasing degree of clustering. These results confirm that increasing the degree of asymmetry in partitioning of non-functional proteins (in this case via clustering) decreases mean division times. Interestingly, these results agree with reports that the fusion of protein aggregates supports population survival in yeast [7].

Next, we tested the effects of preferential polar segregation and partitioning in division of non-functional proteins produced by the cell (again, inherited proteins are not affected, remaining at the old pole). For this, we introduced an additional parameter in the model, so as to regulate the degree of bias in partitioning, named ‘Bias’, and ran the model for values of Bias from 0 to 1. For Bias equal to 0.5, new non-functional proteins will be partitioned unbiasedly and independently between daughter cells. Also, according to this scheme, e.g., for a bias of 0.1, each non-functional protein has a 90% chance to be inherited by the daughter cell inheriting the older pole of the mother cell. Finally, note that since inherited proteins will remain at the old pole throughout a cell’s lifetime, divisions will unavoidably introduce asymmetries. Because of this, a Bias of ∼0.7 (and not 0.5) is expected to result in an unbiased partitioning in division of all non-functional proteins of the mother cell.

We obtained the mean increase in the number of cells in the last 30 minutes of the simulations, as well as the mean and standard deviation of the division time of the cells from 1000 independent lineages, for $15 \times t_{div}$, 5. The results are shown in Table 5. First, the mean number of cells at the end of the simulation is smallest for a Bias of 0.7, since this leads to more symmetry in partitioning in division than setting Bias to 0.5.

Another interesting observation is the difference in mean division time between biases towards older and newer cell poles. In particular, one observes that cells that bias the partitioning of novel non-functional proteins towards the older pole, where the inherited non-functional proteins are located, have faster division times. This is because this asymmetry in pole choice maximizes the asymmetry in the numbers of non-functional proteins (inherited plus produced) at the poles.

To validate this further, we loosened the condition of polar retention of inherited non-functional proteins. For this, we replace the constant ‘Retention’ by a parameter, named ‘R’, which can range from 0 to 1. For R = 1, all inherited non-functional proteins remain retained at the old pole. Decreasing R increases the number of inherited proteins that escape the old pole during a cell’s lifetime. If escaping, they ‘merge’ with the population of newly formed non-functional proteins at midcell, and then partitioned in division accordingly. For R = 0, all inherited non-functional proteins are partitioned following the same scheme as produced ones, which controlled by the parameter Bias.

We extracted, as a measure of cell growth rate, the mean increase in number of cells during the last 30 min of each simulation for different values of Bias and R. The results, averaged from 1000 simulations per condition (each $(15 \times t_{div}$, s long), are shown in Fig. 6. The mean increase in number of cells is maximized for a Bias of 0 (which favors the old pole), regardless of the value of R.

From the above, we conclude that the most advantageous solution for cells is to place all non-functional proteins, produced and inherited, into the older pole. Note that, combining a Bias of 1 (all new non-functional proteins move to the new pole) with an R of 0 (no inherited protein is retained at the poles), one attains similar growth rates as for a Bias of 0 (since all proteins, inherited and produced, will be placed at the new pole, creating a full bias). However, we do not expect this latter scheme to be found in nature (at least, not commonly) since it would likely require an energy-consuming mechanism to be implemented to ensure that inherited proteins moved to the newer pole, while the other option is less energy consuming.

Table 5

| Bias | Mean increase in no. of cells at the last 30 min of simulation time | Mean Division time (m) | Standard deviation of division time (m) |
|------|------------------------------------------------------------|-------------------|----------------------------------|
| 0.00 | 888.2                                                      | 75.0              | 31.4                             |
| 0.10 | 572.3                                                      | 80.7              | 33.9                             |
| 0.30 | 238.8                                                      | 95.4              | 39.3                             |
| 0.50 | 109.1                                                      | 123.8             | 42.3                             |
| 0.70 | 95.9                                                       | 118.0             | 41.6                             |
| 0.90 | 217.3                                                      | 99.2              | 38.6                             |
| 1.00 | 384.9                                                      | 88.0              | 35.8                             |
division of the population, leading to higher cell numbers. In particular, in the present case, we observe an increase of 10.2% in the number of cells that appear in the last 30 min when increasing $R$ from 0 to 0.5, and an increase of 34.1% when increasing $R$ from 0 to 1.

Also visible in the inset are the points at which the various lines cross. From these, one can observe that the Bias at which the crossing occurs depends on the values of $R$. Visibly, the higher the difference in values of $R$ of the two models, the higher is the value of Bias at which these models exhibit equal cell division rates.

3. Conclusion

We explored, in silico, using a stochastic model and parameter values extracted from measurements (Table 1), possible selective advantages of different partitioning schemes of non-functional proteins in E. coli. Assuming that the accumulation of non-functional proteins is harmful by decreasing cell division rate, we observed that increasing the degree of asymmetry in their partitioning in division increases the overall division rates of cell lineages.

Relevantly, a source for such asymmetries in E. coli has been identified. Namely, the nucleoid at midcell is denser than the cytoplasm [55]. Evidence suggests that its presence appears to force protein aggregates to travel to and then remain at the poles, by generating anisotropies in the borders between nucleoid and cytoplasm that favor the motion towards the poles [23]. Cell divisions subsequently introduce asymmetries between the number of such non-functional proteins at the old and new poles of the daughter cells, favoring the older pole. The latter was observed to occur to aggregates of non-functional proteins as well [8, 23, 34, 62].

These observations also inform on the means used by cells to differentiate between functional and non-functional proteins. Namely, as the segregation is based on occlusion, one natural means of differentiation is to aggregate the unwanted, non-functioning proteins (and only these), into large clusters, thus ensuring segregation from midcell. Such aggregation has been observed in live cells, and is likely enhanced by chaperones such as IbpA [34]. Here, we further observed that increasing the efficiency of the process of clustering enhances the generation of asymmetries between the numbers of non-functional proteins at the two poles of a cell. As such, we expect the cellular mechanisms of detection and aggregation of non-functional proteins to be under selective pressure for efficiency.

To further verify that the generation of asymmetries in non-functional protein numbers across the cell populations enhances the mean division rate within cell lineages, we tested combining retention of inherited proteins at the old pole with a mechanism able to force newly formed non-functional proteins to move into the old pole as well, prior to cell division. In agreement with the hypothesis, division rates further increased. However, there is no known mechanism of selective transport of unwanted protein aggregates to a specific cell pole (old or new) in E. coli. Also, there is no evidence for asymmetries in the choice of pole (see e.g. [23]). It may be that the division rate of E. coli renders such selective, energy-consuming, transport mechanism not selectively advantageous.

Finally, we tested the effects of removing the polar retention mechanism. Overall, the test showed that this mechanism alone is selectively advantageous, as expected, given that it is a non-energy consuming source of functional asymmetries. Interestingly, its effects could be combined with partitioning schemes capable of enhancing asymmetries in the number of non-functional proteins at the old and new pole of the cells to further enhance the mean rate of cell divisions.

It is worth noting that our model assumes that the accumulation of damage in cells (in this case,
non-functional proteins) leads to a gradual increase in division times, as suggested by previous studies [33, 34, 36, 44, 52]. However, one recent study suggests that the accumulation of damage leads only to death, rather than affecting growth rates gradually [57]. In this study, death occurs when a lethal “factor” in a cell crosses a threshold. This factor was hypothesized to correspond to non-functional protein aggregates [57]. We expect that, using such a model, rather than a change in mean division times and number of cells at the end of the simulation time as a function of, e.g., the partitioning scheme, one would observe only the changes in the number of cells. These changes should be qualitatively similar to those reported here for high

$n$, in the number of cells at the end of the simulation time as a function of the partitioning scheme, the bias in partitioning and, the degree of clustering.

There are several means by which non-functional proteins can interfere with the normal cellular functionality. For example, reduced ribosomal fidelity due to damage accumulation in the protein assembly machinery [3] would not produce functional proteins but would still waste resources, reducing growth rates [49]. As such, it is not surprising that cells evolved mechanisms to handle non-functional proteins, such as polar retention [62]. This retention, among other things, prevents these proteins from reaching binding sites in the DNA, which, in the case of non-functional transcription factors, could lead to harmful interference with normal gene expression dynamics [9].

While E. coli has a morphologically symmetric division process, and thus we expect our results to be valid for other such organisms [1, 2, 4, 15], we also expect the results to be applicable to species with morphologically asymmetric division processes, such as budding yeast [27, 28, 30]. In the case of morphologically asymmetric divisions, if the daughter cell is smaller (e.g. in Saccharomyces cerevisiae), non-functional proteins and other harmful substances should be easily retained in the mother cell. However, if the daughter cell was to be larger, transport mechanisms may be required. We hypothesize that this is one of many factors that has led to evolutionary processes resulting in the production of daughter cells that are smaller than the mother cells.

In the future, it would be of interest to investigate sources of asymmetries in eukaryotic cells as their more complex internal structure and mechanisms may allow more ingenious means to cope with the effects of aging, among other.

4. Methods

4.1. Simulator

Simulations were performed by SGNS2 [35], a simulator of chemical reaction systems whose dynamics are driven according to the delayed Stochastic Simulation Algorithm [48]. While based on the Stochastic Simulation Algorithm (SSA) [17], the delayed SSA differs in that it allows for multi-delayed reaction events. The main advantage of SGNS2, which was developed from SGN Sim [46], is that it allows multi-delayed reactions to be implemented within hierarchical, interlinked compartments that can be created, destroyed and divided during the simulation. In cell divisions, molecules are randomly segregated into the daughter cells following a specified distribution corresponding to one of several partitioning schemes, applicable on a per-molecule-type basis. As such, SGNS2 allows easy implementation and simulation of different models of partitioning of proteins (functional and/or non-functional) in division.

4.2. Modeling gene expression, cell growth and division, and long-term spatial distributions of non-functional proteins

Our model of a “cell” has four main components: (i) a gene expression mechanism; (ii) a mechanism of generating non-functional proteins; (iii) a mechanism of cell growth and division, including partitioning schemes of molecules and effects of non-functional proteins on the growth rate; and, (iv) a mechanism of polar retention of non-functional proteins.

The first component, the model gene expression mechanism, was proposed in [47] and validated in [65], using measurements of protein production in E. coli cells with single molecule sensitivity [63]. More recently, this model, in particular, reactions (1) and (2), were shown to also capture with accuracy the in vivo kinetics of RNA production, when measured with single-molecule sensitivity by MS2-GFP tagging [29, 42]. This component includes reactions (1)-(5):

\[
\text{ProC} \overset{k_1}{\rightarrow} \text{Pro} \cdot \text{RNAPC}
\]

(1)

\[
\text{Pro} \cdot \text{RNAPC} \overset{k_2}{\rightarrow} \text{ProC} + \text{RBSC}
\]

(2)

\[
\text{RBSC} \overset{k_3}{\rightarrow} \text{RBSC} + \text{Pe}(\text{normal} : \tau_p, \sigma_p)
\]

(3)
values set for the open complex formation are accounted for by the time lengths of the closed complex formation and of some binding site region of the RNA (‘escape by the RNAP and the formation of the ribo-plex, from the closed complex, along with the promoter Reaction (2) models the formation of the open complex (‘Pro

\[ \text{RBS} \xrightarrow{k_{oc}} \text{RNAP} \]

\[ \text{Pro} \xrightarrow{k_{cc}} \text{ribosome} \]

Here, C denotes the index of the cell in which the reaction is taking place. Reaction (1) models the finding by an RNA polymerase (RNAP) of the transcription start site (‘Pro’, located in the promoter region) and the formation of the closed complex (‘Pro-RNAP’). Reaction (2) models the formation of the open complex, from the closed complex, along with the promoter escape by the RNAP and the formation of the ribosome binding site region of the RNA (‘RBS’). The time lengths of the closed complex formation and of the open complex formation are accounted for by the values set for \( k_{oc} \) and \( k_{cc} \), respectively [29]. Also, for simplicity, the RNAP is not explicitly modeled in the left side of reaction (1), since it is assumed that the concentration of free RNAP in the cell does not change significantly during a cell’s lifetime. As such, its release is not represented in the right side of reaction (2).

Reaction (3) models the translation of a protein \( P \) along with its folding and activation [65]. These latter events are accounted for by the delay \( \tau_{P} \) in the release of the protein. Note that the ribosome is not explicitly modeled, for the same reason that the RNAP is not explicitly modeled. Note also that this reaction can initiate as soon as a RBS is formed [39]. As such, and since our model does not include either transcription or translation elongation at the nucleotide level, it is not necessary to represent complete RNA molecules explicitly. Finally, reactions (4) and (5) model the degradation of RNA (by degradation of its RBS region) and proteins, respectively, as first-order events [65].

The second component of the model is responsible for generating non-functional proteins, ‘\( Q \)’, from functional proteins:

\[ \text{PC} \xrightarrow{k_{D}} \text{QC} \]

According to this model, the fraction of non-functional proteins in a cell is controlled by the rate \( k_{D} \) [9]. As noted in the introduction, the percentage of non-functional proteins in a cell at any given time is unknown. Since approximately 20% of formed polypeptides are non-functional and given the existence of several error-correction mechanisms, we chose to set \( k_{D} \), the rate at which functional proteins become non-functional, to \( 1.9 \times 10^{-5} \, \text{s}^{-1} \) (see Table 1), since we observed by inspection that, with this value, usually 1% to 5% of the proteins will become non-functional during the lifetime of a cell. Finally, note that these non-functional proteins, unlike functional ones, do not degrade (as we assume that they ‘survived’ the error correction and subsequent degradation mechanisms).

The third component of the model consists of cell growth, division and partitioning schemes of proteins and RNA molecules in division. As such, this component needs to account for the effects of non-functional proteins on a cell’s division rate. Division and partitioning of cellular components are not represented in the form of chemical reactions (see below).

In optimal conditions, the moment of division of an E. coli cell appears to be strongly correlated with reaching a specific cell length [12, 13], and there is a very small variance in cells’ lifetime. The division process is therefore considered to be largely deterministic [31]. As such, we model division as an instantaneous event that occurs once the cell reaches a specific length. When a division occurs, the DNA (i.e. the promoter region, \( \text{Pro} \)) is replicated and one copy is placed in each daughter cell. If the RNAP has formed the closed complex (i.e. Pro-RNAP,C) at the point of division it is inherited by the cell with the older pole while the cell inheriting the newer pole inherits a free promoter. This arbitrary choice does not affect our conclusions as it only affects (very mildly) the protein numbers in each cell.

Meanwhile, RBS and P molecules are partitioned between the daughter cells according to an unbiased, independent partitioning scheme (unless stated otherwise), resulting in a binomial distribution of molecules inherited by a given daughter cell. We use reaction (7) to simulate cell growth and the effect of non-functional proteins on this growth rate. This reaction controls the quantity ‘cell length’ (denoted \( l_{C} \), where \( C \) is the cell ID).

\[ \text{QC} \xrightarrow{t_{div} \times \text{f}(Q)} \text{PC} \]

In reaction (7), the rate constant \( l_{C} \times \text{f}(Q) \) (where \( t_{div} \) is the division time in the absence of non-functional proteins) is multiplied by a hill function, \( \text{f}(Q) \). This function ranges from 0 to 1 and regulates the degree to which the optimal growth rate is reduced.
in the presence of non-functional proteins (Equation (8)):

\[ f(Q) = \frac{\beta}{Q^n + \beta^n} \]  

(8)

In this equation, \( \beta \) is the number of non-functional proteins required for \( f(Q) \) to equal 0.5 and \( n \) is the exponent of the hill function and will determine the how quickly \( f(Q) \) transitions from high to low as the number of non-functional proteins increases. Here, we assume that the length, \( l_C \), initially equals 100 (arbitrary value). Division, modeled as an instantaneous process, is triggered automatically when \( l_C \) reaches the value 200. Note that this reaction is such that, in the absence of non-functional proteins, the cell length grows exponentially in time [49].

We set the division time of the cells (\( t_{div} \)) to match the empirical mean division times of DH5α-PRO cells in LB media (~60 minutes) [29] when \( \beta(Q) \) equals 1 (i.e. no non-functional proteins present). These times are longer than for other strains (e.g. K12). They were chosen since most of the other rate constants used were obtained from measurements in DH5α-PRO. Also, this rate does not qualitatively affect the results. Namely, only the time-scale of events is affected.

Finally, to implement the ‘cluster formation’ partitioning scheme (see ‘Effects of different partitioning schemes’ in the Results section), the model requires one additional reaction. Clusters form during a cell’s lifetime according to the following reaction, where the index \( C \) is the cell ID:

\[ \emptyset + \frac{1}{C} \text{Cluster}_C \rightarrow \text{Cluster}_C \]  

(9)

At division, clusters are partitioned following an unbiased binomial distribution. The non-functional proteins are partitioned by assigning each non-functional protein to one of the partitioned clusters [26]. All cells begin with Cluster set to 1, which then increases depending on the rate of \( k_3 \). Larger values for \( k_3 \) will produce a larger number of clusters in the cell, and thus a weaker ‘degree of clustering’.

The fourth and final component of the model is a retention mechanism of non-functional proteins at the cell poles. Based on previous observations [23, 34], we assume that non-functional proteins inherited by a cell will remain retained at the old pole of that cell during its lifetime, and will be partitioned accordingly in division. Meanwhile, new non-functional proteins produced during a cell’s lifetime will be partitioned between the two daughter cells following the same mechanism used to partition the functional proteins, namely, the unbiased, independent partitioning scheme.

All parameter values, aside from \( \beta \), along with the references from which they were extracted, are shown in Table 1.

**Author contributions**

A.S.R. and A.G designed research; all authors performed research; A.G and J.L.P. implemented the models. A.G. analyzed data; A.S.R. and A.G wrote the paper.

**Acknowledgments**

Work supported by Academy of Finland (grant no. 257603) (ASR), Portuguese Foundation for Science and Technology (grant no. PTDC/BBB-MET/1094/2012) (ASR), Finnish Cultural Foundation (AG) and TUT President’s graduate programme (JLP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**References**

[1] M. Ackermann, A. Schauerste, S.C. Stearns and U. Jenal, Experimental evolution of aging in a bacterium, *EMBO J* 27 (2007), 126.

[2] M. Ackermann, C.S. Stearns and U. Jenal, Senescence in a bacterium with asymmetric division, *Science* 300 (2003), 1920.

[3] M. Ballesteros, A. Fredriksson, J. Henriksson and T. Nyström, Bacterial senescence: Protein oxidation in non-proliferating cells is dictated by the accuracy of the ribosomes, *EMBO J* 20 (2001), 5280–5289.

[4] M.G. Barker and R.M. Walsby, Replicative ageing in the fission yeast *Schizosaccharomyces pombe*, *BMC Biol* 15 (2019), 1511–1518.

[5] N.G. Bodnarska, J. Schymkowitz, F. Roussel and J. Van Eldere, Protein aggregation in bacteria: The thin boundary between soluble and insoluble forms, *Biochim Biophys Acta* 1434 (1999), 170–176.

[6] M.M. Carnis, J.L. Carchero and A. Villota, Protein digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms, *Biochem Biophys Acta* 1434 (1999), 170–176.

[7] M. Coelho, S.J. Lade, S. Alberti, T. Gross and I.M. Tolić, Fusion of protein aggregates facilitates asymmetric damage segregation, *PLoS Biol* 12 (2014), e1001886.
[8] A.-S. Coquelle, J.P. Jacob, M. Prunet, A. Demarez, M. Dimicoli, T. Jolivet, et al. Localization of protein aggregation in Escherichia coli is governed by diffusion and nuclear macromolecular crowding effect. PLoS Comput Biol 9 (2013), e1003038.

[9] X. Dai, S. Healy, O. Yli-Harja and A.S. Ribeiro, Tuning cell differentiation patterns and single cell dynamics by regulating proteins’ functionalities in a toggle switch. J Theor Biol 261 (2009), 441–448.

[10] E. Deuring, A. Schuler-Specking, T. Tomoya, A. Mok and B. Bakus, Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature 400 (1999), 693–696.

[11] C.M. Dobson, The structural basis of protein folding and its links with human disease. Philos Trans R Soc Lond B Biol Sci 356 (2001), 133–145.

[12] W.D. Donachie and K.J. Begg, “Division potential” in Escherichia coli. J Bacteriol 178 (1996), 5971–5976.

[13] W.D. Donachie, K.J. Begg and M. Victore, Cell length, cell growth and cell division, Nature 264 (1976), 328–333.

[14] H. Dong, L. Nilsson, C.G. Kurland, H. Dong, L. Nilsson and C.G. Kurland, Granulosome overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. J Bacteriol 177 (1995), 1497–1504.

[15] N. Enjar, M. Ciprovic, E. Klupp and T. Nyström, Selective benefits of damage partitioning in unicellular systems and its effects on aging. Proc Natl Acad Sci U S A 105 (2008), 18704–18709.

[16] K.L. Ewalt, J.P. Hendrick, W.A. Hoey and F.U. Hartl, In vivo observation of polypeptide flux through the bacterial chaperonin system. Cell 97 (1999), 491–500.

[17] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81 (1977), 2340–2361.

[18] B.R. Glick, Metabolic load and heterologous gene expression, Biophys J 3 (1972), 247–261.

[19] A.L. Goldberg, A role of aminosac-cDNA in the regulation of protein breakdown in Escherichia coli. Proc Natl Acad Sci U S A 71 (1974), 362–366.

[20] A.L. Goldberg, Degradation of Abnormal Proteins in Escherichia coli. Proc Natl Acad Sci U S A 49 (1972), 422–426.

[21] S. Gottsman and M.R. Maurizi, Regulation by proteolysis: Energy-dependent proteases and their targets. Microbiol Rev 56 (1992), 502–521.

[22] A.J. Gray, E.S. Martin, M.A. Krupenko, M.V. Kashlev and E. Maisonneuve, Protein aggregation: An aging factor involved in cell death. J Bacteriol 190 (2008), 3076–3079.

[23] A.L. Hsu, C.T. Murphy and C. Kenyon, Regulation of aging and age-related disease by DAF-16 and heat-shock factor, Science 300 (2003), 1142–1145.

[24] D. Huh and J. Paulsson, Random partitioning of molecules at cell division. Proc Natl Acad Sci U S A 108 (2011), 15004–15009.

[25] D.M. Jarzynski, Growing old: Metabolic control and yeast aging. Annu Rev Microbiol 56 (2002), 769–792.
[47] A. Ribeiro, R.U.I. Zhu, S.A. Kauffman and A. General, Modelling strategy for gene regulatory networks with stochastic dynamics, J Comput Biol 13 (2006), 1630–1639.

[48] M.R. Roussel and R. Zhu, Validation of an algorithm for delay stochastic simulation of transcription and translation in prokaryotic gene expression, Plos Biol 3 (2005), 274–284.

[49] M. Scott, C.W. Gunderson, E.M. Matzker, Z. Zhang and T. Hwa, Interdependence of cell growth and gene expression: Origins and consequences, Science (80) 330 (2010), 1099–1102.

[50] C. Soti and P. Csermely, Aging and molecular chaperones, Exp Gerontol 38 (2003), 1037–1040.

[51] M.A. Speed, D.I. Wang and J. King, Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition, Nat Biotechnol 14 (1996), 1283–1287.

[52] E.J. Stewart, R. Madden, G. Paul and F. Taddei, Aging and death in an organism that reproduces by morphologically symmetric division, Plos Biol 3 (2005), 6295–6300.

[53] S.A. Teter, W.A. Houry, D. Ang, T. Tradler, D. Rockabrand, G. Fischer, et al., Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. Cell 97 (1999), 755–765.

[54] J. Tyedmers, A. Mogk and B. Bukau, Cellular strategies for controlling protein aggregation, Nat Rev Mol Cell Biol 11 (2010), 777–788.

[55] A. Wang, L. Robert, J. Peltier, W.L. Dang, F. Todd, A. Wright, et al., Robust growth of Escherichia coli, Curr Biol 20 (2011), 1099–1103.

[56] S. Ventura and A. Villaverde, Protein quality in bacterial inclusion bodies, Trends Biotechnol 24 (2006), 570–185.

[57] E. Viaplana, J.X. Felix, J.L. Corchero and A. Villaverde, Reversible activation of a cryptic cleavage site within E. coli beta-galactosidase in beta-galactosidase fusion proteins, Biochem Biophs Acta 1343 (1997), 221–226.

[58] S. Wickner, Posttranslational-quality control: Folding, refolding, and degrading proteins, Science (80) 286 (1999), 1888–1893.

[59] S. Willetts, Intracellular protein breakdown in non-growing cells of Escherichia coli, Biochem J 103 (1967), 453–461.

[60] A. Winkler, A. Seybert, L. König, S. Prugnaller, U. Haselmann, V. Sourjik, et al., Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing, EMBO J 29 (2010), 910–923.

[61] J. Yu, J. Xiao, X. Ren, K. Lao and X.S. Xie, Probing gene expression in live cells, one protein molecule at a time, Science (80) 311 (2006), 1060–1063.

[62] P. Zhang, C.M. Khursegur, L.M. Hartnell and S. Subramanian, Direct visualization of Escherichia coli chemotaxis receptor arrays using cryo-electron microscopy, Proc Natl Acad Sci USA 104 (2007), 3777–3781.

[63] R. Zhu, A.S. Ribeiro, D. Salahub and S.A. Kauffman, Studying genetic regulatory networks at the molecular level: Delayed reaction stochastic models, J Theor Biol 246 (2007), 725–745.