The functional half-life of an mRNA depends on the ribosome spacing in an early coding region

Pedersen, Margit; Nissen, Søren; Mitarai, Namiko; Lo Svenningsen, Sine; Sneppen, Kim; Pedersen, Steen

The Functional Half-Life of an mRNA Depends on the Ribosome Spacing in an Early Coding Region

Margit Pedersen[1,2], Søren Nissen[1,2], Namiko Mitarai[2], Sine Lo Svenningsen[1], Kim Sneppen[2] and Steen Pedersen[1]*

[1]Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark
[2]Center for Models of Life, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

Bacterial mRNAs are translated by closely spaced ribosomes and degraded from the 5′-end, with half-lives of around 2 min at 37 °C in most cases. Ribosome-free or "naked" mRNA is known to be readily degraded, but the initial event that inactivates the mRNA functionally has not been fully described. Here, we characterize a determinant of the functional stability of an mRNA, which is located in the early coding region. Using literature values for the mRNA half-lives of variant *lacZ* mRNAs in *Escherichia coli*, we modeled how the ribosome spacing is affected by the translation rate of the individual codons. When comparing the ribosome spacing at various segments of the mRNA to its functional half-life, we found a clear correlation between the functional mRNA half-life and the ribosome spacing in the mRNA region approximately between codon 20 and codon 45. From this finding, we predicted that inserts of slowly translated codons before codon 20 or after codon 45 should shorten or prolong, respectively, the functional mRNA half-life by altering the ribosome density in the important region. These predictions were tested on eight new *lacZ* variants, and their experimentally determined mRNA half-lives all supported the model. We thus suggest that translation-rate-mediated differences in the spacing between ribosomes in this early coding region is a parameter that determines the mRNAs functional half-life. We present a model that is in accordance with many earlier observations and that allows a prediction of the functional half-life of a given mRNA sequence.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

The stability of an mRNA is important for the concentration of the mRNA and thus for the expression of the encoded protein. Early, two aspects of the mRNA instability were defined: (a) its functional half-life, measured by following the expression of the encoded protein after halting DNA transcription and (b) its chemical half-life, measured by following the mRNA decay by, for instance, Northern blot or microarray analysis after halting transcription (e.g., Refs. 1 and 2). Obviously, the mRNA's chemical half-life can be longer than its functional half-life, but not *vice versa*.

Most mRNAs in *Escherichia coli* decay with functional half-lives close to 2 min at 37 °C, but a few mRNA species differ substantially in their half-life resulting in a span of close to 100-fold.[3–5] The stability of the very stable mRNA depends on

sequestering of the mRNAs 5′-end in a secondary structure.[6,7] Characterization of mutants with altered mRNA half-lives and biochemical analyses has led to models for chemical mRNA degradation that involves a multienzyme complex, the degradosome, that has been extensively reviewed.[8–11]

Pinpointing the event that inactivates the mRNA functionally has been elusive,[10] but recently, the hydrolysis of the 5′-triphosphate group to a 5′-monophosphate group at the end of the mRNA, catalyzed by the RppH enzyme, was suggested to be an initial and rate-limiting step in mRNA degradation. Secondary mRNA structures in the 5′-untranslated region (UTR) were shown to protect the 5′-triphosphate group and to stabilize the mRNA, but this study also showed that mRNA decay in the absence of RppH was possible with a reduced rate, indicating alternative pathways.[12,13]

Previously, Petersen constructed eight variants of the *lacZ* mRNA, where short sequences of 4 to 19 codons were inserted after codon 5 in *lacZ*.[14] These mRNAs should have similar structures in the 5′-UTR because the initial 52 nucleotides of the *lacZ* mRNA were identical. Nevertheless, these changes in the coding region decreased the mRNA functional half-life up to fourfold. When the *lacZ* ribosome-binding site was substituted with sites expected to bind ribosomes with higher affinity, the levels of protein expression were increased. Analysis showed only a marginally increased rate of translation initiation but that this increased the mRNA functional half-life twofold to threefold.[15,16] Mathematical modeling, see below, explains this result by the dominance of the occlusion time: if a ribosome masks the ribosome-binding site, this will prevent binding of the following ribosome and prevents a better binding site to be used to its full capacity. These results also indicated that a determinant for the functional mRNA stability was located in the coding region. The aim of our work was to further characterize this determinant.

The rate of translation of individual codons varies at least fivefold.[17] We recently modeled the translation process and focused, in particular, on an experiment that indicated the formation of ribosome queues upstream from inserts of slowly translated codons in *lacZ*.[17,18] This allowed us to estimate the translation initiation frequency for the wild-type (WT) *lacZ* mRNA in living *E. coli* cells rather precisely to 1 initiation per 2.3 s under the conditions used. Our analysis also indicated that stochastic collisions between ribosomes are normal, frequent and probably harmless events.[18] Because it takes approximately 1 s to translate the 11 codons that is covered by a ribosome, the distance between the ribosomes translating the *lacZ* mRNA is on average just above one ribosome diameter, subject to varying local translation rates and to stochastic fluctuations.

The variation in translation rate among individual codons is sufficient to result in large local variations in the spacing of ribosomes even with a constant translation initiation frequency.[17,18] Therefore, we analyzed the ribosome distribution in quantitative terms by mathematical modeling of the above-mentioned *lacZ* variant mRNAs with altered ribosome-binding sites or with short inserts in the early coding region. We found a correlation between the functional half-life of mRNA and the fraction of time an initial region of the mRNA is uncovered by ribosomes. Thus, spacing of ribosomes should be incorporated in models for mRNA breakdown as initially proposed,[19] particularly the spacing between the two most upstream ribosomes on the mRNA.

Results

Variations in the translation rate of individual codons along an mRNA may cause ribosomes to collide, for instance, if slowly translated codons are preceded by rapidly translated codons.[17] The probability of collisions is expected to rise dramatically with the translation initiation frequency. Changes either in the Shine–Dalgarno sequence or in the mRNA coding sequence might therefore affect ribosome spacing quite far from the sequence change itself.

To model the distribution of ribosomes along the mRNA in detail, we applied our previous model of the ribosome traffic[18] to the corresponding sequences and analyzed both the fraction of time a codon is occupied by a ribosome and the fraction of time a specified stretch of mRNA is not masked by ribosomes and accessible for nucleases. The model is described in Materials and Methods†. The codon-specific translation rates used in this modeling were fast (A), middle (B) and slow rate codons (C), translated with a rate of 35; 8; and 4.5 codons per second, respectively.[18] The "A" codons correspond roughly to the codons preferentially used in mRNAs encoding highly expressed proteins and the "C" codons correspond to the codons read by rare tRNA species. These translation rate values are approximations but reproduce closely our previous measurements of the translation time of several individual genes and mRNA inserts in living *E. coli* cells.[18]

We first analyzed how varying local translation rates will affect the ribosome spacing by modeling the artificial situation in three mRNAs with 20 repeats of A, B and C codons in the proportions 5:5:1 that approximates the distribution of these codons in the

---

† The applet can be visited at http://www.cmol.nbi.dk/models/mRNALifeTime/mRNALifeTime.html

lacZ mRNA. The mRNA1 was (AAAAABBBBBC)$_{20}$; mRNA2 was (AAAAA)$_{20}$(BBBBBC)$_{20}$ and mRNA3 was (BBBBBC)$_{20}$(AAAAA)$_{20}$. Figure 1 shows that the relative positions of the slow and fast codons will have large effects on ribosome spacing: mRNA1, (a) has all codons occupied by ribosomes in about 43% of the time whereas for mRNA2 (b) the fast codons are covered approximately 98% of the time because the ribosome queue formed at the slowly translated distal codons extended upstream. For mRNA3 (c), the fast codons are almost completely unoccupied (occupied 8% of the time). The kinetic parameters used to create these readouts were the same. However, for mRNA2, the ribosome queue extended back and masked the ribosome-binding site, thereby lowering the time that this ribosome-binding site was accessible for new initiations. This reduced the actual initiation rate. Varying the on-rate to compensate for this did not change the conclusion that the relative location of slowly and rapidly translated codons will affect the spacing between ribosomes even at a distance.
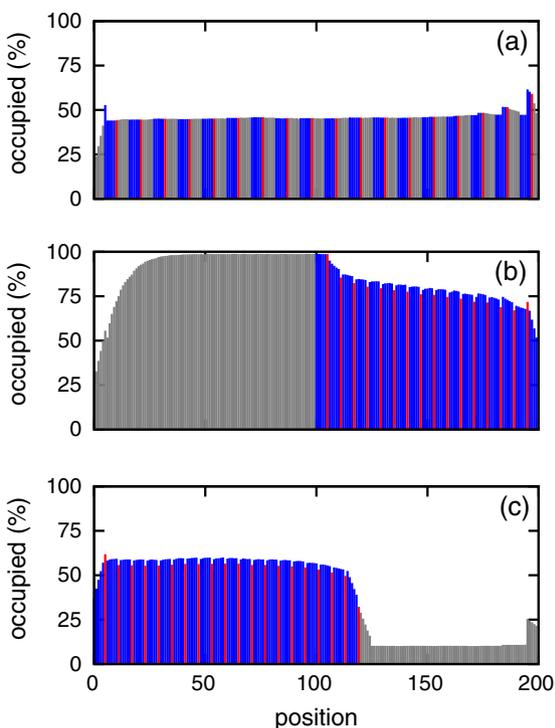

Fig. 1. The spacing between ribosomes varies in response to the distribution of slow and fast translated codons. Readouts from the applet modeling the completely artificial situation, where three mRNAs with the same overall codon composition and the same kinetic parameters were analyzed. These mRNAs were 220 nucleotides long and consist of 11 nucleotide repeats of A, B and C codons, shown in gray, blue and red, respectively. The mRNA1 (a) contained 20 repeats of AAAAABBBBBC rate codons: (AAAAABBBBBC)$_{20}$; mRNA2 (b) was (AAAAA)$_{20}$(BBBBBC)$_{20}$ and mRNA3 (c) was (BBBBBC)$_{20}$(AAAAA)$_{20}$.

To analyze more natural mRNAs, we turned to the eight variants of the *lacZ* mRNA analyzed previously.[14] Here, short sequences inserted between codon 5 and codon 10 in the *lacZ* mRNA were found to decrease the functional mRNA half-life twofold to fourfold. Also, we analyze translation of *lacZ* expressed from the two plasmids (pIV18 and pIV1), where the *lacZ* ribosome-binding site was substituted with sequences from highly expressed genes expected to give a stronger ribosome binding compared to *lacZ*.[15,16] To model the ribosome spacing on these mRNA variants, we estimated the Gibbs free energy, ΔG, for the interaction between their *lacZ* Shine–Dalgarno sequences and the 3'-end of the 16S ribosomal RNA as described previously.[20] This interaction affects the off-rate of the 30S ribosomes and changes the effective on-rate (the rate that ribosomes start elongation, if there is no occlusion by other ribosomes, K_s, as detailed in Materials and Methods) from 1.1/s to 22/s for pIV1 and from 1.1/s to 25/s for pIV18. We assumed that the effective on-rate K_s is proportional to $e^{-\Delta G/RT}$, where R is the gas constant and T is temperature. In addition, we assumed a slower start for initiating on codon GUG in pIV1 compared to AUG (see Materials and Methods). All together, we modeled data from 10 variants of *lacZ* mRNA, which experimentally had been shown to give a 10-fold change in the functional mRNA half-life.[14–16] The inserted sequences are so late in the gene that the translation rate do not affect the initiation rate significantly. All experiments were done using the same background strain, plasmid copy number, temperature and growth medium.

Modeling using these on-rates and the parameters determined earlier[18] show that the resulting initiation rates for *lacZ* mRNA translation in the plasmids pIV18 and pIV1 should be only marginally increased, by 14% relative to the *lacZ* WT initiation rate, in good agreement with the experimental values.[15] This is due to the dominance of the occlusion time. The presence of a ribosome within the first 11 codons prevents binding of the following ribosome and prevents the binding site to be used to its full capacity.

Correlation between the ribosome density in an early mRNA region and the functional half-life of the mRNA

The simulation results for 4 of these 10 *lacZ* variants are shown in Fig. 2. All panels in Fig. 2 show large variations in the ribosome occupancy that result from the distribution of rapidly and slowly translated codons. Examining these readouts show that varying the initiation frequency or having different translation rates of the codons inserted between codon 5 and codon 10 in the WT sequence does indeed affect the ribosome spacing downstream to about 50 codons from the changed
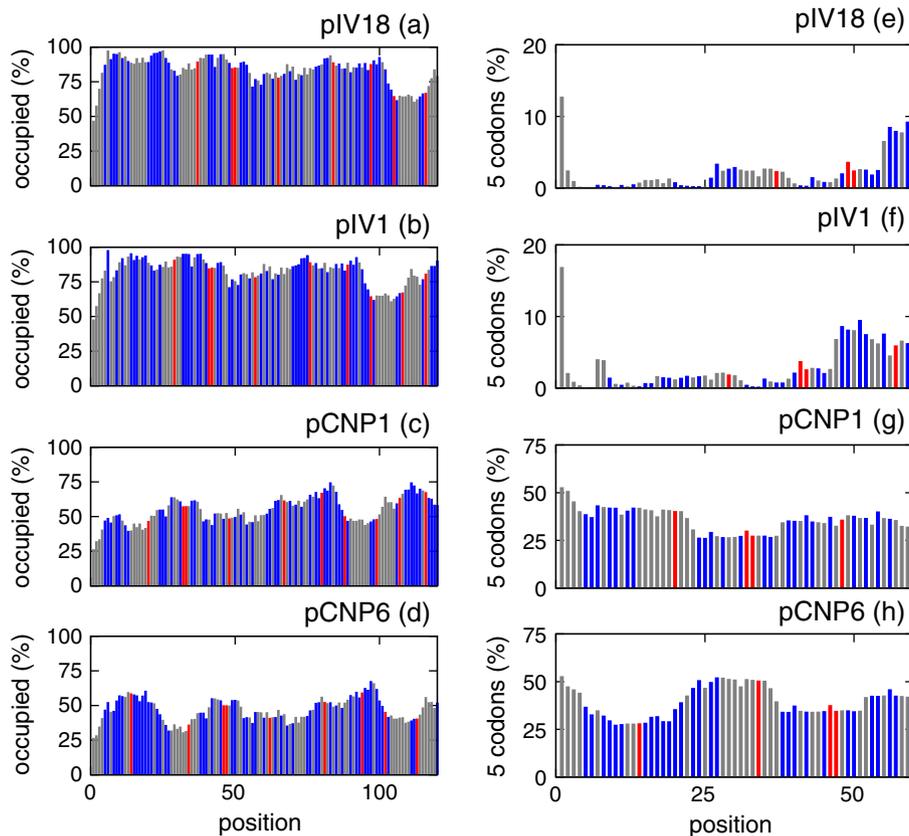
Fig. 2. Modeling ribosome occupancy during translation of the first 120 codons in variants of the *lacZ* mRNA. (a) to (d) show the fraction of time each codon is covered by a ribosome for the *lacZ* variant in (a) pIV18, (b) pIV1, (c) pCNP1 and (d) pCNP6, having functional half-lives of 6 min, 3 min, 117 s and 28 s, respectively. The locations of slow, medium and rapidly translated codons are indicated with red, blue and gray bars, respectively. The corresponding right panels (e–h) give the percent of the time a 5-codon window is free at varying positions in the first 60 codons of the gene.

sequence. For clarity, only the ribosome distribution on the first 120 codons of the mRNAs is shown. Figure 2 also illustrates that the more stable mRNAs have a higher ribosome density on the initial part of the mRNA and that this reduces the activity of the ribosome-binding site. Because of the occurrence of slowly translated codons in *lacZ*, only the first part of the mRNA has a higher ribosome density. On the right panels of Fig. 2, a window of five codons was moved down the mRNA, and the fraction of time where these five codons were uncovered by ribosomes was estimated. The width of this window was chosen to dampen the noise in these plots and to be of a length that allowed binding of a fairly large enzyme complex as the degradosome. The right panels in Fig. 2 would not change significantly if a window of 10 or 20 codons had been used. For all 10 *lacZ* variants, we find a correlation between the fraction of time these five-codon windows are uncovered in the region between codon 20 and codon 45 and the mRNA's functional half-life. For other parts of the mRNA, this correlation is not found (see below). The best correlation we find is for

the mRNA stretch from codon 27 to codon 31. Figure 3 shows the mRNA half-lives (filled circles) for the 10 mRNAs that were used to find the correlation as a function of the probability that the mRNA region from codon 27 to codon 31 is free. Given that our modeling is based on approximate values for the codon-specific translation rates,[18] we expected some noise in this analysis but found that the correlation was worth investigating further.

Predictions from the model and test of these predictions

Our results indicated that the ribosome occupancy of the mRNA from codon 20 to codon 45 is of special importance for the mRNA half-life. From this model, we can predict:

(1) that insertion of slowly translated codons before codon 20 in the WT *lacZ* gene should decrease the functional stability because the region distal to the insert would be less occupied by ribosomes,
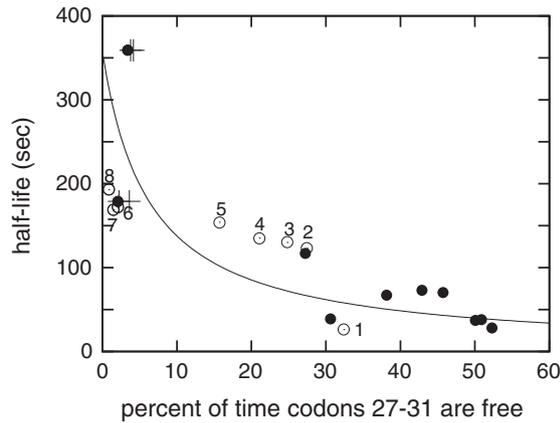
Fig. 3. Correlation between mRNA half-life and the ribosome occupancy in an early part of the mRNA. The mRNA half-life is plotted as a function of the fraction of the time the mRNA from codon 27 to codon 31 is uncovered by a ribosome. Values for the 10 *lacZ* variants used to find the correlation are shown as filled circles. The open circles, numbered from 1 to 8, give the ribosome occupancy and half-life for the new constructs with inserts of slowly translated codons. The points 1–8 show values for *lacZ* encoded by pSN4, pMAP217, pMAP210, pMAP211, pMAP212, pIV1, pMAP206 and pMAP207, respectively (see Table 1 for sequences). The indicated curve gives the best fit to a straight line in the reciprocal plot 1/half-life, R *versus* the percent of time codon 27–31 is free, f. The fitting curve is $R = af + b$ with $a = 0.0005$ and $b = 0.002$. Because b is different from zero, our correlation does not describe all aspects of *lacZ* mRNA degradation. Also, we included modeling data for pIV1 and pIV18 (crosses) with twofold, fourfold or eightfold less effective on-rate compared to the calculated. As long as the effective on-rate is roughly threefold above the WT value, we see that this does not significantly affect the results of the modeling.

(2) that insertion of slowly translated codons after codon 45 should increase the stability because ribosomes would form a queue in front of these codons and protect the region of interest and

(3) that a distal mRNA region, uncovered by ribosomes, should have little effect on the mRNA functional half-life.

These predictions were tested experimentally by constructing the following eight *lacZ* variants: pSN4 has the normal codons at position 16–18 exchanged with slowly translated codons. In the *lacZ* variants in pMAP210, pMAP211 and pMAP212, the normal codons at positions 42, 42–44 or 42–46 were replaced with one, three or five slowly translated codons, respectively. In pMAP205, pMAP206 and pMAP207, the stronger *tufA* Shine–Dalgarno sequence from pIV1 replaced the *lacZ* Shine–Dalgarno

region in pMAP210, pMAP211 and pMAP212. The sequences of the relevant regions in these plasmids are given in Table 1. Finally, we made pMAP217 with an insert of 36 fast translated GAA codons at position 927 in *lacZ*. This insert was modeled to create an unoccupied stretch of mRNA much longer than the region in pSN4.

The WT *lacZ* gene contains two slowly translated codons at positions 31 and 32. In the variants with slowly translated codons inserted downstream of codon 42, the mRNA stability should be affected only slightly according to the model because it is difficult to create a bottleneck after an existing bottleneck. In order to distinguish the expected small changes in the mRNA half-lives, we improved the accuracy in the experiments by performing the half-life determinations on a mixture of two cultures: the *lacZ* variant to be tested and the pMAP217 variant that encodes a β-Gal protein with a higher molecular weight, which allowed separation of the two β-Gal proteins by one-dimensional SDS gel electrophoresis. Thus, the time of sampling, temperature and other experimental conditions were identical for the two plasmids in the mixed culture. The pMAP217 *lacZ* variant was therefore used for normalization between experiments. The internal reference had a half-life of 123 s, not different from the 113 s determined for the WT *lacZ* mRNA.[14] The alteration of the ribosome density after codon 927 in pMAP217 had therefore no influence on the functional half-life. The functional half-lives for these eight new *lacZ* mRNA variants were measured as described previously.[14] Figure 3 shows the normalized values (open circles, numbered from 1 to 8), plotted as function of the modeled ribosome occupancy from codon 27 to codon 31. From Fig. 3, we see that the slowly translated codons 16–18 in pSN4 destabilized the mRNA fourfold, whereas exchange of the natural 1, 3 or 5 codons at codons 42, 42–44 or 42–46, respectively, with slowly translated codons, increased the mRNA functional half-life by approximately 5%, 9% and 23%, both when the mRNA had the *lacZ* ribosome-binding site and the *tufA* ribosome-binding site. We conclude that the measured functional half-lives of these eight new *lacZ* variants correspond well to the predicted values. We modeled the ribosome occupancy along the mRNA for these variants. Figure 4 shows some of these to illustrate that the occupancy in the region from codon 20 to codon 45 closely follows the functional half-life.

### Relevance of the model for other functional mRNA half-lives

Finally, we tested if our model had relevance for other mRNAs for which the functional half-life has been determined, for instance, the *ompA* and *lpp* mRNAs that have an above average stability.

Table 1. Relevant sequences of the constructed *lacZ* variants

Plasmid	Sequence of codon 1–30 in the *lacZ* variants.
pMAS2 (*lacZ* WT)	ATG ACC ATG ATT ACG GAT TCA CTG GCC GTC GTT TTA CAA CGT CGT GAC TGG GAA AAC CCT GGC GTT ACC CAA CTT AAT CGC CTT GCA GCA
pSN4	ATG ACC ATG ATT ACG GAT TCA CTG GCC GTC GTT GGT CAC CGT CGT AGG CGG GGA AAC CCT GGC GTT ACC CAA CTT AAT CGC CTT GCA GCA
pIV1, codon 1–15 *tufA*; from codon 16, *lacZ* WT from its codon 7	GTG TCT AAA GAA AAA TTT GAA CGT ACA AAA CCG CAC GTT GGG AAT TCA CTG GCC GTC GTT TTA CAA CGT CGT GAC TGG GAA AAC CCT GGC
pIV18, codon 1–23 *rpsA*; from codon 24, *lacZ* WT from its codon 7	ATG ACT GAA TCT TTT GCT CAA CTC TTT GAA GAG TCC TTA AAA GAA ATC GAA ACC CGC CCG GCC GGG AAT TCA CTG GCC GTC GTT TTA CAA CGT CGT GAC TGG GAA AAC CCT GGC GTT ACC CAA CTT AAT CGC CTT
	Sequence of codon 33–47 in the *lacZ* variants.
pMAP210	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG GCC CGC ACC GAT CGC
pMAP211	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG CGG GGA ACC GAT CGC
pMAP212	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG CGG GGA AGG CGG CGC
	Sequence of codon 33–47, codon 1–15 from *tufA*, remaining sequences *lacZ*
pMAP205	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG GCC CGC ACC GAT CGC
pMAP206	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG CGG GGA ACC GAT CGC
pMAP207	CCT TTC GCC AGC TGG CGT AAT AGC GAA AGG CGG GGA AGG CGG CGC

Altered nucleotides are underlined.

However, the functional stability of many membrane protein mRNAs is influenced by complex formation to small RNAs.[21,22] In the case of *ompA*, the stability of the mRNA is modulated by binding of small RNA species to the untranslated 5′-end of the mRNA.[23] The availability of proposed binding site of the small RNA is close to the ribosome-binding site and might be influenced by the ribosome occupancy. Our model only describes the occupancy in the
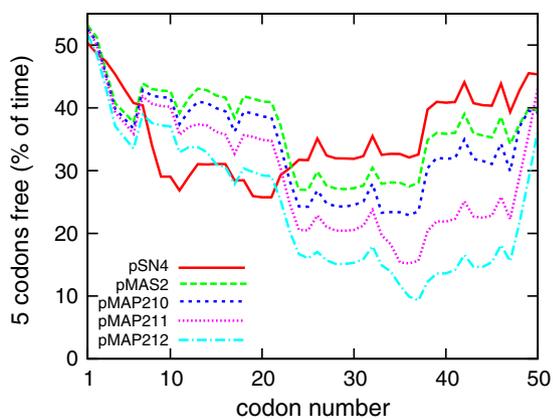
Fig. 4. Effect of inserting slowly translated codons. Modeling the effect of the slowly inserted translated codons on the fraction of time, we unoccupy a moving window of 5 codons along the first 50 codons in the *lacZ* mRNA. *lacZ* WT (green, $t_{0.5} = 113$ s). Three slowly translated codons inserted at codon 16–18 (red, $t_{0.5} = 26$ s); one, at codon 42 (blue, $t_{0.5} = 116$ s); three, at codon 42–44 (violet, $t_{0.5} = 120$ s) or five, at codon 42–46 (turquoise, $t_{0.5} = 136$ s).

translated part of the mRNA, but due to the strong Shine–Dalgarno interaction, the *ompA* mRNA should have a high density of ribosomes in its early coding region. The same holds for the rather stable *bla* mRNA.[24] With these caveats, we have modeled translation of the *ompA* and *bla* mRNAs, and the results are shown in Fig. 5. Modeling the fraction of time codon 27–31 is accessible for the *bla* and *ompA* mRNAs gives values of between 10% and 20%, which according to Fig. 3 would indicate mRNA half-lives well above average for these two mRNAs in agreement with the experimental values.

Discussion

By mathematical modeling, we have analyzed how the translation rate of individual codons influences the spacing of ribosomes on an mRNA. We have focused on modeling ribosome traffic in the early part of the coding region because breakdown of the mRNA takes place from the 5′-end[25,26] and sequence changes here affect the half-life.[1,14–16] We have kept all experimental variables constant except for the small sequence changes to be studied, and this enabled the observation of a correlation between the mRNA's functional half-life and the ribosome occupancy in the coding region of the mRNA approximately from codon 20 to codon 45. To our knowledge, this is the first correlation between an mRNA's functional stability and a recognizable property of the mRNA coding sequence.

The results presented in Fig. 3 were done by analyzing the occupancy of the mRNA from codon
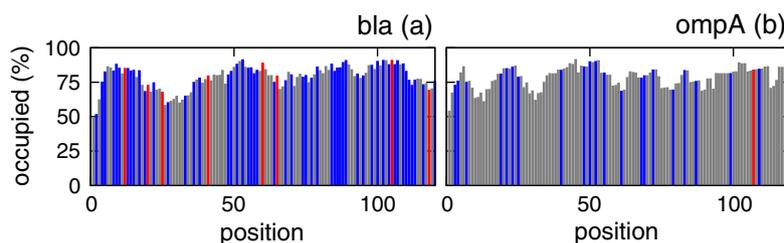
Fig. 5. The modeled ribosome occupancies for the *ompA* mRNA (a) top and the *bla* mRNA (b) bottom plotted as in Fig. 1.

27 to codon 31, which gives the best correlation to the mRNA half-life, but other mRNA stretches, as for example, the stretch from codon 20 to codon 25 or from codon 25 to codon 40, give results that are only slightly different. However, it is only for this initial part of the coding region, codon 20 to codon 45, that such correlation can be observed; ribosome occupancy in the mRNA at codon 1–20 or after codon 50 (Fig. 4) or late in the gene as after codon 927 (Fig. 3) does not correlate with the functional half-life.

What might be special for the codon 20–45 region? As modeled previously,[18] there is a denser packing of the ribosomes early in mRNAs because of the higher density of the rare and presumably slowly translated codons here. Ribosomes initiate once per 2.3 s and physically cover about 11 codons.[18] Therefore, the mRNA segment from codon 20 to codon 45 will often represent the space between the two ribosomes closest to the 5′-end of the mRNA at any time.

The degradosome model for degradation of mRNA[13] has an endonucleolytic cut of the mRNA between two translating ribosomes as one of the options for the initial event that leads to inactivation of the mRNA. We consider this option unlikely. As discussed previously,[18] the estimates for how many times an mRNA is being used in its lifetime are around 30–50. If the degradosome had an obligatory cut between two ribosomes, at least one of these 30–50 ribosomes would need the transfer-messenger RNA (tmRNA) mechanism to become untrapped. The tmRNA mechanism would therefore be involved in at least one of these 30–50 translations plus its known involvement in the special cases where the mRNA is cleaved at, for instance, starved codons.[27] If the degradosome cut between translating ribosomes, the tmRNA should be involved in rescuing 2–3.3% of all ribosomes plus the special cases. In addition to being wasteful, this is an order of magnitude higher than the estimated amount of tagged peptides: 0.4% of the total number of nascent peptides.[28] We conclude that the tmRNA mechanism only acts in the special cases where the mRNA is fragmented and that there, normally, are no degradosome cut between ribosomes.

Our characterization of a half-life determinant in the mRNAs coding region therefore requires an alternative model: The distribution of slow and fast translated codons in the mRNA causes the space between the first two ribosomes to be relatively free (Fig. 6, top). The degradosome (colored blue in Fig. 6) can bind to the free region. The RNase E enzyme has been noted to require such unstructured region for its activity.[29] Due to the bound degradosomes vicinity to the 5′-end, the 5′-triphosphate will be hydrolyzed. This model is compatible with the observation that structures in the UTR might inhibit this interaction and that ribosome occupancy in distal parts of the mRNA will not influence the mRNA stability as in the region at codon 927 in pMAP212. Following its interaction with the 5′-end, the degradosome now leaves the mRNA coding region and continues its activity upstream from ribosome 1 that now finishes its translation without a need for tmRNA rescuing (Fig. 6, bottom). The codon 1–20 region is rarely accessible for degradosome binding because the ribosome spends a relatively long time here due to the initiation process and the many relatively slowly translated codons located here. Recently, small RNAs have been found to cause cleavage of mRNA by binding to a region
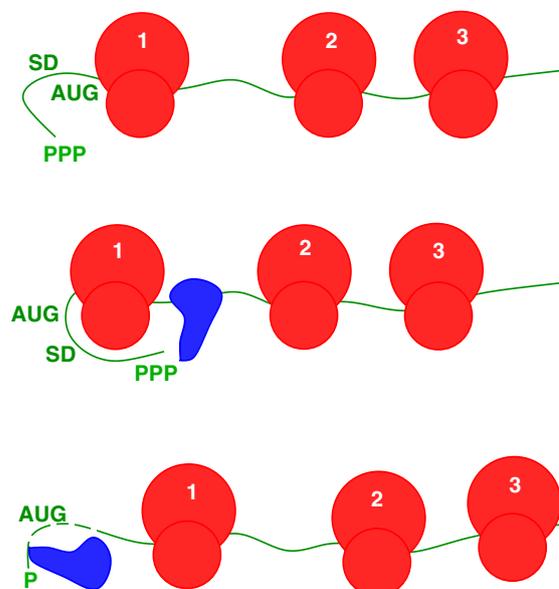

Fig. 6. Model for mRNA degradation.

upstream of the translation start.[21] Such a mechanism might provide an alternative for generating a 5′-monophosphate group on the mRNA. Because we have mainly studied mRNAs that are almost identical, we cannot exclude that additional parameters such as the mRNA length, sequence and structure also influences the stability.

Recently, the impact of the codon usage on the expression of the *gfp* reporter gene was analyzed, and it was concluded that the absence of mRNA structure, and not codon usage in the first 28 codons of an insert, was the major determinant for gene expression.[30] This result is not in contradiction with our model. The *gfp* gene comes from a jelly fish with a codon adaptation index much different from that of *E. coli*. The *gfp* mRNA is therefore expected to be translated slowly,[18] and the ribosomes will most likely back up and cover the inserted 28 codons. In this situation, translation initiation will be determined by the bottleneck of slow codons in *gfp*, and the codons in the inserted region will have little impact on the gene expression level, precisely as was observed.[30]

Our analysis of ribosome spacing depends on a correct estimate of the resulting ribosome on-rate. As mentioned above, the plasmids used[14] all had the same Shine–Dalgarno sequence and the same first five codons in the coding region. For the plasmids pIV1 and pIV18, with a presumed higher affinity for the initiating 30S ribosome, we tested the robustness of our determination of the spacing by using resulting on-rates that were twofold, fourfold or eightfold below the estimated values we estimated. These results are included in Fig. 3. As seen, such changes in the on-rates have only a minor influence on the modeled ribosome spacing in the relevant early region of the mRNA. A general analysis of whether the more stable natural mRNAs are more densely occupied by ribosomes compared to the unstable natural mRNAs is not possible because we lack information about the translation initiation rate or about the functional half-lives of the mRNAs.

The study of Ringquist *et al.* provided a detailed study of how the varying of the Shine–Dalgarno interaction affected *lacZ* expression.[31] However, no functional half-life was measured directly in this study. It is therefore not known if the observed effects on *lacZ* expression were due to an altered on-rate for translation initiation, an altered on-rate that changed the mRNA half-life or an altered transcriptional polarity. These data are not in contradiction with our model because it is very likely that they resulted from an altered on-rate that changed the mRNA half-life via an influence on the ribosome spacing.

There are several known examples where a specific mRNA sequence affects the mRNA half-life. One such example is the finding that a ribosomal protein S1 binding AU-rich mRNA sequence can stabilize an mRNA.[32] According to our modeling, the mechanism behind this mRNA stabilization might well be that avid binding of ribosomal protein S1 to such an mRNA sequence increases the resulting on-rate for 30S ribosome binding, which, in turn, might decrease the ribosome spacing and increase the mRNA half-life.

We found that the preponderance of slowly translated codons in the 5′-end of the mRNA was a highly conserved feature and suggested that this conservation had to do with fine-tuning the translation initiation frequency or had importance for the overall ribosome efficiency by preventing queues later on the mRNA that would trap ribosomes.[18] In addition, the present modeling suggests that the conserved codon usage in the early part of the mRNA also might have evolved to provide the mRNA with a suitable functional half-life.

It is a common observation that the activity of an enzyme often is insensitive to amino acid changes in the N-terminus such as the addition of a tag sequence. The β-Gal protein is a well-known example of this,[33] where a plethora of fusion proteins to its N-terminus retain enzymatic activity. It is therefore conceivable that genes might have few constraints on evolving both the amino acid usage and the codon usage in the initial part that results in a suitable mRNA half-life.

Finally, we note that for at least some eukaryotic mRNAs including plants, the accessibility to the mRNA target for components that affects mRNA degradation [proteins[33] or small RNA (reviewed by Chen *et al.*[34])] might be determined by the codon usage in specific regions of the mRNA.

Materials and Methods

The functional stabilities of the *lacZ* mRNA variants were all carried out in the strain NF1830: *E. coli* MC1000 Δ*ara-leu recA*/F′*lacI*^{q1} *lacZ*: :Tn*5*, essentially as described previously.[14] The cells were grown in a glycerol minimal medium at 37 °C and carried each a *lacZ* gene on a plasmid derived from pBR322. The cells were fully induced by 1 mM IPTG for 15 min, after which the inducer was removed by filtration. The residual rate of synthesis of the β-Gal variant proteins was determined by pulse labeling for 15 s with carrier-free [^{35}S]Met at various time points after removal of the inducer. Samples were chased for 3 min with excess unlabeled Met. After separation of the labeled peptides by one-dimensional SDS gel electrophoresis, the radioactivity in the bands was determined by phosphoimage scanning followed by quantitation by the public domain software ImageJ 1.41o‡. The radioactivity in the β-Gal bands was normalized to that in the RNA polymerase subunits β and β′, assumed to be unaffected by *lac* induction and the functional half-life determined from a semilog plot of these values *versus* the time for the pulse labeling. The standard error of the mean of the mRNA half-life determinations were in the range 5–11%.

Construction of plasmids

All of the new *lacZ* variants were constructed by recombineering using single-stranded DNA primers (DNA Technology A/S, Denmark) with 35 base homologies to both sides of the sequence alteration using the plasmids pMAS2 or pIV1[15,17] as template and in *E. coli* HME70 essentially as described previously.[35–37] First a TAG stop codon was introduced in *lacZ* at position 13 or at position 42. After overnight incubation in rich medium at 30 °C with agitation, the culture was spread on plates containing 100 µg ampicillin and 40 µg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside per milliliter to screen for cells containing a defective *lacZ* gene on either pMAS2 or pIV1. White colonies were cross-streaked with phage Φ80*supF* that restores the activity of *lacZ* amber mutants. The presence of the TAG stop codon at the desired positions was verified by sequencing (Eurofins MWG Operon, Germany). The desired codon changes were again introduced by recombineering, screening for blue colonies on plates containing 100 µg ampicillin and 40 µg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside per milliliter.

The plasmid pMAP217, with an insert of 36 GAA codons at position 927 in *lacZ*, was constructed by first introducing a unique XhoI restriction site at position 927 by recombineering in *lacZ* on pMAS2. A 146-base-long oligo containing 36 GAA codons was used to produce a double-stranded DNA fragment with XhoI restriction site in both ends. The 146-base-pair DNA fragment was cloned using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen). The resulting plasmid was digested with XhoI, and the 123-base-pair XhoI DNA fragment was inserted and cloned in XhoI restricted pMAP201.

Modeling

We performed Monte Carlo simulations of the ribosome traffic model with a constant time step dt. We simulate the traffic of ribosome on one mRNA with infinite half-life and calculated the presented quantities. Ribosomes can bind to the mRNA with the rate K_s, as long as the binding site is accessible. At each time step, all ribosome proceed to the next codon with a probability $R_x dt$ as long as the distance to the preceding ribosome is larger than the occluding distance d. In addition, at the start codon, a ribosome waits the time τ before it starts translation. When a ribosome reaches the 3′-end, it terminates the translation with a rate K_t. More detailed procedures are presented in the previous paper.[18] The parameters used for WT *lacZ* are $K_s = 0.91/\text{s}$, $\tau = 0.2$ s, $d = 11$ codons and $K_t = 2/\text{s}$. The codon dependent rate R_x is given according to the mRNA sequence as denoted in the main text and in Ref. 18. For the mRNAs whose Shine–Dalgarno sequences are different from the WT *lacZ*, we calculated the binding free energy ΔG and changed the binding rate K_s by assuming that K_s is proportional to $e^{-\Delta G/RT}$ (see the next section). If the start codon is not AUG but GUG (pIV1 and its variants), we set $\tau = 0.5$ s (see the next section).

Determination of the resulting on-rates

The interaction between the 16S RNA and the Shine–Dalgarno sequence affects the off-rate of the 30S ribosomes and the resulting ribosome initiation rate. When k_{on} is the diffusion-limited on-rate of ribosomes and ΔG is the binding free energy of a ribosome, the equilibrium off-rate of the ribosome is given by $k_{on}e^{\Delta G/RT}$. Here, R is the gas constant. When k_{start} is the rate at which the ribosome bound to the Shine–Dalgarno site starts elongation, the probability for a ribosome bound to the ribosome-binding site, P_{on}, is given by $dP_{on}/dt = +k_{on}(1 - P_{on}) - (k_{on}e^{\Delta G/RT} + k_{start})P_{on}$. Here, we assume that the occlusion time by elongating ribosome is negligible. In the steady state, this gives the resulting ribosome initiation rate, K_s, to be $K_s = k_{start}P_{on} = (k_{start}k_{on})/(k_{on}e^{\Delta G/RT} + k_{start} + k_{on})$. Thus, the initiation rate K_s will depend on ΔG, especially when ΔG is weak. For strong ΔG ($\Delta G < 0$), the initiation rate might be limited by the on-rate and approach a constant. Our experiments indicate that the initiation rate depends on ΔG; therefore, the off-rate should be rate limiting to some extent. In our analysis, we estimate the initiation rate by a factor $e^{-\Delta G/RT}$ but also tried weaker dependences. The factor $e^{-\Delta G/RT}$ results in the relative resulting on-rates and can be estimated to 1:28:24 for *lacZ* WT, *tufA* (pIV1) and the −9G mutant *rpsA* (pIV18) mRNAs that resulted in a 1.5-fold to 3-fold increase, respectively, in the mRNA half-life. The resulting effective on-rate for *lacZ* WT mRNA translation under these conditions is 0.9/s,[18] and the initiation rates for pIV18 and pIV1 from the above estimate are therefore 22/s and 25/s, respectively. We tested the robustness of our estimate of the resulting on-rates by analyzing rates that were twofold, fourfold or eightfold below these. These results are included in Fig. 3. Note that these resulting effective on-rates do not include the occlusion time: at the high initiation rate, the actual translation frequency is limited by the occlusion time. As seen from Fig. 3, as long as the initiation rate is at least threefold larger than the WT, it does not affect the modeled ribosome occupancies significantly.

References

1. Yarchuk, O., Jacques, N., Guillerez, J. & Dreyfus, M. (1992). Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. *J. Mol. Biol.* **226**, 581–596.
2. Bernstein, J. A., Lin, P. H., Cohen, S. N. & Lin-Chao, S. (2004). Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl Acad. Sci. USA*, **101**, 2758–2763.

3. Blundell, M., Craig, E. & Kennell, D. (1972). Decay rates of different mRNA in *E. coli* and models of decay. *Nat. New Biol.* **238**, 46–49.

4. Pedersen, S. & Reeh, S. (1978). Functional mRNA half lives in *E. coli*. *Mol. Gen. Genet.* **166**, 329–336.

5. Gerdes, K., Thisted, T. & Martinussen, J. (1990). Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Mol. Microbiol.* **4**, 1807–1818.

6. Emory, S. A., Bouvet, P. & Belasco, J. G. (1992). A 5′-terminal stem–loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* **6**, 135–148.

7. Franch, T., Gultyaev, A. P. & Gerdes, K. (1997). Programmed cell death by *hok/sok* of plasmid R1: processing at the *hok* mRNA 3′-end triggers structural rearrangements that allow translation and antisense RNA binding. *J. Mol. Biol.* **273**, 38–51.

8. Dreyfus, M. & Régnier, P. (2002). The poly(A) tail of mRNAs: bodyguard in eukaryotes, scavenger in bacteria. *Cell*, **111**, 611–613.

9. Kushner, S. R. (2002). mRNA decay in *Escherichia coli* comes of age. *J. Bacteriol.* **184**, 4658–4665.

10. Deana, A. & Belasco, J. G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* **19**, 2526–2533.

11. Steege, D. A. (2000). Emerging features of mRNA decay in bacteria. *RNA*, **6**, 1079–1090.

12. Celesnik, H., Deana, A. & Belasco, J. G. (2007). Initiation of RNA decay in *Escherichia coli* by 5′ pyrophosphate removal. *Mol. Cell*, **27**, 79–90.

13. Deana, A., Celesnik, H. & Belasco, J. G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature*, **451**, 355–358.

14. Petersen, C. (1987). The functional stability of the *lacZ* transcript is sensitive towards sequence alterations immediately downstream of the ribosome binding site. *Mol. Gen. Genet.* **209**, 179–187.

15. Vind, J., Sørensen, M. A., Rasmussen, M. D. & Pedersen, S. (1993). Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes. Expression from reporter genes does not always reflect functional mRNA levels. *J. Mol. Biol.* **231**, 678–688.

16. Petersen, C. (1991). Multiple determinants of functional mRNA stability: sequence alterations at either end of the *lacZ* gene affect the rate of mRNA inactivation. *J. Bacteriol.* **173**, 2167–2172.

17. Sørensen, M. A. & Pedersen, S. (1991). Absolute *in vivo* translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J. Mol. Biol.* **222**, 265–280.

18. Mitarai, N., Sneppen, K. & Pedersen, S. (2008). Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J. Mol. Biol.* **382**, 236–245.

19. Apirion, D. (1973). Degradation of RNA in *Escherichia coli*. A hypothesis. *Mol. Gen. Genet.* **122**, 313–322.

20. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986). Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl Acad. Sci. USA*, **83**, 9373–9377.

21. Guillier, M., Gottesman, S. & Storz, G. (2006). Modulating the outer membrane with small RNAs. *Genes Dev.* **20**, 2338–2348.

22. Bossi, L. & Figueroa-Bossi, N. (2007). A small RNA downregulates LamB maltoporin in *Salmonella*. *Mol. Microbiol.* **65**, 799–810.

23. Rasmussen, A. A., Eriksen, M., Gilany, K., Udesen, C., Franch, T., Petersen, C. & Valentin-Hansen, P. (2005). Regulation of *ompA* mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol. Microbiol.* **58**, 1421–1429.

24. Nilsson, G., Belasco, J. G., Cohen, S. N. & von Gabain, A. (1984). Growth-rate dependent regulation of mRNA stability in *Escherichia coli*. *Nature*, **312**, 75–77.

25. Jacquet, M. & Kepes, A. (1971). Initiation, elongation and inactivation of *lac* messenger RNA in *Escherichia coli* studied by measurement of its beta-galactosidase synthesizing capacity *in vivo*. *J. Mol. Biol.* **60**, 453–472.

26. Cannistraro, V. J. & Kennell, D. (1985). Evidence that the 5′ end of *lac* mRNA starts to decay as soon as it is synthesized. *J. Bacteriol.* **161**, 820–822.

27. Keiler, K. C., Waller, P. R. & Sauer, R. T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, **271**, 990–993.

28. Moore, S. D. & Sauer, R. T. (2005). Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*. *Mol. Microbiol.* **58**, 456–466.

29. Mackie, G. A. (1998). Ribonuclease E is a 5′-end-dependent endonuclease. *Nature*, **395**, 720–723.

30. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*, **324**, 255–258.

31. Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G. D. & Gold, L. (1992). Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol. Microbiol.* **9**, 1219–1229.

32. Komarova, A. V., Tchufistova, L. S., Dreyfus, M. & Boni, I. V. (2005). AU-rich sequences within 5′ untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *J. Bacteriol.* **187**, 1344–1349.

33. Lemm, I. & Ross, J. (2002). Regulation of c-*myc* mRNA decay by translational pausing in a coding region instability determinant. *Mol. Cell. Biol.* **22**, 3959–3969.

34. Chen, X. (2009). Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44.

35. Oppenheim, A. B., Rattray, A. J., Bubunenko, M., Thomason, L. C. & Court, D. L. (2004). *In vivo* recombineering of bacteriophage λ by PCR fragments and single-strand oligonucleotides. *Virology*, **319**, 185–189.

36. Thomason, L. C., Costantino, N., Shaw, D. V. & Court, D. L. (2007). Multicopy plasmid modification with phage λ Red recombineering. *Plasmid*, **58**, 148–158.

37. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. (2009). Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223.