8472–8483 *Nucleic Acids Research, 2012, Vol. 40, No. 17* *Published online 22 June 2012*
doi:10.1093/nar/gks583

Dynamics of transcription driven by the *tetA* promoter, one event at a time, in live *Escherichia coli* cells

Anantha-Barathi Muthukrishnan[1], Meenakshisundaram Kandhavelu[1], Jason Lloyd-Price[1], Fedor Kudasov[1], Sharif Chowdhury[1], Olli Yli-Harja[1,2] and Andre S. Ribeiro[1,*]

[1]Laboratory of Biosystem Dynamics, Computational Systems Biology Research Group, Department of Signal Processing, Tampere University of Technology, FI-33101 Tampere, Finland and [2]Institute for Systems Biology, 1441N 34th St, Seattle, WA, 98103-8904, USA

Received February 3, 2012; Revised May 22, 2012; Accepted May 23, 2012

ABSTRACT

In *Escherichia coli*, tetracycline prevents translation. When subject to tetracycline, *E. coli* express TetA to pump it out by a mechanism that is sensitive, while fairly independent of cellular metabolism. We constructed a target gene, P_{tetA}-mRFP1-96BS, with a 96 MS2-GFP binding site array in a single-copy BAC vector, whose expression is controlled by the tetA promoter. We measured the *in vivo* kinetics of production of individual RNA molecules of the target gene as a function of inducer concentration and temperature. From the distributions of intervals between transcription events, we find that RNA production by P_{tetA} is a sub-Poissonian process. Next, we infer the number and duration of the prominent sequential steps in transcription initiation by maximum likelihood estimation. Under full induction and at optimal temperature, we observe three major steps. We find that the kinetics of RNA production under the control of P_{tetA}, including number and duration of the steps, varies with induction strength and temperature. The results are supported by a set of logical pairwise Kolmogorov-Smirnov tests. We conclude that the expression of TetA is controlled by a sequential mechanism that is robust, whereas sensitive to external signals.

INTRODUCTION

Tetracycline is a polycyclic naphthacene carboxamide (1), isolated from Streptomyces genus of Actinobacteria (2), which prevents bacterial cell growth by binding to the 30 S ribosomal subunit. This binding interferes with the attachment of aminoacyl-tRNA to the mRNA-ribosome translation complex, inhibiting protein synthesis (3,4). In *Escherichia coli*, resistance to tetracycline is conferred by the extra-chromosomal Tn*10* transposon encoded class B molecular determinants, forming the *tet* operon (5–8) (also named divergon (9) or regulon (10)). This operon consists of two structural genes, *tetA* and *tetR*. *tetA* is essential for tetracycline resistance (11), as it encodes for a membrane-targeted antiporter protein, TetA, responsible for active efflux of tetracycline, whereas *tetR* codes for TetR that regulates the *tet* operon.

In the absence of tetracycline, TetR binds to the operator sites of *tetA* and *tetR*, preventing their transcription (12). In the presence of tetracycline, TetR binds as a dimer to the biologically active tetracycline–Mg^{2+} complex, causing an allosteric conformational change in the repressor protein (13). This releases the repressor from the DNA, allowing RNA polymerase to bind and initiate transcription of *tetA* and *tetR*. The *tet* operon is thus a self-repressing system (12,14), capable of fast and efficient response to tetracycline. Its expression activity is largely independent of the metabolic state of the host cell, making it a preferential system to control recombinant gene expression (15) and to study mechanisms of gene expression (13). Studies of the β-galactosidase activity of the Tn*10* *tetR-tetA* promoter have showed that P_{tetA} is the strongest of the three promoters of this operon (16). It has thus been isolated, modified, and used in several studies (namely the *tetR/O* region) of gene expression (4), and its derivatives have been used in several synthetic circuits (17–23).

Studies suggest that gene expression kinetics in prokaryotes appears to be mainly controlled at the transcriptional level, particularly in initiation (24). *In vitro* measurements suggest that, in *E. coli,* transcription initiation is a sequential process (25). The first step, named 'closed complex formation', is the binding of the RNA polymerase enzyme to the promoter region and finding of the transcription start site (TSS). This is followed by isomerization, DNA unwinding and loading of the

*To whom correspondence should be addressed. Tel: +358331153928; Fax: +358331154989; Email: andre.ribeiro@tut.fi

nucleotide strand (the 'open complex formation') (25). Afterwards, elongation of the RNA strand takes place (26), and once the termination site is reached, the single-stranded RNA molecule and the RNA polymerase are released (27). This sequential process (28) is subject to tight regulation that takes place at one or more stages.

Evidence suggests that, for most promoters in *E. coli*, the open complex formation is the main rate-determining step in initiation (29,30). However, the durations of other steps, namely the closed complex formation, isomerization, and promoter clearance are also sequence-dependent, differing widely between promoters and in different conditions (31). Environmental factors can also affect this kinetics (32,33). For example, apart from induction (34), factors such as temperature and pH can also affect the ''rate-limiting'' steps in initiation (30,35).

Changes in the kinetics of these steps allows the overall expression rate of a gene to be changed (8,25). For example, *in vivo* measurements of the expression of the synthetic promoter $P_{LtetO-1}$ using the luciferase reporter system have shown that the rate of RNA production can change by 5000-fold with induction by anhydrotetracycline (aTc) (8). However, it is noted that this promoter was engineered with the aim of allowing tight regulation (wide range of induction) (8). Among the several tetracycline analogs, aTc was found to be an effective inducer even at very low concentrations (<50 ng/ml), as it binds to the TetR protein with very high affinity (~35-fold higher than tetracycline) (36–38) and is less toxic to cells with a minimal inhibitory concentration of 4 μg/ml (39).

If transcription initiation is the rate-determining step in RNA production, and if it is composed of several sequential steps (30–32,40,41), provided that these are approximately exponentially distributed in duration, one would expect that transcript production is a sub-Poissonian process (40). However, recent measurements of *in vivo* cell-to-cell diversity in RNA numbers have exhibited higher diversity in RNA numbers than would be expected from a sub-Poissonian process of transcript production (42–44). In (44), it was hypothesized that this was because of either super-Poissonian RNA production or non-Poissonian RNA degradation. In (42,43), it was hypothesized that the cause is the existence of periods of activity and inactivity of the promoter. It is noted that measurements of cell-to-cell diversity in either RNA or protein numbers are affected by several processes other than transcription and translation. Specifically, the kinetics of degradation of RNA and proteins can be complex (45), that is, non-Poissonian, and there may be non-negligible differences in measurements using *in vivo* and *in vitro* techniques (24). Finally, another event that affects cell-to-cell diversity in RNA numbers is cell division, particularly because the intracellular environment of *E. coli* is not well-stirred (46). Even if the RNA molecules are partitioned in an unbiased fashion, the stochastic nature of this process will enhance the cell-to-cell diversity in the numbers of these molecules following division events (47,48). In addition, bacteria are known to partition unwanted protein aggregates in a biased fashion (49,50), including, for example, RNA tagged with MS2 coat protein fused with Green Fluorescent Protein (MS2-GFP) (51,52), and fluorescent proteins such as Tsr-Venus (53,54,55). This bias in partitioning ought to exacerbate cell-to-cell diversity in RNA and protein numbers when assessed by these methods. The degree of stochasticity in transcription thus needs to be assessed without the interference of subsequent events, by measuring transcript production one event at a time (40).

Here, using *in vivo,* single-molecule based techniques, we characterize the kinetics of initiation of P_{tetA} (15,56), including its stochasticity and how it changes with induction strength and temperature. Namely, we assess the transcript production dynamics at the single event level. This is possible by a method recently developed to tag RNA *in vivo* in *E. coli* with MS2-GFP proteins, which allows individual transcription events to be detectable shortly after production (55,57), and the behaviour is similar to that of the unlabelled system (57).

We report measurements of time intervals between consecutive productions of RNA molecules under the control of P_{tetA} when subject to several induction strengths and temperatures. From the distribution of these intervals, we analyse the dynamics of this promoter. We address the following questions. What is the *in vivo* kinetics of RNA production, one event at a time, under the control of P_{tetA}, when fully induced? How does the *in vivo* kinetics of transcription change with induction? How noisy is this process? Finally, from the inference of number and durations of the rate-limiting steps, we address how the kinetics of the rate-limiting steps changes with temperature. In the end, we compare the kinetics of initiation of P_{tetA} with that of the $P_{lac/ara-1}$ promoter, also named P_{lar}, which has been recently characterized (40) using the same methods.

MATERIALS AND METHODS

Chemicals

For routine cultures, the components of Luria-Bertani (LB) broth (Tryptone, Yeast extract and NaCl) were purchased from LabM (UK) and antibiotics from Sigma-Aldrich (USA). Phusion high-fidelity polymerase and other PCR reagents are from Finnzymes (Finland). Fermentas kits (Finland) for Plasmid isolation and PCR product extraction and purification were made as per the instructions provided. To perform qPCR, cells were fixed with RNAprotect bacteria reagent (Qiagen, USA). The Tris and EDTA for lysis buffer were purchased from Sigma-aldrich (USA) and lysozyme from Fermentas (USA). The total RNA extraction was done with RNeasy RNA purification kit (Qiagen, USA). DNase I, RNasefree for RNA purification, was purchased from Promega (USA). iScript Reverse Transcription Supermix for cDNA synthesis and iQ SYBR Green supermix for qPCr were purchased from Biorad (USA). Agarose for microscopic slide gel preparation and electrophoresis and isopropyl b-D-1-thiogalactopyranoside (IPTG) and aTc for induction of cells are from Sigma-Aldrich (USA). For staining DNA and RNA on gels, SYBR-Safe from Invitrogen (USA) was used.

Bacterial strain and growth conditions

The strain *E. coli* DH5α-PRO (identical to DH5α-Z1) (57) was used to clone and express the target and reporter genes. For overnight cultures, the strain from glycerol stock was inoculated in LB broth 10 g of tryptone, 5 g of yeast extract and 5 g of NaCl per litre, pH 7.0) (56) with appropriate antibiotics (100 μg/ml ampicillin and 35 μg/ml chloramphenicol) and incubated at 37°C with shaking (250 rpm).

Genetic constructs

We constructed the target gene P_{tetA}-*mRFP1-96BS* with a 96 MS2-GFP binding site array in a single-copy BAC vector by restricting out the P_{lar} promoter with BamH1 restriction endonuclease from a BAC clone carrying a target gene P_{lar}-*mRFP1-96BS* (42) (a kind gift from Ido Golding, University of Illinois, IL), and replacing it with P_{tetA} amplified from the pTetLux1 plasmid (56). The primers (Forward: 5′GGGATCCCTCACATGACCCGA CAC 3′ and Reverse: 5′GGGATCCACTGCAATCGCG ATAGC 3′) were designed to amplify the P_{tetA} promoter with BamHI restriction site flanking regions. The amplicon and the BAC vector were subjected to BamH1 restriction digestion, followed with ligation of the amplified product into the BAC vector. Thus, we obtained a single copy F-based plasmid carrying the target region P_{tetA}-*mRFP1-96BS*. This product was transferred into the competent *E. coli* strain DH5α–PRO host cells. The recombinants were selected with antibiotic screening and further confirmed with sequence analysis. The reporter molecules to visualize the target RNA were expressed from the pZS12MS2-GFP plasmid (55) (SC101 origin, 6–8 copies per cell, Amp^R, $P_{LlacO-1}$ promoter) cloned into the host strain, a kind gift from Philippe Cluzel, University of Chicago, IL. The *tetR* gene that encodes for a regulatory protein TetR, is integrated into the chromosome of *E. coli* strain DH5α–PRO, under the control of a strong promoter P_{N25}, that ensures appropriate levels of repressor proteins for tight regulation and full induction, in spite of the residual binding affinity of tetR-aTc complex to DNA (8,55).

A detailed map of the *tetA* promoter sequence with the crucial elements, such as the TetR binding site, the −10/–35 regions, the TSS and the ribosome binding site region, as well as the beginning of the MS2-GFP binding region is shown in Figure 1. Note that there is no *tetA* gene in the genome of *E. coli* DH5α–PRO or in the genetic constructs that were transformed into the strain.

Induction of expression of the target gene and of the reporter gene

From the overnight culture, cells were inoculated into a fresh LB medium supplemented with antibiotics, with initial OD of 0.1 at 600 nm and incubated at a specific temperature (24°C or 37°C) to mid-logarithmic phase with 0.5 OD. To induce the production of MS2-GFP proteins, IPTG (1 mM) was added in the medium at 0.35 OD. The target mRNA from P_{tetA}-*mRFP1-96BS* was then induced by adding aTc to the liquid culture. The target mRNA is rapidly tagged by the MS2-GFP proteins in the cytoplasm and can be detected as fluorescent spots soon after transcription occurs (57).

Quantitative PCR for mean mRNA quantification

The quantification of changes in the mean transcripts production rate of the target gene with induction strength and temperature, relative to a reference gene, were validated with qPCR. For the experimental samples, 10 ml of cells with 0.5 OD at 600 nm were induced with aTc (5–25 ng/ml) alone for 1 hour in liquid culture at a specific temperature (24°C or 37°C). Cells were then immediately fixed with RNAprotect bacteria reagent followed by enzymatic lysis with Tris-EDTA lysozyme buffer (pH 8.3). From the lysed cells, total RNA was isolated with RNeasy RNA purification kit. The total RNA was separated by electrophoresis through a 1% agarose gel and stained with *SYBR® Safe* DNA Gel Stain. The RNA was found intact with discreet bands for 16 S and 23 S ribosomal RNAs. To remove DNA contamination, RNA samples were treated with DNase I, RNase-free enzyme as per manufacturer's instructions. The A 260 nm/280 nm ratio for the RNA samples assessed using GeneQuant pro UV/Vis Spectrophotometer (80-2114-98) were 2.0–2.1, indicating highly purified RNA, and the yield was estimated to be 0.4–0.5 μg/μl. cDNA was synthesized from 1 μg of RNA with iScript Reverse Transcription Supermix according to manufacturer's instructions and stored at −20°C. The qPCR master mix contained iQ SYBR Green supermix with primers for the target and reference genes at a final concentration of 200 nM. We used three reference genes (16 S rRNA (42,58), 23 S rRNA (59) and *dxs* (60)) for internal reference, and similar patterns were observed in all cases. In the results section, we show the data relative to the 16 S rRNA reference gene, whereas in the supplement, we show the data relative to *dxs* (Supplementary Figure S2).

The primers for the target mRNA were (Forward: 5′ TA CGAC GCCGAGGTCAAG 3′ and Reverse: 5′ TTGTG GGAGGTGATGTCCA 3′) to the region of *mRFP1* (GenBank Accession Number: AF506027) (61) with amplicon length 90 bp and for the reference gene 16 S rRNA (EcoCyc Accession Number: EG30090) (Forward: 5′ CGTCAGCTCGTGTTGTGAA 3′ and Reverse: 5′ GGACCGCTGGCAACAAAG 3′), with amplicon length 74 bp (40), and primers were obtained from Thermo Scientific. The template for the reaction was 20 ng of cDNA with similar PCR efficiencies for both the target and reference genes, both greater than 95%. The thermal cycling protocol used was: 94°C for 15 s, 54°C for 30 s, and 72°C for 30 s up to 36 cycles, and in the end, one cycle of 94°C for 15 s. The fluorescence was read at the end of each cycle. These reactions were performed in three experiments, each with three replicates per condition with a final reaction volume of 50 μl. No-RT controls and no-template controls were used to crosscheck non-specific signals and contamination. The reaction was carried out in low-profile tube strips in a MiniOpticon Real time PCR system (Biorad). The Cq values generated

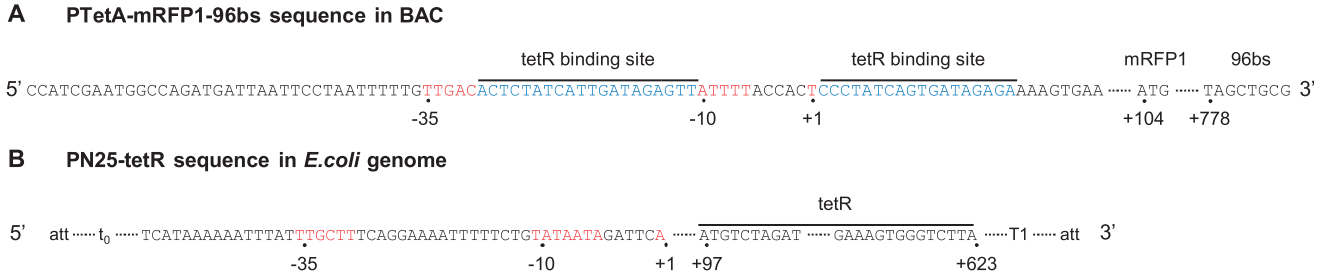

Figure 1. (**A**) P_{tetA}-*mRFP1-96bs*-BAC plasmid map: nucleotide sequence of the *tet* regulatory region on the BAC plasmid, beginning with P_{tetA} (-35 and −10 consensus sequences), TSS (+1), palindromic patterns corresponding to two TetR binding sites (blue), mRFP1 start site at +104, and 96bs start site at +778. (**B**) P_{N25}-*tetR* sequence integrated in the *E. coli* genome: P_{N25} controlled TetR protein coding gene and the lambda *attP* site. Terminators t_0 and T1 prevent transcription from the integrated promoters into the neighbouring regions of the *E. coli* genome.

by CFX Manager™ Software were imported into Microsoft Excel, and the data were analysed following the Livak method (62) to obtain the fold changes in the target gene, normalized to the reference gene, and to calculate the standard error between experiments.

Fluorescent microplate reader measurement for mean protein levels

The mean fluorescence of the mRFP1 protein under the control of P_{tetA} was measured with a Thermo Scientific* Fluoroskan Ascent Microplate Fluorometer. Cells at 0.5 OD_{600nm} were induced with 15 ng/ml of aTc and incubated at a specific temperature (24°C or 37°C) for 1 hour with shaking. The optical density of induced and non-induced cells was measured after 1 hour. From that, 0.5 OD_{600nm} of cells were taken, centrifuged and then re-suspended with fresh medium of 1:200 folds dilution. From this, 150 µl of cells were taken and placed on 96 well microplate and measured for relative fluorescence levels of mRFP1 protein with excitation (584 nm) and emission (607 nm) wavelengths (61). The cell density was kept identical in all wells of the plate for all measured conditions. We performed three independent experiments with three replicates for each condition.

Time-lapse microscopy

Cells were induced with IPTG and aTc as described earlier in the text. Five minutes after induction in liquid culture by aTc, cells were placed on a microscope slide between a coverslip and 1% LB-agarose gel with IPTG (1 mM) and aTc (15 ng/ml), to maintain full induction under the microscope. Cells were visualized in a Nikon Eclipse (TE2000-U, Nikon, Japan) inverted C1 confocal laser-scanning system with a 100× Apo TIRF (1.49 NA, oil) objective. The slide was kept in a temperature-controlled chamber, and the cells were focused within a few seconds under light microscope. Images were collected once per minute up to 1 hour under the fluorescence confocal microscope. Image acquisition began approximately 20 minutes after induction (including the 5 minutes induction in liquid culture). This interval is sufficient to reach a steady state level of induction (55,63). For image acquisition, we used Nikon software EZ-C1, under dark condition to minimize photolysis of aTc (photobleaching and nutrient depletion only become significant after 2 or more hours). GFP fluorescence was measured using a 488 nm argon ion laser (Melles-Griot) and a 515/30 nm detection filter. Images were acquired using medium pinhole, gain 130 and 1.68 µs pixel dwell. On the slide, the division time of the cells was approximately 40 minutes, likely because of the imaging.

We used a recent interacting multiple model filter based autofocus strategy (64). The method relies on the nature of the focal drift and exploits the interacting multiple model filter algorithm to predict the focal drift at time *t* based on the measurement at time *t-1*. It allows a drastic reduction of the number of required z-slices for focal drift correction, thus minimizing photo bleaching.

Cells and spots segmentation and the intensity jump-detection method

Detection of cells from the images is performed by a semi-automatic method (40). It consists of manually masking the regions that cells occupy during the imaging. For each image, the locations, dimensions and orientations of the cells within their masks were estimated by principal component analysis, assuming that the fluorescence inside the cell is uniformly distributed. This assumption is supported by measurements of the pixels intensities inside each cell, which were found to be fairly uniform (40). The segmentation of fluorescent spots (tagged RNAs) was performed by a kernel density estimation method (65) with a Gaussian kernel. An example of the results of the segmentation is shown in Supplementary Figure S1.

From these data, we compute the cell-background-subtracted total spot intensity time traces for each cell. Because the tagged RNAs do not degrade during the experiment (shown in the 'Results' section), this intensity should follow a monotonically increasing piecewise-constant function, where the jumps correspond to the appearance of novel mRNAs. This was verified by inspection (40). We fit such a function to the time trace by least squares, where the number of pieces in the function is determined by an *F*-test with a *P* value of 0.01, thus requiring higher order curves to fit significantly better to justify their use. Some intermediate results of this procedure, along with raw data, are shown in Figure 2. In Supplementary Data, we provide movies of two cells showing the time (in seconds) when each frame was obtained following induction of the target gene.

Inference of the number and duration of the sequential steps in transcription

From the distribution of intervals between productions of consecutive RNA molecules, we infer by maximum-likelihood the number and duration of the sequential steps in transcription initiation (40). We assume that the duration of each of the sequential steps follows an exponential distribution (40). Although the steps in initiation, such as the open complex formation, are likely not elementary (25), it was possible by this method to fit well the measured distributions in the case of P_{lar} using a small number of steps that is consistent with the number of steps believed to be rate-limiting from *in vitro* studies (31,32). To support the results of the inference, we perform Kolmogorov-Smirnov tests to compare the measured distribution with the inferred one that best fits the data. This test is used to determine whether the inferred distribution does not fit the measured distribution.

The inference procedure assumes that the measured intervals between the productions of consecutive RNA molecules are not significantly affected by elongation. This relies on the fact that the mean duration of the intervals between consecutive transcription events is of the order of 600 s or higher, depending on induction strength and temperature (see Results section and *in vitro* studies (8)). On the other hand, elongation of the target gene was measured to take only tens of seconds (57). Possible sequence-dependent events, such as long transcriptional pauses, can also be ruled out as affecting significantly the measured distributions because, if existing, they would last only 10–100 s (e.g. 32 s half-life for the *ops* pause and 52 s for the *his* pause) (66). In addition, target RNA molecules become visible even while elongating (57), further diminishing possible effects of events in elongation in the measured distributions. In any case, although the elongation process may increase the variance of the distributions, it would not affect their means. Finally, the eventuality of possible premature terminations of transcription events can be ruled out in this case because they would generate distributions of intervals between transcription events with multiple peaks, centered on multiples of the mean interval between productions, which were not observed (see 'Results' section).

Here, we assume a d-steps model such that the durations of the d steps are exponentially distributed and are independent, with possibly different rates. The fit for each d-steps model is obtained by maximum likelihood estimation. The likelihoods are compared using likelihood ratio test, and the model with smallest d is selected that cannot be rejected at the significance level 0.01 in favour of a higher order model. This method was found to reliably distinguish the number of steps and the duration between any two steps when they differ by ~25% or more in duration, from 200 intervals sampled from a stochastic model of gene expression with d exponentially distributed steps (40). Note that this method does not allow us to determine the temporal order of the sequential steps inferred. Only their number and durations can be assessed.

RESULTS

We study the *in vivo* kinetics of production of individual RNA molecules, as a function of inducer concentration and temperature, under the control of P_{tetA}. First, we measured the relative mean RNA levels by qPCR as a function of induction and temperature. In Figure 3, we show the results using the 16 S rRNA gene as reference, whereas in supplement, we show the results using *dxs* as reference (Supplementary Figure S2), which are similar to 16 S rRNA. Previous studies suggest that the maximum induction is achieved with a concentration of aTc of 20 ng/ml or higher (8,56). The results (Figure 3 and Supplementary Figure S2) are in agreement, indicating that there is no significant increase in the rate of RNA production beyond 15 ng/ml of aTc. From these figures, we find that both the temperature and the inducer concentration

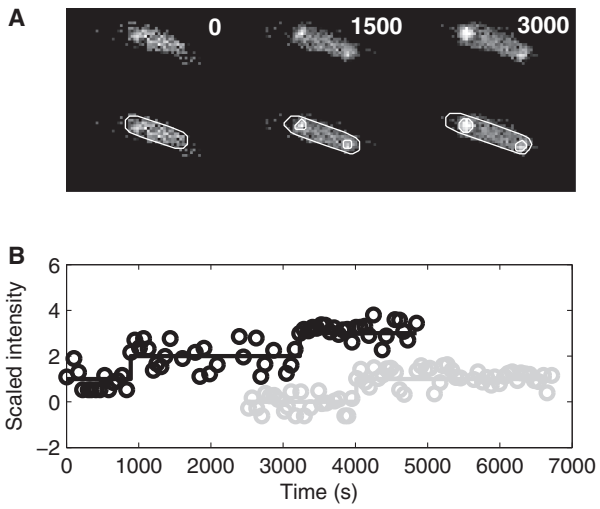

Figure 2. Tagged RNAs in *E. coli* cells. (**A**) Unprocessed frames and segmented cells and RNA spots. The moments when images were taken are shown for each frame. (**B**) Examples of time series of scaled spot intensity levels from individual cells (circles) and the corresponding estimated RNA numbers (solid lines). The cell shown in (A) does not correspond to any of the traces in (B).

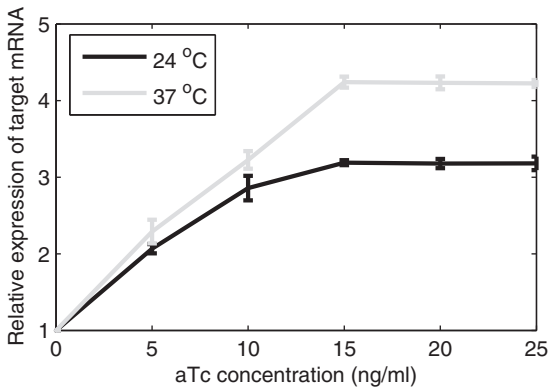

Figure 3. Relative expression level of target mRNA induced with different concentrations of aTc (ng/ml) at 24°C and 37°C, quantified by qPCR using the 16 S rRNA gene as reference. The standard deviation bars are from three independent experiments. In some cases, these bars' lengths are too small to be visible.

significantly affect the rate of transcription under the control of P_{tetA}.

From here onwards, we focus on three conditions. Specifically, we measure gene expression in the absence of aTc at 37°C, and with 15 ng/ml aTc at both 37°C and 24°C so as to study how the kinetics changes with induction and temperature. First, we verify whether, for these conditions, the relative protein expression levels follow those of the RNA. Results of the measurements of relative fluorescence levels by microplate fluorometer are shown in Figure 4, and confirm that the protein levels follow the RNA levels. They also show similarity (same order of magnitude) to the measurements reported in (56) for this promoter using a luminescent reporter system, even though the cells in (56) were in late log-phase and had a 90 min induction period.

We next study the kinetics of RNA production in live, individual cells using the MS2-GFP tagging method (40,42). As described in the Methods section, the expression of the target gene is controlled by P_{tetA} and is induced by aTc. The sequence of the target gene contains 96 binding sites for the MS2 coat protein. Because of these, the reporter proteins (MS2-GFP) can bind to the target RNA, producing a fluorescent spot that is detectable from fluorescence microscopy images.

Using this system, we first measured the cell-to-cell diversity in the number of tagged RNA molecules produced by individual cells over a certain period of time in all three conditions. In both the induced and non-induced cases, cells are placed under the confocal microscope 2 h following induction by IPTG. In the induced cases, the induction by aTc is done 1 h following the induction by IPTG.

From the images, we extracted the number of target RNA molecules in each cell, and calculated the mean and standard deviation of the number of RNA molecules in individual cells (Figure 5). We observe that the measured mean is in agreement with the measurements by qPCR (Figure 3), though the *in vivo* measurements have a slightly smaller relative increase with induction.

We also extracted the fraction of cells with a given number of RNA molecules (Figure 6). In the non-induced case (0 ng/ml aTc, 37°C), the variance of the distribution is 0.62, and the mean is 1.0. In the induced case, the variance is 2.5, and the mean is 3.6. Finally, for cells induced with 15 ng/ml aTc and incubated at 24°C, the variance of the distribution is 1.2, and the mean is 2.2. These values are of significance in that, in all cases, the variance is smaller than the mean (i.e. the Fano factor of the distribution is smaller than 1). If the process of RNA production was Poissonian, that is, if the intervals between consecutive productions were independent and followed an exponential distribution, the variance ought to be equal to the mean (i.e. Fano factor equal to 1). Because the tagged RNA molecules do not degrade, this can only be explained either by the dynamics of RNA production or of RNA partitioning in division (or the combination of these processes).

To investigate this, we obtained distributions of intervals between productions of consecutive RNA molecules in individual cells in the three conditions (Figure 7). These distributions are not affected by the dynamics of RNA partitioning in division because we only count intervals between consecutive RNAs in single cells. In each case, several independent measurements were made, and the results were combined.

The number of cells observed in each condition is shown in Table 1 along with the number of intervals detected, as well as the square of the coefficient of variation (CV^2, defined as the variance over the mean squared) of the intervals. It is noted that the mean duration of the intervals for the aTc 15 ng/ml (37°C) condition is in close agreement with *in vitro* measurements of the time it takes for a transcription initiation event to be completed once initiated (35).

From Figure 7, in all cases, the shapes of the distributions of intervals are not exponential-like. This implies that the process of RNA production under the control of P_{tetA} is not Poissonian. Instead, because the standard

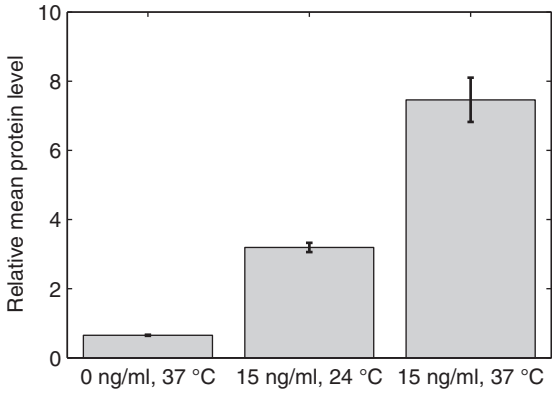

Figure 4. Relative mean expression level of target proteins (mRFP1) estimated by microplate fluorometer in three conditions. The error bars are the standard deviation from the measurements in the different wells. For precise quantities, see Supplementary Table S1.

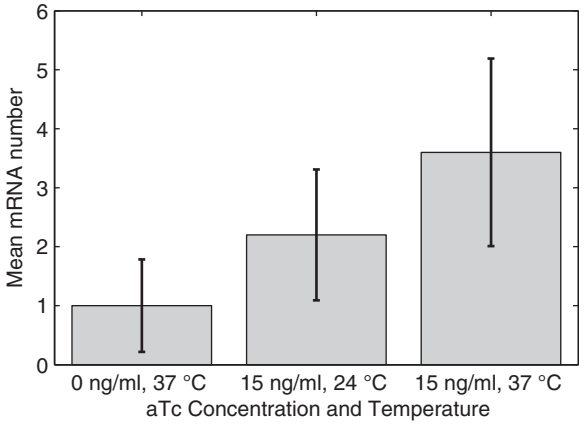

Figure 5. Mean levels of the target mRNA relative to the non-induced condition at 37°C obtained from live cell imaging of the RNA tagged with MS2-GFP under the confocal microscope. Images taken 1 h after induction by aTc in three conditions. The error bars are the standard deviation of the number of RNA molecules in each cell. For precise quantities, see Supplementary Table S2.

deviations of the distributions are smaller than the means, resulting in a CV^2 below 1 (Table 1), we can conclude that this process is sub-Poissonian. This explains the low values of cell-to-cell diversity in RNA numbers observed in cell populations (Figure 6).

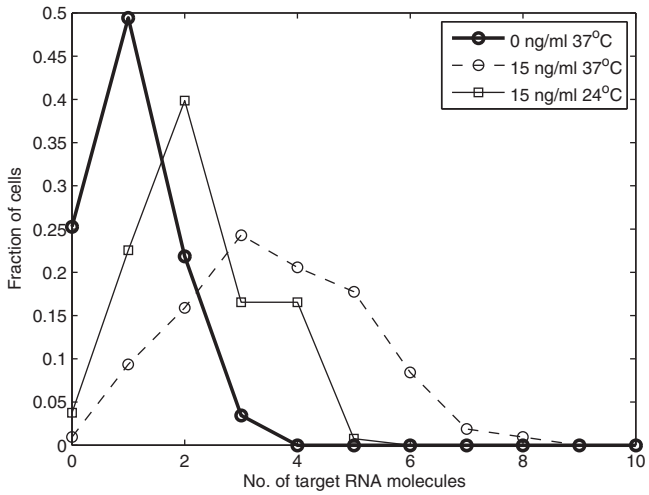

Figure 6. Distribution of the fraction of cells with a given number of mRNA molecules, 1 h following induction by aTc (when induced) obtained from live cell imaging in three conditions.

From the comparison of the distributions A and C in Figure 7, we can assess the effects of induction in the dynamics of transcription under the control of P_{tetA}. Note how distribution C, aside from having a smaller mean, is also slightly more exponential-like, explaining why the CV^2 in RNA numbers is closer to 1 in case C than in case A (Table 1). From the comparison of the distributions B and C, we can assess the effects of lowering the temperature in the dynamics of transcription under the control of P_{tetA}. When the temperature is reduced from 37°C (distribution C) to 24°C (distribution B), the mean interval between transcription events increases, and the shape of the distribution changes significantly. Namely, distribution B, corresponding to full induction at 24°C, is more exponential-like than distribution C.

These results rely on the intensity jump-detection method. This method assumes that the target RNA molecules are quickly bound by the MS2-GFP tagging proteins once transcribed (which was verified in (55)) and that, once that occurs, they do not degrade during the measurement. To validate this assumption of ''immortality'', we studied the kinetics of degradation of these complexes. The MS2-GFP proteins used here are an assembly-defective mutant with the FG loop deleted (55,67). A recent study suggested that when using a target RNA with 48 binding sites for these MS2-GFP

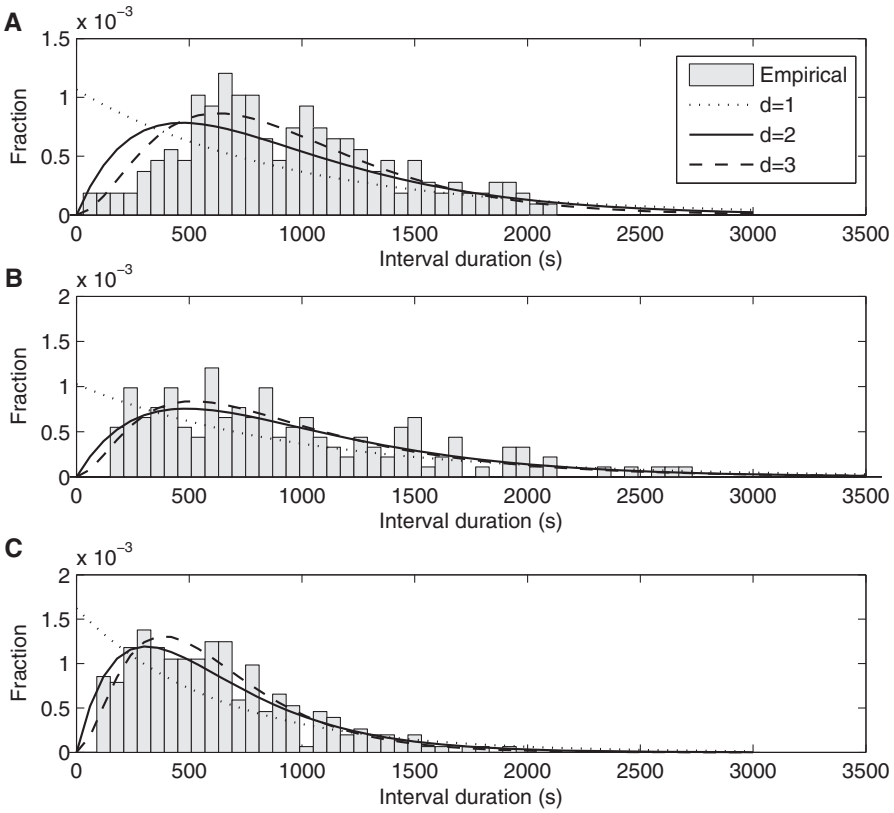

Figure 7. Distributions of time intervals between productions of consecutive mRNA molecules in individual cells under the control of P_{tetA} in conditions: (**A**) 0 ng/ml aTc at 37°C, obtained from 157 cells and 43 intervals (**B**) 15 ng/ml aTc at 24°C, obtained from 119 cells and 100 intervals, and (**C**) 15 ng/ml aTc at 37°C, obtained from 113 cells and 254 intervals. The probability density functions of inferred models of transcription initiation with differing number of rate-limiting steps are also shown in each plot. Note the different scales in the *y*-axis.

proteins, the resulting RNA-MS2-GFP complex may degrade during measurement sessions of a couple of hours (46). This would interfere, to some extent, with the intensity jump-detection method (see 'Materials and Methods' section). However, our target mRNA has 96 binding sites, and therefore the kinetics of degradation likely differs. We studied the degradation kinetics of the target RNA when bound the MS2-GFP and observed ~120 RNA-MS2-GFP complexes produced in 50 cells for up to 2 h. We did not observe any degradation event of these complexes nor did we observe any significant loss of brightness in individual spots. We conclude that the additional MS2-GFP molecules provide sufficient stability to the complex, for us to consider them to be immortal for all practical purposes, implying that the total brightness of the spots within a single cell monotonically increases with time.

We next study the kinetics of the underlying processes responsible for shaping the interval distributions (Figure 7) and thus the observed cell-to-cell diversity in RNA numbers (Figure 6). For this, certain assumptions are necessary. First, we assume that the distributions are mainly shaped by the kinetics of transcription initiation. This relies on the following: in all distributions in Figure 7, the mean interval duration is higher than ~500 s. Therefore, elongation is not expected to play a significant role in shaping the distributions because it lasts only tens of seconds (57). Also, as noted, the target RNA, when bound by MS2-GFP proteins, does not degrade during the measurements. Because of this, we assume that the distributions are shaped by transcription initiation, which includes steps such as the closed complex formation, isomerization and open complex formation (8,25–32,35,40,41,68).

We estimate by maximum likelihood the number and duration of the most prominent "rate-limiting" steps, that is, the ones that shape the distributions of intervals between transcription events (40). The results are shown in Figure 7 for each condition when assuming one, two and three steps. In Table 2, we show the log-likelihood values and the durations of the inferred steps for 1-step, 2-step and 3-step models, for each condition. The number of steps can be determined by using a likelihood-ratio test between pairs of models to reject a lower-degree model in favour of a higher-degree one (55). In Table 3, we show the results of the likelihood-ratio tests. For all distributions, the test rejects the 1-step (exponential) model in favour of the 2-step model (P values $< 8.32 \times 10^{-7}$). For distributions A and C, the 2-step model is also rejected in favour of the 3-step model (P values < 0.0019).

The time scales of the steps (for d = 2) are identical for all cases. As discussed in a previous work (40) this may be because of some unknown artefact of the inference method or be representative of the real kinetics of transcription initiation of this promoter. The method of inference was found to reliably distinguish the duration of each step when they differ by approximately 25% or more in duration, from 200 intervals sampled from a model of gene expression (40). For smaller differences, the solution can be biased toward inferring steps with identical durations, for unknown reasons. Nevertheless, given the number of intervals measured, it is possible to conclude that the steps do not differ by more than approximately 25%.

From Tables 2 and 3, provided that the assumed sequential model of transcription initiation (8,25–32,35,40,41,68) is correct, we conclude that, when not induced with aTc, transcription initiation controlled by P_{tetA} has 3 rate-limiting steps, which are similar in duration (differing by less than 100–150 s between them). When fully induced, at 24°C, there are two dominant rate-limiting steps, similar in duration. Finally, at 37°C under full induction, there are three rate-limiting steps, two longer and similar in duration, and a (clearly) shorter third step. No significant improvement was obtained in the fit with more steps in any of the conditions.

We conclude that there are three rate-limiting steps in transcription initiation of P_{tetA}. By lowering the temperature, two of the steps become longer in duration, whereas the third step remains unaltered and, because of its now relatively much shorter duration, it becomes barely detectable (P value of 0.0268, Table 3). Interestingly, the other two steps are not significantly affected by temperature (compare cases B and C in Table 2). Induction on the

Table 1. Number of cells analysed, number of intervals between productions of consecutive RNA molecules detected in individual cells, mean interval duration, standard deviation of interval durations and CV^2 of the interval durations for each condition

Condition	aTc 0 ng/ml (37°C)	aTc 15 ng/ml (24°C)	aTc 15 ng/ml (37°C)
No. cells	504	178	113
No. intervals	180	152	254
Interval mean (s)	939	974	617
Interval std (s)	459	676	367
Interval CV^2	0.21	0.48	0.35

Table 2. Log-likelihood and durations of the steps of the inferred models with d steps, for each condition

	aTc 0 ng/ml (37°C)		aTc 15 ng/ml (24°C)		aTc 15 ng/ml (37°C)	
d	Log-likelihood	Durations (s)	Log-likelihood	Durations (s)	Log-likelihood	Durations (s)
1	–1412	939	–1198	975	–1886	617
2	–1369	(470, 470)	–1174	(487, 487)	–1836	(309, 309)
3	–1356	(313, 313, 313)	–1171	(620, 240, 115)	–1828	(254, 254, 109)

There is no implied temporal order for the steps.

Table 3. Likelihood-ratio test *P* values between pairs of models for each condition

(d_0, d_1)	aTc 0 ng/ml (37°C)	aTc 15 ng/ml (24°C)	aTc 15 ng/ml (37°C)
(1, 2)	0	3.02×10^{-12}	0
(2, 3)	7.27×10^{-7}	0.027	1.20×10^{-4}
(3, 4)	0.342	0.268	0.370

The null model is the d_0 step model (where d_0 is 1, 2, or 3) while the alternative model is the d_1 step model (where $d_1 = d_0 + 1$)

Table 4. *P* values from the Kolmogorov-Smirnov test between the empirical distribution and the various inferred models with *d* steps

d	aTc 0 ng/ml (37°C)	aTc 15 ng/ml (24°C)	aTc 15 ng/ml (37°C)
1	4.84×10^{-13}	4.81×10^{-5}	1.27×10^{-11}
2	1.10×10^{-4}	0.725	0.0485
3	0.0782	0.844	0.378
4	0.662	0.842	0.272

other hand affects the duration of the three steps (compare cases A and C in Table 2).

It is possible to provide additional support to these results, as well as to the assumption that each step follows an exponential distribution in duration, using a set of logical pairwise Kolmogorov-Smirnov (K–S) tests. One can compare the distributions of intervals between transcription events, for each condition, between the empirical cumulative distribution function of each case and the corresponding cumulative distribution function of the inferred models with d-steps, for all values of *d*. If the model accurately describes the measurements, the empirical and the inferred distributions should be indistinguishable by the K–S test. The comparisons (i.e. the *P* values) are shown in Table 4. Usually, for *P* values smaller than 0.01, it is concluded that the two distributions differ significantly.

From Table 4, in case A (aTc 0 ng/ml, 37°C), the models with less than 3 steps do not accurately match the measured data. In case B (aTc 15 ng/ml, 24°C), only the 1-step model does not accurately match the data. This, as noted, is because of the increase in duration in two of the steps because of the lower temperature, rendering the effects of the third step much less significant in the overall distribution of intervals. Finally, in case C, we have the same result as in case A, that is, models with less than 3 steps do not accurately match the measured data. These results support the previous conclusions, using the likelihood ratio test, regarding the number of sequential steps that determine the shape of the distribution of intervals in each condition.

Finally, we tested whether during our *in vivo* measurements, the kinetics of production of RNA changed over time because of possible changes in the intracellular concentration of aTc. This could occur if degradation because of light sensitivity of intracellular aTc or its slow diffusion across the membrane were significant. That is, if either effect is significant, it should be possible to distinguish between the distributions of intervals obtained in the first 30 min and those obtained in the last 30 min of the hour-long measurements. We extracted these two sub-distributions from each induction condition and compared them with the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test was unable to differentiate the two distributions in any condition (all *P* values > 0.1), demonstrating that the measurements were done at an approximate steady state. In particular, for case A (aTc 0 ng/ml, 37°C), the *P* value was 0.32; for case B (aTc 15 ng/ml, 24°C); it was 0.16 and for case C (aTc 15 ng/ml, 37°C), it was 0.78.

DISCUSSION

Previous studies have shown that the kinetics of transcription initiation of P_{tetA} is heavily dependent on induction and on environmental factors such as temperature (8,16,35,56). Other studies have shown that RNA production in bacteria is a stochastic (24,69) and a multi-step process (30), generally subject to complex regulatory mechanisms (8). Finally, recent studies have shown that the cell-to-cell diversity in RNA numbers, and likely protein numbers, can be significantly affected by processes other than gene expression (47,48,51). Therefore, the assessment of the kinetics of gene expression and regulation requires *in vivo* measurements of RNA production dynamics in individual cells, one event at a time, under various induction and environmental conditions (40).

The measured *in vivo* distributions of intervals between consecutive productions of RNA molecules in single cells are found to be sub-Poissonian, for the induction levels and temperatures tested. This was also observed in the case of P_{lar} (40). The sub-Poissonian nature of the kinetics explains the low cell-to-cell diversity in RNA numbers observed at the cell population level. Relevantly, the distributions of intervals, including their mean and variance, are found to respond readily and discernibly to induction as well as temperature, revealing the plasticity of the expression mechanism of TetA. The plasticity appears to arise from the diversity of the changes in the durations of the various steps in response to differing induction levels and temperature.

Our results assume that the measurements are only affected by intrinsic noise sources in transcription. Downstream events, such as translation or RNA degradation, do not affect the results, as we detect RNA molecules as soon as these are produced and study only the time intervals between these events. It is necessary to discuss if other noise sources, aside intrinsic noise in transcription, could affect diversity in number of produced RNA molecules per cell. It may be that differences in the amounts of TetR and/or aTc in each cell could contribute to this diversity (i.e. be a significant source of extrinsic noise (70)). First, the strain used here overexpresses TetR (DH5αPRO produces constitutively around 7000 dimeric Tet repressors per cell during logarithmic growth (8)); thus, we expect the contribution of

diversity in TetR numbers to the diversity in RNA numbers to be negligible. As for aTc, we found no evidence, at least at the population level, of varying rates of transcription initiation because of varying concentrations of aTc over time (e.g. because of depletion), as we found no difference in the kinetics of RNA production during the first and second half of the measurement period. Additionally, the *tetA* gene, responsible for expressing TetA, which confers tetracycline resistance (11) (by pumping it out of the cell) is not present in our strain. Finally, diversity in the numbers of either aTc or TetR in the cells is expected to only increase diversity, not decrease it (and thus cannot be the explanation for the observed sub-poissonian kinetics). Nevertheless, it is stressed that our measurements are not sufficient to predict the cell to-cell-diversity in number of RNA molecules under the control of P_{tetA} as other factors, such as RNA degradation (45) and RNA partitioning in cell division (47) need to be considered.

In this regard, measurements by fluorescent *in situ* hybridization of RNA numbers under the control of 20 *E. coli* promoters (42,43) exhibited high values of cell-to-cell diversity. Another work reported Fano factors of mRNA numbers for 137 highly expressed genes in *E. coli*, ranging from 1 to 3, using a yellow fluorescent protein fusion library for *E. coli* (44). Although these results could be because of super-Poissonian transcript production from the strongly expressing promoters studied, other explanations cannot be ruled out. For example, it may be that although the production is sub-Poissonian, the observed diversity is a result of the subsequent contribution from complex RNA degradation mechanisms (45) and imperfect partitioning of RNA molecules in cell division ([47,48,51), among other things such as the biased segregation of unwanted substances (e.g. the fluorescent molecules used to tagged the RNA) (49,52). In any case, we note that our measured distributions of intervals between transcription events cannot be explained by models, such as the ''on-off'' model of transcription initiation (43,71), as this model entails super-Poissonian kinetics.

The kinetics of RNA production under the control of P_{tetA} can be explained by a model of transcription initiation with successive ''rate-limiting'' steps, each of which is exponentially distributed in duration. From the inference of the number and duration of these steps in several conditions, we found that induction with aTc significantly changes the RNA production kinetics by reducing the duration of all rate-limiting steps, to various degrees. In particular, one of the steps becomes almost indiscernible. Meanwhile, lowering temperature under full induction by aTc increases the duration of two of the steps, but not of the third step, causing the kinetics to be well-fit by a two-step model under these conditions. Note that it is not possible to determine which steps [e.g. closed complex, open complex or isomerization (28)] are affected by aTc and temperature. Novel experimental techniques are necessary to perform this study *in vivo*. However, we can rule out TetR dissociation from the promoter as one of the rate-limiting steps. The complex TetR-promoter has a half-life of 12 s (37), which is a much faster process than the measured rate-limiting steps. Additionally, although TetR, when bound by aTc, may retain some ability to bind to the DNA [although this ability is reduced by approximately 9 orders of magnitude (72)], there is no evidence that this complex would have a longer half-life than when the DNA is bound by TetR alone.

A recent work used the methods used here to analyse the kinetics of P_{lar} (40). The measurements were made at 24°C. It is of interest to compare them with our measurements regarding the response to induction. First, the kinetics of RNA production of P_{lar} is also sub-Poissonian. Also, induction of P_{lar} with IPTG and arabinose reduces the duration of the rate-limiting steps. The main differences between these two promoters are in the mean duration of the intervals at 24°C ([for P_{lar}, this mean duration is ~1500 s (40), whereas for P_{tetA}, it is ~1000 s]) and in the variability of the intervals. For P_{lar}, the CV^2 of the durations of these intervals is 0.70, whereas for P_{tetA}, it is 0.52. From this, we conclude that the kinetics of transcription initiation of P_{tetA} is less noisy than that of P_{lar}. It is worthwhile to mention that the observations of the behaviour of P_{tetA} (a native promoter) suggest that the sub-Poissonian kinetics of RNA production is not, for example, an artefact of the synthetic nature of P_{lar}, and that it may be a common feature of the dynamics of transcript production in *E. coli*. Also, the results support several previous observations on the effect of temperature on the kinetics of transcription, but further show that the changes in kinetics are due, in part, to the alteration of the mean duration of the intermediate steps in transcription initiation.

P_{tetA} controls the expression of TetA, which is responsible for the active efflux of tetracycline-Mg^{2+} complexes. This protein's function justifies the need for such a stringent regulatory mechanism so as to ensure that TetA is present in the appropriate amount because both tetracycline and TetA (in high amounts) are harmful to the cell (12). We find that this control is achieved not only by the negative feedback mechanism of the *tet* operon (12) but also by a sub-Poissonian kinetics of transcription initiation. Relevantly, although robust (less noisy than a Poisson process), this system is nevertheless sensitive to external stimuli, such as tetracycline, and temperature because its behaviour discernibly changes with temperature and inducer concentration.

In the future, it will be of interest to further analyse how the dynamics of P_{tetA} differs in other environmental conditions, such as in differing concentrations of hydrogen ions and metabolites. Such studies may provide insights on the plasticity of the kinetics of gene expression in bacteria and thus guide the engineering of synthetic genetic circuits with specific behavioural patterns.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2, Supplementary Figures 1 and 2 and Supplementary Movies 1 and 2.

ACKNOWLEDGEMENTS

The authors thank F.G. Biddle for useful advice. A.-B. M. executed the qPCR measurements.

FUNDING

Academy of Finland [126803 to A.S.R.]; Finnish Funding Agency for Technology and Innovation (TEKES) [40284/08 to O.Y.-H.]; TUT President's Doctoral Programme. Funding for open access charge: TEKES [40284/08 to O.Y.-H.].

Conflict of interest statement. None declared.

REFERENCES

1. McNaught,A.D. and Wilkinson,A. (1997) IUPAC. Compendium of Chemical Terminology. (the "Gold Book"). 2nd edn. Blackwell Scientific Publications, Oxford.
2. Duggar,B.M. (1948) Aureomycin: a product of the continuing search for new antibiotics. *Ann. N. Y. Acad. Sci.*, **51**, 177–181.
3. Connel,S.R., Tracz,D.M., Nierhaus,K.H. and Taylor,D.E. (2003) Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob. Agents Chemother.*, **47**, 3675–3681.
4. Schnappinger,D. and Hillen,W. (1996) Tetracyclines: antibiotic action, uptake, and resistance mechanisms. *Arch. Microbiol.*, **165**, 359–369.
5. Hillen,W., Scholimeier,K. and Gatz,C. (1984) Control of expression of the Tn10-encoded tetracycline resistance operon II. Interaction of RNA polymerase and TET repressor with the tet operon regulatory region. *J. Mol. Biol.*, **172**, 185–201.
6. Gossen,M. and Bujard,H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA.*, **89**, 5547–5551.
7. Nguyen,T.M.N., Phan,Q.G., Duong,L.P., Bertrand,K.P. and Lenski,R.E. (1989) Effects of carriage and expression of the Tn10 tetracycline-operon on fitness of *E. coli cells*. *Mol. Biol. Evol.*, **6**, 213–225.
8. Lutz,R. and Bujard,H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.*, **25**, 1203–1210.
9. Lazarevic,V., Margot,P., Soldo,B. and Karamata,D. (1992) Sequencing and analysis of the Bacillus subtilis lytRABC divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. *J. Gen. Microbiol.*, **138**, 1949–1961.
10. Collier,G.B., Mattson,T.L., Connaughton,J.F. and Chirikjian,J.G. (1994) A novel Tn10 tetracycline regulon system controlling expression of the bacteriophage T3 gene encoding S-adenosyl-L-methionine hydrolase. *Gene*, **148**, 75–80.
11. Beck,C.F., Mutzel,R., Barbe,J. and Müller,W. (1982) A multifunctional gene (*tetR*) controls Tn*10*-encoded tetracycline resistance. *J. Bacteriol.*, **150**, 63–642.
12. Hillen,W. and Berens,C. (1994) Mechanisms underlying expression of Tn*l0* encoded tetracycline resistance. *Annu. Rev. Microbial.*, **48**, 345–369.
13. Orth,P., Schnappinger,D., Hillen,W., Saenger,W. and Hinrichs,W. (2000) Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.*, **7**, 215–219.
14. Biliouris,K., Daoutidis,P. and Kaznessis,Y.N. (2011) Stochastic simulations of the tetracycline operon. *BMC Syst. Biol*, **5**, 9.
15. Skerra,A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene*, **151**, 131–135.
16. Bertrand,K.P., Postle,K., Wray,L.V. Jr and Reznikoff,W.S. (1984) Construction of a single-copy promoter vector and its use in analysis of regulation of the transposon Tn*10* tetracycline resistance determinant. *J. Bacteriol.*, **158**, 910–919.
17. Gardner,T.S., Cantor,C.R. and Collins,J.J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature*, **403**, 339–342.
18. Elowitz,M.B. and Leibler,S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335–338.
19. Becskei,A. and Serrano,L. (2000) Engineering stability in gene networks by autoregulation. *Nature*, **405**, 590–593.
20. Becskei,A., Séraphin,B. and Serrano,L. (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.*, **20**, 2528–2535.
21. Deans,T.L., Cantor,C.R. and Collins,J.J. (2007) A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell*, **130**, 363–372.
22. Nevozhay,D., Adams,R.M., Murphy,K.F., Josic,K. and Balázsi,G. (2009) Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 5123–5128.
23. Nandagopal,N. and Elowitz,M.B. (2011) Synthetic biology: integrated gene circuits. *Science*, **333**, 1244–1248.
24. Kaern,M., Elston,T.C., Blake,W.J. and Collins,J.J. (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, **6**, 451–464.
25. McClure,W.R. (1985) Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.*, **54**, 171–204.
26. Uptain,S., Kane,M. and Chamberlin,M. (1997) Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem.*, **66**, 117–172.
27. Browning,D.F. and Busby,S.J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, **2**, 57–65.
28. deHaseth,P.L., Zupancic,M.L. and Record,M.T. (1998) RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.*, **180**, 3019–3025.
29. Nierman,W.C. and Chamberlin,M.J. (1979) Studies of RNA chain initiation by *Escherichia coli* RNA polymerase bound to T7 DNA. Direct analysis of the kinetics and extent of RNA chain initiation at T7 promoter A1. *J. Biol. Chem.*, **254**, 7921–7926.
30. McClure,W.R. (1980) Rate-limiting steps in RNA chain initiation. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 5634–5638.
31. Lutz,R., Lozinski,T., Ellinger,T. and Bujard,H. (2001) Dissecting the functional program of *Escherichia coli* promoters: the combined mode of action of Lac repressor and AraC activator. *Nucleic Acids Res.*, **29**, 3873–3881.
32. Buc,H. and McClure,W.R. (1985) Kinetics of open complex formation between *Escherichia coli* RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry*, **24**, 2712–2723.
33. Kandhavelu,M., Lihavainen,E., Muthukrishnan,A.B., Yli-Harja,O. and Ribeiro,A.S. (2012) Effects of Mg2+ on in vivo transcriptional dynamics of the lar Promoter. *Biosystems*, **107**, 129–134.
34. Smolander,O.P., Kandhavelu,M., Mannerström,H., Lihavainen,E., Kalaichelvan,S., Healy,S., Yli-Harja,O., Karp,M. and Ribeiro,A.S. (2011) Cell-to-cell diversity in protein levels of a gene driven by a tetracycline inducible promoter. *BMC Mol. Biol.*, **12**, 21.
35. Bertrand-Burggraf,E., Lefèvre,J.F. and Daune,M. (1984) A new experimental approach for studying the association between RNA polymerase and the tet promoter of pBR322. *Nucleic Acids Res.*, **12**, 1697–1706.
36. Degenkolb,J., Takahashi,M., Ellested,G.A. and Hillen,W. (1991) Structural requirements of tetracycline-Tet repressor interaction: determination of equilibrium binding constants for tetracycline analogs with the Tet repressor. *Antimicrob. Agents Chemother.*, **35**, 1591–1595.
37. Murphy,K.F., Adams,R.M., Wang,X., Balázsi,G. and Collins,J.J. (2010) Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Res.*, **38**, 2712–2726.
38. Schubert,P., Pfleiderer,K. and Hillen,W. (2004) Tet repressor residues indirectly recognizing anhydrotetracycline. *Eur. J. Biochem.*, **271**, 2144–2152.
39. Oliva,B., Gordon,G., McNicholas,P., Ellestad,G. and Chopra,I. (1992) Evidence that tetracycline analogs whose primary target is not the bacterial ribosome cause lysis of *Escherichia coli*. *Antimicrob. Agents Chemother.*, **36**, 913–919.

40. Kandhavelu,M., Mannerström,H., Gupta,A., Häkkinen,A., Lloyd-Price,J., Yli-Harja,O. and Ribeiro,A.S. (2011) In vivo kinetics of transcription initiation of the lar promoter in *Escherichia coli*. Evidence for a sequential mechanism with two rate limiting steps. *BMC Sys. Biol.*, **5**, 149.
41. Saecker,R.M., Record,M.T. and deHaseth,P.L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.*, **412**, 754–771.
42. Golding,I., Paulsson,J., Zawilski,S.M. and Cox,E.C. (2005) Real-time kinetics of gene activity in individual bacteria. *Cell*, **123**, 1025–1036.
43. So,L.H., Ghosh,A., Zong,C., Sepúlveda,L.A., Segev,R. and Golding,I. (2011) General properties of transcriptional time series in *Escherichia coli*. *Nat. Genet.*, **43**, 554–560.
44. Taniguchi,Y., Choi,P.J., Li,G.W., Chen,H., Babu,M., Hearn,J., Emili,A. and Xie,X.S. (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, **329**, 533–538.
45. Yarchuk,O., Jacques,N., Guillerez,J. and Dreyfus,M. (1992) Interdependence of translation, transcription and mrna. *J. Mol. Biol.*, **5**, 581–596.
46. Llopis,P.M., Jackson,A.F., Sliusarenko,O., Surovtsev,I., Heinritz,J., Emonetand,T. and Jacobs-Wagner,C. (2010) Spatial organization of the flow of genetic information in bacteria. *Nature*, **466**, 77–81.
47. Huh,D. and Paulsson,J. (2010) Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.*, **43**, 95–100.
48. Huh,D. and Paulsson,J. (2011) Random partitioning of molecules at cell division. *Proc. Natl Acad. Sci. USA.*, **108**, 15004–15009.
49. Lindner,A.B., Madden,R., Demarez,A., Stewart,E.J. and Taddei,F. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc. Natl Acad. Sci. USA.*, **105**, 3076–3081.
50. Stewart,E., Madden,R., Paul,G. and Taddei,F. (2005) Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol.*, **3**, 295–300.
51. Lloyd-Price,J., Lehtivaara,M., Kandhavelu,M., Chowdhury,S., Muthukrishnan,A.B., Yli-Harja,O. and Ribeiro,A.S. (2011) Probabilistic RNA partitioning generates transient increases in the normalized variance of RNA numbers in synchronized populations of *Escherichia coli*. *Mol. Biosyst.*, **8**, 565–571.
52. Lloyd-Price,J., Häkkinen,A., Kandhavelu,M., Marques,I.J., Chowdhury,S., Lihavainen,E., Yli-Harja,O. and Ribeiro,A.S. (2012) Asymmetric disposal of individual protein aggregates in *Escherichia coli*, one aggregate at a time. *J. Bact.*, **194**, 1747–1752.
53. Yu,J., Xiao,J., Ren,X., Lao,K. and Xie,X.S. (2006) Probing gene expression in live cells. *Science*, **311**, 1600–1603.
54. Xie,X.S., Choi,P.J., Li,G.W., Lee,N.K. and Lia,G. (2008) Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.*, **37**, 417–444.
55. Le,T.T., Harlepp,S., Guet,C.C., Dittmar,K., Emonet,T., Pan,T. and Cluzel,P. (2005) Real-time RNA profiling within a single bacterium. *Proc. Natl Acad. Sci. USA.*, **102**, 9160–9164.
56. Korpela,M., Kurittu,J.S., Karvinen,J.T. and Karp,M.T. (1998) A recombinant *Escherichia coli* sensor strain for the detection of tetracyclines. *Anal. Chem.*, **70**, 4457–4462.
57. Golding,I. and Cox,E. (2004) RNA dynamics in live Escherichia coli cells. *Proc. Natl Acad. Sci. USA.*, **101**, 11310–11315.
58. Nevo-Dinur,K., Nussbaum-Shochat,A., Ben-Yehuda,S. and Amster-Choder,O. (2011) Translation-independent localization of mRNA in *Escherichia coli*. *Science*, **25**, 1081–1084.
59. Chini,V., Foka,A., Dimitracopoulos,G. and Spiliopoulou,I. (2007) Absolute and relative real-time PCR in the quantification of tst gene expression among methicillin-resistant Staphylococcus aureus: evaluation by two mathematical models. *Lett. Appl. Microbiol.*, **45**, 479–484.
60. Lee,C., Kim,J., Shin,S.G. and Hwang,S. (2006) Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *J. Biotec.*, **123**, 273–280.
61. Campbell,R.E., Tour,O., Palmer,A.E., Steinbach,P.A., Baird,G.S., Zacharias,D.A. and Tsien,R.Y. (2002) A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. USA.*, **99**, 7877–7882.
62. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
63. Shales,S.W., Chopra,I. and Ball,P.R. (1980) Evidence for more than one mechanism of plasmid-determined tetracycline resistance in *Escherichia coli*. *J. Gen. Microbiol.*, **121**, 221–229.
64. Chowdhury,S., Kandhavelu,M., Yli-Harja,O. and Ribeiro,A.S. (2011) An interacting multiple model filter based autofocus strategy for confocal time-lapse microscopy. *J. Microscopy*, **245**, 265–275.
65. Chen,T.B., Lu,H., Lee,Y.S. and Lan,H.J. (2008) Segmentation of cDNA microarray images by kernel density estimation. *J. Biomed. Info.*, **41**, 1021–1027.
66. Toulokhonov,I. and Landick,R. (2006) The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol. Cell*, **12**, 1125–1136.
67. Peabody,D.S. (1993) The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.*, **12**, 595–600.
68. China,A., Tare,P. and Nagaraja,V. (2010) Comparison of promoter-specific events during transcription initiation in mycobacteria. *Microbiology*, **156**, 1942–1952.
69. Arkin,A., Ross,J. and McAdams,H.M. (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics*, **149**, 1633–1648.
70. Elowitz,M.B., Levine,A.J., Siggia,E.D. and Swain,P.S. (2002) Stochastic gene expression in a single cell. *Science*, **297**, 1183–1186.
71. Peccoud,J. and Ycart,B. (1995) Markovian modelling of gene product synthesis. *Theor. Popul. Biol.*, **48**, 222–234.
72. Lederer,T., Takahashi,M. and Hillen,W. (1995) Thermodynamic analysis of tetracycline-mediated induction of Tet repressor by a quantitative methylation protection assay. *Anal. Biochem.*, **232**, 190–196.