Absolute quantitative measurement of transcriptional kinetic parameters in vivo

Sykanya Iyer, Emory University
Bo Ryoung Park, Emory University
Minsu Kim, Emory University

Journal Title: Nucleic Acids Research
Volume: Volume 44, Number 18
Publisher: Oxford University Press (OUP): Policy C - Option B | 2016-07-04,
Pages e142-e142
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1093/nar/gkw596
Permanent URL: https://pid.emory.edu/ark:/25593/rrw6s

Final published version: http://dx.doi.org/10.1093/nar/gkw596

Copyright information:
© The Author(s) 2016.
This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/).

Accessed April 6, 2018 8:42 AM EDT
Absolute quantitative measurement of transcriptional kinetic parameters \textit{in vivo}

Sukanya Iyer$^{1,†}$, Bo Ryoung Park$^{1,†}$ and Minsu Kim$^{1,2,*}$

$^1$Department of Physics, Emory University, Atlanta, GA 30322, USA and $^2$Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, GA 30322, USA

Received November 17, 2015; Revised June 13, 2016; Accepted June 20, 2016

ABSTRACT

mRNA expression involves transcription initiation, elongation and degradation. In cells, these dynamic processes are highly regulated. However, experimental characterization of the dynamic processes \textit{in vivo} is difficult due to the paucity of methods capable of direct measurements. We present a highly sensitive and versatile method enabling direct characterization of the dynamic processes. Our method is based on single-molecule fluorescence \textit{in situ} hybridization (smFISH) and quantitative analyses of hybridization signals. We hybridized multiple probes labelled with spectrally distinct fluorophores to multiple subregions of single mRNAs, and visualized the kinetics of synthesis and degradation of the sub-regions. Quantitative analyses of the data lead to absolute quantification of the lag time of mRNA induction (the time it takes for external signals to activate transcription initiation), transcription initiation rate, transcription elongation speed (i.e. mRNA chain-growth speed), the rate of premature termination of transcripts and degradation rates. Applying our method to three different biological problems, we demonstrated how our method may be applicable to reveal dynamics of mRNA expression that was difficult to study previously. We expect such absolute quantification can greatly facilitate understanding of gene expression and its regulation working at the levels of transcriptional initiation, elongation and degradation.

INTRODUCTION

Gene expression is a process by which genetic information is transcribed into mRNAs and translated into proteins. Living cells adapt to their surrounding environments by regulating this process and, especially, regulation of mRNA expression is the primary means (1). Thus, cellular adaptation strategies to environments are reflected in changes of mRNA expression and the characterization of mRNA expression is central in a wide range of biological research activities.

mRNA expression is dynamic, involving transcription initiation, elongation and degradation, and each process is tightly regulated (2–6). Recent biophysical and biochemical studies \textit{in vitro} have revealed intricate, molecular-level events in the processes, emphasizing that mRNA abundance in cells are determined by a complex interplay of these dynamic processes (7–10). It is becoming increasingly clear that to better understand mRNA expression in cells, \textit{in vivo} characterization of these dynamic processes is important. For such characterization, absolute quantification of kinetic parameters, such as transcription initiation rate, elongation speed, premature termination rate and degradation rate, is needed. With technological developments, it became straightforward to measure mRNA abundance \textit{in vivo}, e.g. using qPCR, northern blot, microarray, RNAseq, etc. (11,12). However, application of these methods to measure kinetic parameters related to mRNA expression is limited. Previously, the Northern blot and its variations (such as dot blots) have been used to measure transcription elongation speed (13,14). This method has poor detection limits, requiring considerable amounts of target mRNAs, and involves the use of hazardous radioactive isotopes (12). Moreover, absolute quantification of other kinetic parameters is not possible. There have been efforts to deduce such kinetic parameters from quantitative modelling of mRNA expression kinetics (15,16), but the values obtained are model-dependent.

Here, we introduce a highly sensitive and versatile method enabling absolute quantification of these kinetic parameters \textit{in vivo}. Our method builds on recent development of single-molecule fluorescence \textit{in situ} hybridization (smFISH) that allows for detection of individual mRNA molecules of interest (17–20). We hybridized multiple probes labelled with spectrally distinct fluorophores to multiple sub-regions of single mRNAs, and analysed the hybridization signals quantitatively. Revealing the kinetics of mRNA expression, the method allowed for absolute quan-
ification of various kinetic parameters with unprecedented accuracy. Using three biological problems as examples, we demonstrated the usefulness of this method.

MATERIALS AND METHODS

Bacterial strain and growth conditions

*Escherichia coli* K-12 strain NCM3722 (NMK1) or its derivative (NMK80) were used. In the NMK80 strain, the expression of the lacZ gene was controlled by the synthetic promoter P_{lac-O1} (See Supplementary Methods for the details of strain construction). TetR repressor proteins were constitutively expressed in the strain. Inducers, anhydrotetracycline (aTc) or chlorotetracycline (cTc) inhibit the repression of the promoter by TetR, thereby activating the expression of the lacZ mRNA molecules. aTc has a high affinity for TetR; 3 ng/ml gives induction at half maximum. Hence, controlling low levels of expression using aTc is tricky. In comparison, cTc has a low affinity; 40 ng/ml gives induction at half maximum. Thus, to gradually change the expression in the dose response curve in Figure 2, we used cTc. But, aTc diffuses into cells rapidly (~6 s as estimated below). Hence, to obtain the temporal induction curve after maximum induction (Figures 1B, C and 3), we used aTc.

Cells were grown in a 37°C water bath shaker shaking at 250 rpm. NMK80 cells were cultured in a Luria-Bertani (LB) medium (seed culture). For experiments using minimal media, cells were transferred to a N-C- minimal medium (21) with 20 mM glycerol as the carbon source and 20 mM NH₄Cl as the nitrogen source, and grown overnight (pre-culture). The next morning, cells were transferred to a fresh medium with the same composition to the OD₆₀₀ of 0.02 (experimental culture). To obtain the dose response curve in Figure 2, various concentrations of cTc were added to the fresh medium at the time of the transfer and samples for FISH were collected at OD₆₀₀ of ~0.5. To obtain the temporal induction curve (Figures 1B, C and 3), 100 ng/ml of aTc was added to the culture at an OD₆₀₀ of ~0.5 and samples for FISH were collected at various time points after the addition. For the diauxic shift experiment (yielding the data plotted in Figure 4A), NMK1 was cultured in a minimal medium with 2.5 mM of glucose and 10 mM of glycerol as carbon sources, and 20 mM of ammonium as the nitrogen source. For the ammonium starvation experiment (yielding the data shown in Table 1 and Figure 4B), NMK1 was cultured in a minimal medium with 3 mM of ammonium as the nitrogen source and 10 mM of glycerol.

Fluorescence in situ hybridization

We followed the procedure described in (17); please see Supplementary Methods for the details of the procedure. Briefly, for the probe design, the software, Stellaris Probe Designer, from Biosearch Technologies was used. The sequences of the probes used were provided in Supplementary Table. The sets of probes binding to the head, body and tail regions were labelled with Atto 647N NHS ester (AD488-35, ATTO-TEC GmbH), Atto 488 NHS ester (AD647N-35, ATTO-TEC GmbH) and 6-TAMRA NHS ester (C6123, Invitrogen), respectively. Cell fixation and probe hybridization procedures are as follows. A total of 2.7 ml of the cell culture was collected at various time points (as noted in Figures) and mixed with 300 μl of 37% formaldehyde immediately. After incubating the culture in formaldehyde for 30 min, cells were spun down at 1000 g for 8 min. The supernatant was discarded and the cells were washed twice with 1× PBS. Finally cells were re-suspended in 300 μl of DEPC-treated water and 100% ethanol was added to the final volume of 1 ml. The cells were incubated either at room temperature for 1 h or stored at 4°C overnight. Then, cells were spun down again and re-suspended in the hybridization buffer (2x SSC, 40% formamide, 0.1% w/v dextran sulfate, 10 μg tRNA, 0.1% BSA and 0.5 mM ribonucleoside vanadyl complex). Fluorescent probes to a final concentration of 120 nM (total) were used to hybridize mRNAs in the cells and incubated overnight at 30°C. The following day, cells were re-suspended in 1 ml of wash buffer and spun down at 800 g for 7 min. The supernatant was removed and the cell pellet was washed three times using the wash buffer to remove unhybridized excess probes. Then, cells were re-suspended in 10 μl of 2x SSC buffer before imaging.

Microscope imaging and analysis

The cells were placed on the microscope (Olympus IX83P2Z) and imaged using a cooled cSMOS camera (Andor Neo). The best focal plane for imaging (z-position) was first determined using phase contrast. Then, varying the z-position at a spacing 300 nm, nine fluorescence images were acquired. The image analysis was performed using a custom-built program (based on MATLAB). Briefly, cell boundaries were first determined using a cell segmentation program. Intensities of fluorescence foci within cells were quantified. The intensity values were converted to the copy number of mRNAs per cell as described in Supplementary Figure S1.

β-galactosidase assay

One hundred microlitres of cell culture were collected and fast-frozen on dry ice. Four samples were collected during exponential growth. The β-Galactosidase activity of each sample was measured at 37°C by the traditional Miller method (22). The activities obtained were plotted against the OD₆₀₀ at which the sample was collected. The slope from the plot was taken to be the mean activity and was used in Figure 2.

RESULTS

Visualization of dynamics of mRNA expression

Using lacZ mRNAs in *E. coli* as an example, we describe how multicolour labelling of multiple sub-regions of single mRNAs can visualize the dynamics of mRNA expression. We designed three sets of oligonucleotide probes hybridized to the head, body and tail of the lacZ mRNA molecules; see Figure 1A, and Supplementary Table for the sequences of the probes. Each set is comprised of 25–28 probes and each probe is 20 nucleotides (nt) long. The three sets of the probes
are labelled with red (for head), green (for body) and blue (for tail) fluorophores; note that these are pseudo-colours assigned for illustration. See Supplementary Table for the sequences of the probes. (B) At time zero, 100 µg/ml of aTc (inducer) was added to the culture. Red fluorescence foci (head), green foci (body) and blue foci (tail) appear sequentially in cells. The scale bar in the image of the tail at 90 s (black line) represents 1 µm and applies to all the images in Figure 1B. (C) Even after the mRNA levels reach their steady state (approximately after 10 min of induction), the colours of fluorescence foci in cells are not uniform, revealing the dynamics of synthesis and degradation of different sub-regions of mRNA molecules. The scale bar (black line) represents 1 µm. See the text for details.

**Figure 1.** Visualization of dynamics of mRNA expression. (A) The three sets of probes were hybridized to the head, body and tail of the lacZ mRNA molecules. They were labelled with red (for head), green (for body) and blue (for tail) fluorophores; note that these are pseudo-colours assigned for illustration. See Supplementary Table for the sequences of the probes. (B) At time zero, 100 µg/ml of aTc (inducer) was added to the culture. Red fluorescence foci (head), green foci (body) and blue foci (tail) appear sequentially in cells. The scale bar in the image of the tail at 90 s (black line) represents 1 µm and applies to all the images in Figure 1B. (C) Even after the mRNA levels reach their steady state (approximately after 10 min of induction), the colours of fluorescence foci in cells are not uniform, revealing the dynamics of synthesis and degradation of different sub-regions of mRNA molecules. The scale bar (black line) represents 1 µm. See the text for details.

**Figure 2.** Comparison of the mean number of the head, body and tail of the lacZ mRNA molecules per cell (N_{head}, N_{body} and N_{tail}) and LacZ protein amounts. A synthetic promoter P_{tet-O1} drives the expression of the lacZ gene in our strain (23). We added various concentrations of the inducer, cTc, to exponentially-growing cultures (the cTc is preferable for induction at low levels due to its low affinity; see Materials and Methods). After several doublings, the expression of lacZ mRNAs and LacZ proteins reaches the steady state. We observed a good agreement between the mRNA copy number per cell (red circles for N_{head}, green squares for N_{body} and blue triangles for N_{tail}; left axis) and LacZ protein amounts (right axis). Note that the left and right axes have the same scale.

**Quantifying the transcription elongation speed**

We wish to quantify various kinetic parameters related to mRNA expression from these images. To do so, we first converted the fluorescent foci intensities to mRNA counts using the procedure that was previously established (17,24), and obtained the mean copy number of the head, body and tail of lacZ mRNAs per cell, N_{head}, N_{body} and N_{tail}, respectively; see Supplementary Figure S1 for illustration. As a check of consistency, we plotted the steady state value of N_{head}, N_{body} and N_{tail}, together with the steady state levels of LacZ proteins (measured from a standard β-galactosidase assay) at various inducer concentrations in Figure 2, and observed a good agreement.

The sequential accumulation of hybridization signals from the head, body and tail of the mRNA in Figure 1B contains the information about the lag time of mRNA induction, transcription initiation rate, and transcription elongation speed. To quantify these parameters, we plotted N_{head}, N_{body} and N_{tail} as a function of time in Figure 3A (note that mRNA expression was fully induced by adding a saturating concentration of the inducer, aTc, to the medium at time zero). We see that N_{head}, N_{body} and N_{tail} begin to increase at different time points, revealing the duration of the time needed for cells to synthesize each segment of the mRNA. Fitting the linear portion of the increase, we obtained the X intercept of the head and tail, T_{head} = 25.8 (±0.5) s and T_{tail} = 69.8 (±2.9) s; the standard deviation is
Nucleic Acids Research, 2016, Vol. 44, No. 18

Figure 3. Quantification of mRNA expression and degradation dynamics. (A) Analysing hybridization signals from the head, body and tail of the lacZ mRNA (typical images used were shown in Figure 1B), we obtained \(N_{\text{head}}\) (red circles), \(N_{\text{body}}\) (green squares) and \(N_{\text{tail}}\) (blue triangles) as a function of time in cells growing exponentially in a minimal medium with glucose and ammonium as the sole carbon and nitrogen sources. \(T_{\text{head}}, T_{\text{body}}\) and \(T_{\text{tail}}\) indicate the times at which \(N_{\text{head}}, N_{\text{body}}\) and \(N_{\text{tail}}\) begin to rise; the actual values were obtained from the fit of the regions of the linear increase (dashed lines). From these values, we obtained the transcription elongation speed. From the speed and \(T_{\text{head}}\), we estimated \(T_{\text{lag}}\) as the time it takes for the inducer to activate transcription initiation. From the slopes, we obtained the transcription rate, as well as the rate of premature termination of transcripts. See the text for the details of the analyses. (B) We inhibited the mRNA synthesis by adding rifampicin to an exponentially-growing culture; the addition of rifampicin marks the time zero. \(N_{\text{head}}, N_{\text{body}}\) and \(N_{\text{tail}}\) decrease exponentially (note a semi-log scale). \(N_{\text{head}}\) begins to decrease at earlier time points than \(N_{\text{tail}}\); see the insets.

calculated from the measurement of three independent experiments. The time difference, \(T_{\text{lag}} - T_{\text{head}}\), represents the ‘transit time’ of RNA polymerases on the body and tail segments of the lacZ mRNA molecules. From the difference, considering the known length of the sequence hybridized by the probes (Figure 1A), we obtain the transcription elongation speed of \(48.6\ \text{nt/s} = (1079 + 1061\ \text{nt}) / (69.8 - 25.8\ \text{s})\). This value agrees with a previous estimate of \(45\ \text{nt/s}\) in a similar growth condition (doubling time \(\sim 60\ \text{min}\) using the Northern blots (Table 1 and in (25)).

The analyses of different combinations, e.g. \(N_{\text{head}}\) and \(N_{\text{body}}\), can be performed in an identical manner and yield a similar result. However, we note that the body of the lacZ mRNA was hybridised with probes labelled with the Atto 488 that emits green fluorescence. This emission spectrum overlaps with auto-fluorescence spectrum of E. coli cells, resulting in a relatively high background fluorescence. Thus, we found it generally difficult to accurately determine the fluorescence intensities of green foci and, hence, \(N_{\text{body}}\).

Quantifying the lag time of mRNA induction

Next, we estimated the lag time of mRNA induction, \(T_{\text{lag}}\), the time delay between the addition of the inducer to the medium and transcription initiation; see Figure 3A. Assuming that the transcription elongation speed remains constant, we deduced that it would take \(T_{\text{synthesis}} = 19.2\ \text{s} = 935/48.6\) to synthesize the head of the lacZ mRNA molecule. This is shorter than \(T_{\text{head}} = 25.8\ \text{s}\), the time it takes for the head to appear after the addition of the inducer at the time zero, because \(T_{\text{head}}\) includes not only \(T_{\text{synthesis}}\), but also the time delay, \(T_{\text{lag}}\). Hence, from the difference between \(T_{\text{head}} = 25.8\ \text{s}\) and \(T_{\text{synthesis}} = 19.2\ \text{s}\), we obtain \(T_{\text{lag}} = 6.6\ \text{s} = 25.8 - 19.2\). When we calculate \(T_{\text{lag}}\) using the tail, we obtained a similar value; it would take \(63.3\ \text{s} = 3075/48.6\) to synthesize the tail, and the difference of this value from \(T_{\text{lag}} = 6.5\ \text{s} = 69.8 - 63.3\), showing a good agreement with the estimate above. However, there are other factors contributing to this difference. First, our procedure involves formaldehyde fixation that may take some time (\(\sim\)s). Also, as elaborated in Supplementary Note 1, there is an uncertainty in determination of a full-length segment of the head of lacZ mRNA molecules, though the error from this uncertainty is estimated to be small. Hence, the actual time delay may be a bit longer than \(6.6\ \text{s}\). Note that these factors are expected to contribute equally to the rise of the head and tail, and hence, when the time difference \((T_{\text{head}} - T_{\text{tail}})\) is used to calculate the transcription elongation speed, the resultant value (of the speed) is not affected by these factors.

Quantifying the transcription rate and premature termination rate

Next, we deduced the transcription rate, i.e. how frequently each segment of mRNAs (head, body and tail) is synthesized. The law of mass action describes that the change of the mRNA concentration, \(m\), is determined by the balance between the transcription rate, \(\alpha\), and the degradation rate, \(\beta\), in the following way,

\[
\frac{dm}{dt} = \alpha - \beta \cdot m
\]  

Hence, immediately after induction, when the mRNA level is low \((m \approx 0)\), we have \(\frac{dm}{dt} \approx \alpha\) and \(m \approx \alpha \cdot t\). Thus, the slope in the linear increase of the mRNA level after induction in Figure 3A (dashed lines) is equal to the transcription rate, \(\alpha\). Analysing the slopes, we obtained \(5.0 (\pm 0.3)\) (min for the transcription rate of the head and tail of the mRNA, respectively.

Comparison of the transcription rates of the head and tail can reveal premature termination of transcription. For example, consider that once RNA polymerases initiate transcription, they continue through head, body and tail segments, completing transcription. In that case, the transcription rate of the three segments will always match (after time delay determined by the transcription elongation speed). Hence, the close match of the transcription rates observed...
Effects of cell physiology on the dynamics of mRNA expression

The physiological state of cells exerts significant effects on gene expression, forming an additional layer of control of gene expression (26). There exist extensive studies to characterize these effects by growing cells at different rates in different steady-state conditions (25,27,28). However, because it was difficult to measure dynamics of mRNA expression, these studies have focused on the effects at the protein level and much less is known about the effects at the mRNA level. Here, repeating the experiments above in cells growing at different growth rates and characterizing the kinetic parameters of mRNA expression, we revealed the effects of physiology on mRNA expression. The results are summarized in the Table and the raw data are plotted in Supplementary Figure S2. The results show that the values of all the kinetic parameters are dependent on the rate of cell growth, revealing the effects of physiology on gene expression at the levels of transcriptional initiation, elongation and degradation.

**Quantifying the degradation rate**

Next, we characterized the degradation dynamics of the head, body and tail of the mRNA molecules. As described above, fluorescence foci at the steady state of mRNA expression in Figure 1C exhibit various colours, indicating the dynamics of synthesis and degradation of different sub-regions of mRNA molecules. Detailed analyses of such images can reveal degradation dynamics of mRNAs. Here, we took a different approach. To quantify the degradation rate, $\beta$, of the head, body and tail of the lacZ mRNAs (see Equation 1), we inhibited the mRNA synthesis by adding rifampicin to an exponentially-growing culture. We observed that $N_{\text{head}}$, $N_{\text{body}}$ and $N_{\text{tail}}$ decrease exponentially in Figure 3B (note a semi-log scale); the addition of rifampicin marks the time zero. Fitting the region of the exponential decrease, we obtained that for the cells growing at the rate of 0.7/h (in a medium with glycerol and ammonium as the sole carbon and nitrogen sources), the degradation rates, $\beta$, for the head, body and tail, are 0.47 (±0.05)/min, 0.49 (±0.06)/min and 0.5 (±0.02)/min; see Equation (1). From these values, we obtain the half-lives of the head, body and tail ( = ln2/$\beta$) are 1.5 min, 1.4 min and 1.4 min, respectively. We note a time delay between the onset of decrease of $N_{\text{head}}$ and that of $N_{\text{tail}}$; see the inset in Figure 3B (as described above, it is difficult to accurately determine $N_{\text{body}}$ due to a relatively high background fluorescence, which is why the data points of $N_{\text{body}}$ are more scattered and it is difficult to notice the time delay from $N_{\text{body}}$). As will be shown later, such a delay is more pronounced in cells in a different physiological condition (Supplementary Figure S2). One factor contributing to the time delay may be the presence of on-going transcription. If rifampicin inhibits transcription initiation, but not elongation, immediately after the rifampicin treatment, there will be RNA polymerases that are already transcribing, and will continue to produce the mRNAs for some time. This could result in a time delay in the decrease of mRNA levels, and the delay will be more pronounced for the tail. Alternatively, this delay could result from distinct degradation kinetics of the different sub-regions of mRNAs.

Alternatively, we can deduce the degradation rate using the steady state level of mRNAs together with the transcription rate, $\alpha$, obtained above; at the steady state, the Equation (1) yields $\beta = \alpha/m_s$, where $m_s$ refers to the copy number of mRNAs in the steady state. When mRNA expression is fully induced, we observed that $m_s$ of the head is ~6.7 (the value is very similar for the tail). The value, together with $\alpha = 5.0$/min (as obtained above for the head and also similarly for the tail), yields the degradation rate of 0.75/min. This is less than 2-fold difference from the value obtained using rifampicin above, showing reasonable agreement.

Having demonstrated how our method allows for absolute quantification of dynamics of mRNA expression, we now describe its potential use in different biological problems.

**Transcription elongation and degradation**

Effects of cell physiology on the synthesis rate of mRNAs

The transcription rate, $\alpha$, can be estimated using a set of probes hybridized to narrower sub-regions of the mRNA molecules. Detailed analyses of such images can reveal transcription elongation dynamics of mRNAs. Here, we took a different approach. To quantify the transcription elongation rate, $\alpha$, of the head, body and tail of the lacZ mRNAs (see Equation 1), we inhibited the mRNA synthesis by adding rifampicin to an exponentially-growing culture. We observed that $N_{\text{head}}$, $N_{\text{body}}$ and $N_{\text{tail}}$ decrease exponentially in Figure 3B (note a semi-log scale); the addition of rifampicin marks the time zero. Fitting the region of the exponential decrease, we obtained that for the cells growing at the rate of 0.7/h (in a medium with glycerol and ammonium as the sole carbon and nitrogen sources), the transcription rates, $\alpha$, for the head, body and tail, are 7.6 (±0.3)/min and 4.9 (±0.2)/min for the head and tail, respectively (see Supplementary Figure S2). One factor contributing to the time delay may be the presence of on-going transcription. If rifampicin inhibits transcription initiation, but not elongation, immediately after the rifampicin treatment, there will be RNA polymerases that are already transcribing, and will continue to produce the mRNAs for some time. This could result in a time delay in the decrease of mRNA levels, and the delay will be more pronounced for the tail. Alternatively, this delay could result from distinct degradation kinetics of the different sub-regions of mRNAs.

Alternatively, we can deduce the transcription elongation rate using the steady state level of mRNAs together with the transcription rate, $\alpha$, obtained above; at the steady state, the Equation (1) yields $\alpha = \beta m_s$, where $m_s$ refers to the copy number of mRNAs in the steady state. When mRNA expression is fully induced, we observed that $m_s$ of the head is ~6.7 (the value is very similar for the tail). The value, together with $\alpha = 5.0$/min (as obtained above for the head and also similarly for the tail), yields the transcription elongation rate of 49.5 nt/s. This is less than 2-fold difference from the value obtained using rifampicin above, showing reasonable agreement.

Having demonstrated how our method allows for absolute quantification of dynamics of mRNA expression, we now describe its potential use in different biological problems.

Effects of cell physiology on the dynamics of mRNA expression

The physiological state of cells exerts significant effects on gene expression, forming an additional layer of control of gene expression (26). There exist extensive studies to characterize these effects by growing cells at different rates in different steady-state conditions (25,27,28). However, because it was difficult to measure dynamics of mRNA expression, these studies have focused on the effects at the protein level and much less is known about the effects at the mRNA level. Here, repeating the experiments above in cells growing at different growth rates and characterizing the kinetic parameters of mRNA expression, we revealed the effects of physiology on mRNA expression. The results are summarized in the Table and the raw data are plotted in Supplementary Figure S2. The results show that the values of all the kinetic parameters are dependent on the rate of cell growth, revealing the effects of physiology on gene expression at the levels of transcriptional initiation, elongation and degradation.
**Figure 4.** Dynamics of expression of *glpK* mRNA and *amtB* mRNA molecules in dynamically-changing environments. (A) Cells were cultured in a minimal medium with glucose and glycerol as carbon sources (ammonium as the sole nitrogen source). Cells initially grew on glucose (the shaded region). At the time zero, their growth abruptly stopped (indicating depletion of glucose), and after some time, cells began to grow on glycerol (see Supplementary Figure S3A for growth curve). During this period, the expression of mRNA molecules of glycerol kinase *glpK* was measured. The two sets of probes were hybridized to the head and tail of *glpK* mRNA molecules. The sequences of the probes and probe-binding regions were provided in Supplementary Table. 

* N <sub>head</sub> (red circles) increases before cell growth stops. 

* N <sub>tail</sub> (blue triangles) increases after a time delay. See the text for the details.

(B) Cells were cultured with a limited amount of ammonium as the nitrogen source (and glycerol as the carbon source). As cells use ammonium to grow, the ammonium level in the medium decreases and eventually cell growth stops due to depletion of ammonium (see Supplementary Figure S3D for growth curve). The onset of the growth stop defines the time zero. The two sets of probes were hybridized to the head and tail of *amtB* mRNA molecules. The sequences of the probes and probe-binding regions were provided in Supplementary Table. We see that *N* <sub>head</sub> (red circles) increases before cell growth stops due to ammonium depletion, indicating that cells anticipate impending ammonium depletion. 

* N <sub>tail</sub> (blue triangles) increases after a time delay. The shaded area indicates the period of cell growth, adopted from Supplementary Figure S3D. See the text for the details.

**Dynamics of mRNA expression in changing environments**

Natural environments change dynamically, and to survive and proliferate under such environments, organisms must regulate gene expression accordingly. Here, using changing nutrient availability as an example, we demonstrate the potential use of our method. When presented with two carbon sources, microbes exhibit diauxic growth (29,30). They first grow on the preferred carbon source. When the preferred source is depleted, their growth stops and after some period, they begin to grow on the less-preferred carbon source. In such conditions, the fitness of microbes depends on how rapidly they can adapt to the second carbon source after the depletion of the preferred carbon source, and the dynamics of the adaptation is determined by dynamics of expression of genes required to utilize the second carbon source, e.g. how early and how efficiently these genes are transcribed. However, such dynamics at the mRNA level is not well known. Here, we characterized the dynamics of mRNA expression of the glycerol kinase (*glpk*), an essential gene for glycerol metabolism, as an example.

We measured the kinetic parameters of its mRNA expression (driven by its own endogenous promoter), as cells undergo diauxic shift in the medium containing glucose (the preferred carbon source) and glycerol. We hybridized one set of probes labelled with Atto 647N to the first half of the *glpk* mRNA molecule (head) and the other set labelled with 6-TAMRA to the second half (tail); see Supplementary Table for the detail. The onset of the growth stop due to the depletion of glucose defines the time zero (Supplementary Figure S3A). As discussed above, because how rapidly cells can adapt to the second carbon source depends on how early and efficiently the required genes are expressed, we wish to determine when transcription of the *glpK* operon is activated, how fast it is transcribed (elongation speed), and if transcription terminates prematurely or not. The dynamics of *N* <sub>head</sub> and *N* <sub>tail</sub> of *glpK* mRNA molecules plotted in Figure 4A reveals that *N* <sub>head</sub> begins to increase before cell growth stops at the time zero; see also Supplementary Figure S3B for the direct comparison of the dynamics of cell growth and *N* <sub>head</sub> of *glpK* mRNA molecules. The increase of *N* <sub>head</sub> is followed by the increase of *N* <sub>tail</sub> with a time delay. The initial increase of the *N* <sub>head</sub> and *N* <sub>tail</sub> can be approximated as linear and, fitting the first 4 points after
the onsets of the increase with a linear line (dashed line), we obtained that the onset of the increase of \( N_{\text{head}} \) and \( N_{\text{tail}} \) of the mRNA molecule is \(-4.6\) min and \(-1.9\) min, respectively. Using this delay, together with the known length of the sequence hybridized in the tail region (Supplementary Table), we obtain the transcription elongation speed of 4.5 nt/s \( \left[ = \frac{723\ \text{nt}}{(4.6 - 1.9)\ \text{min}} \times \frac{1\ \text{min}}{60\ \text{s}} \right] \). Assuming that the transcription elongation speed is constant within the gene, with the known length of the head (754 nt), we estimate that the transcription of the gene glpK is initiated 2.8 min \( \left[ = \frac{754\ \text{nt}}{(4.5\ \text{nt/s})} \times \frac{1\ \text{min}}{60\ \text{s}} \right] \) before the onset of the increase of \( N_{\text{head}} \), i.e. 7.4 min \( (4.6 + 2.8) \) min before cell growth stops due to the glucose depletion, revealing that cells anticipate the impending shift. The transcription rate obtained from the slopes of the dashed lines in Figure 4A are 0.23/min and 0.22/min for \( N_{\text{head}} \) and \( N_{\text{tail}} \), respectively. The match of the transcription rate indicates that premature termination of transcription is negligible. After the initial linear increase of \( N_{\text{head}} \) and \( N_{\text{tail}} \), however, the increase becomes steeper, deviating from the linear dashed line. This steep increase indicates that the rate of the transcription increases. Indeed, when we estimated the rate from the increase of the mRNA copy numbers between consecutive time points, we see that the rate increases over time (Supplementary Figure S3C). In addition, cell-to-cell variability of protein expression during diauxic shift has attracted much attention in recent years (31–33). Our method based on smFISH can reveal the mRNA copy number with single-cell resolution (Figure 1C and Supplementary Figure S4) and, if needed, can be used for studies of cell-to-cell variation of mRNA expression during diauxic shift.

Lastly, when the amounts of nutrients in an environment are decreasing, it is important for bacteria to efficiently scavenge the limiting nutrients. The efficient scavenging depends on gene expression of nutrient transporters required for scavenging.AmtB is the ammonium transporter required to transport ammonium present at low levels (34,35), and here we examined dynamics of its mRNA expression (driven by its endogenous promoter), as the ammonium level in the environment is reduced. We hybridized one set of probes labelled with Atto 647N to the first half of the \( \text{amtB} \) mRNA molecule (head) and the other set labelled with 6-TAMRA to the second half (tail); see Supplementary Table for the details. The dynamics of \( N_{\text{head}} \) and \( N_{\text{tail}} \) of the \( \text{amtB} \) mRNA is plotted in Figure 4B. Analysing the data similarly as described for the data from \( \text{glnK} \) above, we obtained that the transcription of the \( \text{amtB} \) gene is activated 16.6 min before cell growth stops due to ammonium depletion (see Supplementary Figure S3D for the growth curve) and proceeds at the elongation speed of 33.9 nt/s. The examples here demonstrate how our method may be used to characterize dynamics of mRNA expression in various conditions, including changing environments.

**DISCUSSION**

Regulation of mRNA expression according to external environments is crucial for organisms’ survival and proliferation. A growing number of biophysical and biochemical studies in vitro, revealing intricate molecular events during transcription, emphasizes that complex interplay of transcription initiation, elongation, premature termination and degradation governs the expression of mRNAs (7–10). Hence, to better understand the mRNA expression in vivo, absolute quantitative measurement of the kinetic parameters is crucial.

We have shown that our method to label sub-regions of single mRNAs with different fluorophores and quantitatively analyse the signals is applicable for such absolute quantification of the kinetic parameters. Importantly, our method based on smFISH is capable of identifying single mRNA molecules. Thus, it can be used to characterize kinetic parameters even when the expression of mRNAs is very low. Investigating the effects of cell physiology on mRNA expression, as well as dynamics of mRNA expression in changing environments as examples, we demonstrated how our method may be used to reveal new knowledge of gene expression and its regulation which was difficult to obtain previously. In addition, we expect various other applications of our method. For instance, a certain sub-region within a mRNA molecule can be particularly important for regulation of mRNA expression; a good example is the formation of hairpin structures of \( \text{trp} \) mRNAs that results in premature termination of transcription (36). In such cases, probes can be designed to hybridize the upstream and downstream of the sub-region to pinpoint the role of such regulations. Also, although we characterized expression kinetics of one gene here, with appropriate combinations of spectrally distinct fluorophores, the expression kinetics of more than one gene may be examined simultaneously. We note that there exist previous studies to use multiple-labelling of single mRNAs using FISH (37,38). However, in the studies, such multiple-labelling was used to uniquely determine the identities of mRNAs, and hence, different from our application.

Another advantage of our method based on smFISH is its low detection limits and sensitivity. Because a single mRNA molecule can be visualized (Figure 1), the dynamics of mRNA expression at very low levels can be characterized quantitatively. Furthermore, our method allows discernment of the copy number of mRNAs in individual cells. For example, analysing image sequences shown in Figure 1B, we plotted how the number of cells expressing mRNAs changes after induction, and how the distribution of the mRNA copy number per cell changes after induction in Supplementary Figure S4. Hence, our method can be applicable to measure the dynamics of mRNA expression at single-molecule and single-cell resolution.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors thank Hedia Maamar for her help during the early development of this method, and Ido Golding for kindly answering questions regarding the previously-published protocol of smFISH from his laboratory and providing us with test probes early on.
FUNDING

MK’s Emory start-up fund. Funding for open access charge: Emory start-up fund.

Conflict of interest statement. None declared.

REFERENCES