

## Coarse-Grained Dynamics of Protein Synthesis in a Cell-Free System

Eyal Karzbrun,<sup>1</sup> Jonghyeon Shin,<sup>2</sup> Roy H. Bar-Ziv,<sup>1</sup> and Vincent Noireaux<sup>2</sup>

<sup>1</sup>*Department of Materials and Interfaces, the Weizmann Institute of Science, Rehovot, Israel*

<sup>2</sup>*Department of Physics, University of Minnesota, Minneapolis, Minnesota, 55455, USA*

(Received 10 September 2010; published 24 January 2011)

A complete gene expression reaction is reconstituted in a cell-free system comprising the entire endogenous transcription, translation, as well as mRNA and protein degradation machinery of *E. coli*. In dissecting the major reaction steps, we derive a coarse-grained enzymatic description of biosynthesis and degradation, from which ten relevant rate constants and concentrations are determined. Governed by zeroth-order degradation, protein expression follows a sharp transition from undetectable levels to constant-rate accumulation, without reaching steady state.

DOI: 10.1103/PhysRevLett.106.048104

PACS numbers: 87.18.Vf

Protein synthesis is among the most important and complex sets of reactions in the cell, with over 100 different components involved in transcription and translation (TX/TL) [1]. The interplay between synthesis and degradation creates a source-sink nonequilibrium system leading to pattern formation when coupled to nonlinearity and feedback, as observed in gene circuits [2], morphogenesis [3], and decision making [4]. Quantitative description of the complete reaction dynamics and information flow between its stages is, therefore, essential for understanding the physics of gene expression in isolated circuits [5] and in whole-cell function [6]. Despite the significant advances in single-cell measurements, it remains a challenge to determine the numerous biochemical parameters, to deal with unavoidable cross talk and resource sharing between expression reactions, and to monitor the changes in cellular states, all simultaneously. Consequently, our understanding of the reaction dynamics is often based on *ad hoc* assumptions and incomplete knowledge of the important parameters. A cell-free system offers a means to reconstitute complete TX/TL reactions, and to investigate the space of parameters. *In vitro* systems based on cell extract or purified components [7] are emerging tools to construct TX circuits [8–10], artificial cells [11], active biochips [12,13], and DNA-based devices [14,15]. Hence, improved phenomenology of cell-free reactions is bound to impact our ability to design novel systems.

In this Letter we investigate the dynamics of a complete TX/TL/degradation reaction *in vitro*. In contrast to commonly used cell-free systems, in which machinery from different organisms is combined to optimize overall production, the one of choice here is based on cell extract comprising *E. coli* machinery solely [16,17]. The system thus uniquely offers use of *E. coli*'s promoters and regulatory repertoire for cascading gene networks, and here we dissected the dynamics of one gene only (reporter GFP). In addition, the extract demonstrates consistent expression dynamics over a wide range of nutrients and building blocks (rNTP, amino acids, tRNA), rendering the system

suitable for quantitative study [16]. Our objective was to formulate a qualitative description of the mRNA and protein dynamics and to determine the typical scales. Despite the complexity of synthesis and degradation reactions, each involving several enzymes and substrates, a coarse-grained model based on four enzymes and ten free parameters was sufficient for a consistent phenomenology. Our main results are the following. (i) TX/TL machinery concentration and rates are 1–2 orders of magnitude lower than *in vivo* values [18], consistent with the volume dilution from *E. coli* (300 mg/ml) to cell extract (10 mg/ml), and with partial loss of specific activity *in vitro*. (ii) mRNA exhibits exponential degradation and reaches steady state in the presence of TX. (iii) The mRNA per DNA ratio at steady state is of order 1. Yet, the overall protein per DNA amplification is roughly 50-fold after 1 h without protein degradation. (iv) *E. coli*'s ClpXP protein degradation complex in the extract exhibits zeroth-order enzymatics. (v) Consequently, protein levels do not achieve steady state, and a sharp transition between undetectable levels and constant-rate accumulation is observed when TL rate exceeds degradation.

*Experimental.*—The cell extract is a crude cytoplasmic extract from *E. coli* which contains soluble proteins (above a 10 kDa cutoff) including the entire endogenous TX/TL machinery, as well as mRNA and protein degradation enzymes [16]. Reaction typically took place in volumes of 20–50  $\mu$ l at 28 °C. Two variants of GFP were used: deGFP-ssrA specifically degraded by ClpXP, a key tag-specific endogenous protease, and deGFP-ssrA-DD, a non-degradable variant [17]. mRNA dynamics were followed using radioactively labeled rNTPs and protein dynamics by fluorescence. Details of the experimental system and procedures are in the supplemental material [19].

*Enzyme-substrate dynamics.*—Protein and mRNA dynamics are described using Michaelis-Menten (MM) processes with RNA polymerase  $R_p$ , ribosome  $R$ , and two degradation enzymes, RNase  $X_m$  and protease  $X_p$ . Figure 1(a) illustrates the information flow from DNA

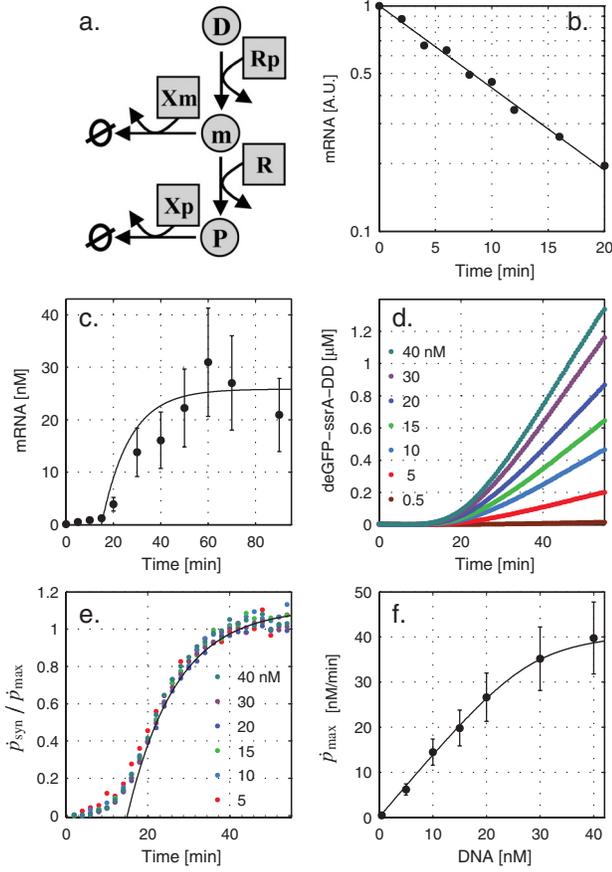


FIG. 1 (color). (a) Information flow from DNA ( $D$ ) to mRNA ( $m$ ) to protein ( $p$ ) is carried by synthesis enzymes RNA polymerase  $R_p$  and ribosome  $R$ . Degradation by RNase  $X_m$  and protease  $X_p$ . (b) Exponential degradation of  $\sim 200$  nM mRNA with lifetime  $\tau_m = 12$  min. (c) mRNA dynamics in the presence of 30 nM DNA exhibiting a delay and exponential accumulation with a time scale  $\tau_m$ . (d) Kinetics of protein synthesis at various DNA concentrations. (e) Normalized protein synthesis rate  $\dot{p}_{\text{syn}}/\dot{p}_{\text{max}}$  fitted to an exponential rise with a time scale  $\tau_m$ . (f) Protein maximal accumulation rate as a function of DNA fitted with a MM curve using  $R_p = 29$  nM and  $K_{\text{TX}} = 1.5$  nM.

( $D$ ) to mRNA ( $m$ ) to protein ( $p$ ). Each enzymatic reaction occurs in two steps: (i) Binding of enzyme  $E$  and substrate  $S$  to form a complex  $[ES] = S^f(K_M + S^f)^{-1}E$  at equilibrium. The MM constant is  $K_M$  and free substrate is  $S^f = S - [ES]$  [2]. (ii) Catalysis of product  $P$  or substrate degradation given by  $\dot{P} = k_{\text{cat}}[ES]$  with catalysis rate  $k_{\text{cat}}$ . For  $S \ll E + K_M$  the product formation rate scales linearly with the substrate (first-order dynamics)  $\dot{P} = k_{\text{cat}}(1 + K_M/E)^{-1}S$ ; for  $S \gg E + K_M$  product formation rate is independent of substrate (zeroth-order dynamics)  $\dot{P} = k_{\text{cat}}E$ . We denote the MM constant  $K_{\text{TX}}$  ( $K_{\text{TL}}$ ) for TX (TL) and  $K_m(K_p)$  for mRNA (protein) degradation. Respectively,  $k_{\text{TX}}$ ,  $k_{\text{TL}}$ , and  $k_{\text{deg}}$  are the catalysis rates for TX, TL, and protein degradation.

**mRNA degradation and synthesis.**—Degradation of mRNA, primarily regulated by the endonuclease RNase

$E$  [20], was monitored by adding  $\sim 200$  nanomolar (nM) 960 base-long transcript (not coding for GFP) that was transcribed in advance and labeled with radioactive rNTPs. The dynamics showed exponential decay with lifetime  $\tau_m = 12 \pm 2$  min [Fig. 1(b)], suggesting  $K_m + X_m \gg 200$  nM. Next, we followed mRNA TX and degradation by expressing the nondegradable protein deGFP-ssrA-DD (780 base transcript), while monitoring the incorporation of labeled rNTPs. Under TX/degradation mRNA appeared after a delay of  $\tau_0 = 15 \pm 5$  min followed by exponential accumulation to steady state  $m_{\text{ss}}$ ,  $\dot{m} = m_{\text{ss}}(1 - \exp[-(t - \tau_0)/\tau_m])$  [Fig. 1(c)]. We attribute the delay to the time required for TX of a single mRNA [21], from which we deduce a TX rate  $k_{\text{TX}} = 1 \pm 0.5$  rNTP  $s^{-1}$  that is smaller by 1–2 orders of magnitude than *in vivo* [18]. For  $D = 10$ –60 nM the mRNA steady state varied between 10 and 30 nM, which is consistent with the TX rate and mRNA lifetime. Therefore, beyond the initial delay, the mRNA dynamics can be described by a constant TX term and first-order degradation,

$$\dot{m}(t) = k_{\text{TX}}N_m^{-1}[R_p D] - \frac{m(t)}{\tau_m}, \quad \forall t > \tau_0, \quad (1)$$

where  $N_m$  is the mRNA length in rNTPs and  $[R_p D]$  is left to be determined. The steady-state solution is given by  $m_{\text{ss}}(D) = \tau_m k_{\text{TX}} N_m^{-1} [R_p D]$ , where  $\tau_m k_{\text{TX}} N_m^{-1} \approx 1$ .

**Protein synthesis.**—We next studied the synthesis dynamics (TX/TL) of deGFP-ssrA-DD ( $p_{\text{syn}}$ ), varying its plasmid concentration [Fig. 1(d)]. The protein accumulated to  $\sim 1$   $\mu\text{M}$  within an hour, after which the synthesis rate  $\dot{p}_{\text{syn}}$  decayed exponentially over 9 h before completely stopping with  $\sim 10$   $\mu\text{M}$  protein expressed. The slow decay in synthesis is a result of reagents depletion and of lack of waste removal, and we, therefore, focus on the first hour of reaction. We compute the normalized protein synthesis rate as the ratio of  $\dot{p}_{\text{syn}}(D, t)$  to its maximal value  $\dot{p}_{\text{max}}(D)$  [Fig. 1(e)]. The normalized protein dynamics was independent of  $D$ , exhibiting a delay of  $\tau_0 + \tau_f \leq 20$  min (first 3–5 min not measured) followed by an exponential rise  $\dot{p}_{\text{syn}}/\dot{p}_{\text{max}} = 1 - \exp[-(t - \tau_0 - \tau_f)/\tau_m]$ . Aside from an additional delay  $\tau_f$ , protein synthesis rate scales linearly with mRNA dynamics [Fig. 1(c)] for all measured  $D$ , suggesting  $m \ll K_{\text{TL}} + R$  and first-order dynamics:

$$\dot{p}_{\text{syn}}(D, t) = k_{\text{TL}}N_p^{-1}m(D, t - \tau_f)/(1 + K_{\text{TL}}/R). \quad (2)$$

Here  $N_p$  is the protein length (in amino acids). Thus,  $\dot{p}_{\text{max}}(D) = k_{\text{TL}}N_p^{-1}m_{\text{ss}}(D)/(1 + K_{\text{TL}}/R)$ . Assuming TX and TL are concurrent [22], we attribute the additional delay  $\tau_f \leq 5$  min to the protein folding time, which is similar to previous measurements [17,23]. For  $D = 30$  nM the mRNA steady state is  $m_{\text{ss}} = 25 \pm 5$  nM [Fig. 1(c)], and the maximal protein synthesis rate is  $\dot{p}_{\text{max}} = 35 \pm 5$  nM  $\text{min}^{-1}$  [Fig. 1(f)]. Thus,  $k_{\text{TL}}N_p^{-1}(1 + K_{\text{TL}}/R)^{-1} = 1$ –2  $\text{min}^{-1}$ . As the ratio  $K_{\text{TL}}/R$  is

undetermined, this sets a lower bound on the translation rate  $k_{TL} > 4$  amino acids  $s^{-1}$ , which is 1 order of magnitude smaller than the *in vivo* values [18] and slightly higher than previous *in vitro* work [24]. We find that the protein accumulation rate saturates at  $D \approx 30$  nM [Fig. 1(f)]. As TL is linear in  $m$ , the saturation is a result of the TX dynamics and suggests  $R_p + K_{TX} \approx 30$  nM. We fit the data using the expression for  $\dot{p}_{max}(D)$  [Fig. 1(f)]. We find  $K_{TX} = 1\text{--}10$  nM, which is similar to biochemical data [25], and  $R_p = 30 \pm 5$  nM which is consistent with the extract dilution of *E. coli* cytoplasm where  $R_p = 1\text{--}10$   $\mu\text{M}$  [18].

**Protein degradation.**—To study the dynamics of protein degradation ( $p_{deg}$ ) we added to the extract varying initial concentrations  $P_0$  of purified His-eGFP-ssrA, and measured the decrease in fluorescence [Fig. 2(a)]. The dynamics did not obey simple exponential decay. The initial rate ( $t < 5$  min) as a function of  $P_0$  was fitted to the MM equation  $\dot{p}_{deg} = -k_{deg}[pX_p]$  [Fig. 2(b)]. The degradation rate saturates around  $P_0 = 100$  nM suggesting  $K_p + X_p \approx 100$  nM with a maximal degradation rate  $k_{deg}X_p = 12 \pm 2$  nM  $min^{-1}$ . However, for  $t > 5$  min, degradation accelerated to a constant rate of  $k_{deg}X_p = 15 \pm 2$  nM  $s^{-1}$  [Fig. 2(c)]. The zeroth-order degradation was observed down to 1–10 nM protein, our detection limit, indicating that  $X_p, K_p < 10$  nM. Similar zeroth-order dynamics were detected in single-cell measurements [26] and in purified form in the presence of SspB (an endogenous specificity-enhancing protein) where  $K_p \sim 100$  nM [27]. This is qualitatively different from the first-order dynamics observed in cell population [26] and in purified form without SspB where  $K_p \sim 1$   $\mu\text{M}$  was measured [28]. We note that at most  $\sim 0.5$   $\mu\text{M}$  protein could be degraded by the cell-free system, beyond which the rate slowed down considerably to a complete stop, possibly due to lack of feeding reservoir.

**Protein synthesis and degradation.**—Theoretically, when both synthesis and degradation reactions take place, a steady-state solution to the equation  $\dot{p} = \dot{p}_{syn} + \dot{p}_{deg}$  is given by  $p = K_p \eta (1 - \eta)^{-1} + X_p \eta$ , where  $\eta(D) = \dot{p}_{max}(D)/(k_{deg}X_p)$ . The solution is possible only if synthesis is slower than the maximal degradation rate  $\eta < 1$ , and scales with  $K_p + X_p$  for small  $\eta$ . For  $\eta \rightarrow 1$  a transition occurs to a regime with zeroth-order degradation and protein accumulation. Based on the two separate measurements of protein synthesis [Figs. 1(d)–1(f)] and degradation [Figs. 2(a)–2(c)] we deduced that combined reactions would lead to protein accumulation for synthesis rate  $\dot{p}_{max} > k_{deg}X_p = 15$  nM  $min^{-1}$  ( $\eta > 1$ ). Conversely, steady state would be possible for  $\dot{p}_{max} < 15$  nM  $min^{-1}$  ( $\eta < 1$ ), but its value bound by  $X_p, K_p < 10$  nM (not detectable). In addition, for  $\eta > 1$  we expected a transition from undetectable protein levels to accumulation at a lag time  $T$  as the protein accumulation rate  $\dot{p}(D, T)$  becomes positive. We obtain  $T$  by solving  $\dot{p}_{syn}(D, T) = k_{deg}X_p$ :

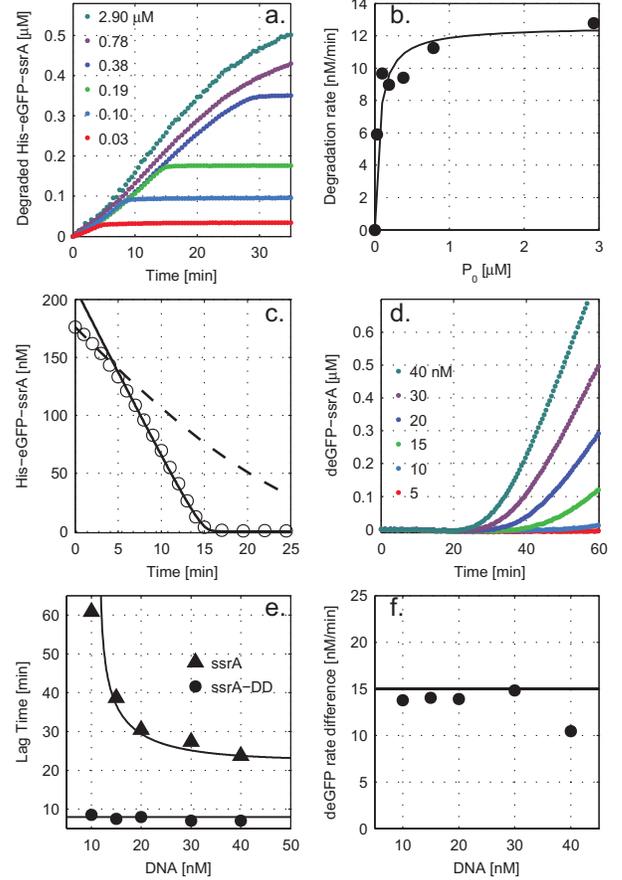


FIG. 2 (color). (a) His-eGFP-ssrA degradation for  $0.03 < P_0 < 2.9$   $\mu\text{M}$ . (b) Initial degradation rate as a function of  $P_0$  fitted to MM kinetics using  $K_p = 60$  nM,  $X_p = 1$  nM, and  $k_{deg}X_p = 12.5$  nM  $min^{-1}$ . (c) Protein degradation for  $P_0 = 190$  nM (open circles) fitted with MM kinetics using  $K_p = 60$  nM,  $X_p = 1$  nM,  $k_{deg}X_p = 10$  nM  $min^{-1}$  (dashed line) and  $K_p = 5$  nM,  $X_p = 1$  nM,  $k_{deg}X_p = 15$  nM  $min^{-1}$  (solid line). (d) deGFP-ssrA accumulation at various DNA concentrations. (e) Lag time  $T$  as a function of DNA for deGFP-ssrA fitted using Eq. (3) and deGFP-ssrA-DD exhibiting a constant delay. (f) Accumulation rate difference between deGFP-ssrA and deGFP-ssrA-DD.

$$T(D) = \tau_0 + \tau_f - \tau_m \log[1 - \eta^{-1}(D)]. \quad (3)$$

Indeed, we followed the dynamics of deGFP-ssrA expression, observing no steady state and a lag time before accumulation at constant rate [Fig. 2(d)]. Protein accumulation was not detected below  $D \leq 5$  nM, for which  $\dot{p}_{max} < 15$  nM  $min^{-1}$  [Fig. 1(f)]. The lag time decreased with  $D$  and fitted well to Eq. (3) [Fig. 2(e)] using the previously derived parameters (Table I). Satisfyingly, the accumulation rate of deGFP-ssrA-DD exceeded that of deGFP-ssrA by a constant value  $15$  nM  $min^{-1}$  [Fig. 2(f)], which is exactly the maximal degradation rate, consistent with active degradation during the transition to constant accumulation.

TABLE I. Biosynthesis and degradation parameters

	This work	Previous	Ref.
$k_{TX}$ (rNTP s <sup>-1</sup> )	1.0 ± 0.5	40–55	[18]
$k_{TL}$ (amino acids s <sup>-1</sup> )	>4	11–21	[18]
$\tau_m$ (min)	12 ± 1.5	1–18	[29]
$X_p k_{deg}$ (nM min <sup>-1</sup> )	15 ± 2	$X_p$ , 1	[27]
$R$ (nM)	>30	10 <sup>3</sup> –10 <sup>5</sup>	[18]
$R_p$ (nM)	30 ± 5	10 <sup>3</sup> –10 <sup>4</sup>	[18]
$X_p$ (nM)	<10	...	
$K_{TX}$ (nM)	1–10	1–10 <sup>2</sup>	[25]
$K_p$ (nM)	<10	10 <sup>2</sup> –10 <sup>3</sup>	[27,28]
$K_m + X_m$ (nM)	>10 <sup>2</sup>	>10 <sup>3</sup>	[30]

**Summary.**—We measured mRNA and protein dynamics in an endogenous *E. coli* extract and derived a consistent coarse-grained model. Additional couplings such as TX/TL [21], TL/mRNA degradation [31], and polysome effects [24] were not considered here and are, therefore, implicit in the parameters. We found that protein TX/TL/degradation dynamics did not achieve detectable steady state, and that degradation occurred at constant rate rather than exponentially in time. Both results stem from the highly efficient function of ClpXP/ssrA system, at a scale of  $K_p \sim 10$  nM. Thus, our data derived from *E. coli* cell extract are consistent with biochemical and single-cell data [26,27]. In *E. coli* the ssrA tag is not coded for in the gene as other degradation peptides, but rather is added post-translationally to defective proteins [32]. Possibly,  $K_p$  is set to low values, 1–10 nM, to allow for complete elimination of defective proteins rather than to maintain steady-state levels. The observed transition to protein accumulation is generic for enzymatic degradation, as steady-state solutions only exist in part of the parameter space. Nonetheless, first-order degradation in the extract should be attainable using tags other than ssrA. The control of degradation kinetics is interesting in systems with spatially localized sources where nonlinearity leads to power law rather than exponential gradients [3]. Thus, the current cell-free system together with a surface platform [12,13] offers a suitable framework to study spatial phenomena.

We wish to thank S. Daube and E. B. Jaoui for assistance and discussions. This work was funded by the U.S.-Israel Binational Science Foundation (V.N., R.H.B.-Z.), the Israel Science Foundation (R.H.B.-Z.), the Minerva Foundation (R.H.B.-Z.), and National Science Foundation Grant No. PHY-0750133 (V.N.).

- [1] J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry. International Edition* (W. H. Freeman and Company, New York, 2002).  
 [2] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, 2007).

- [3] A. Eldar, D. Rosin, B. Z. Shilo, and N. Barkai, *Dev. Cell* **5**, 635 (2003).  
 [4] T. J. Perkins and P. S. Swain, *Mol. Syst. Biol.* **5**, 326 (2009).  
 [5] M. B. Elowitz and S. Leibler, *Nature (London)* **403**, 335 (2000).  
 [6] A. D. Tadmor and T. Tlusty, *PLoS Comput. Biol.* **4**, e1000038 (2008).  
 [7] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, and T. Ueda, *Nat. Biotechnol.* **19**, 751 (2001).  
 [8] V. Noireaux, R. Bar-Ziv, and A. Libchaber, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12 672 (2003).  
 [9] M. Isalan, C. Lemerle, and L. Serrano, *PLoS Biol.* **3**, e64 (2005).  
 [10] J. Kim, K. S. White, and E. Winfree, *Mol. Syst. Biol.* **2**, 68 (2006).  
 [11] V. Noireaux and A. Libchaber, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17 669 (2004).  
 [12] A. Buxboim, M. Bar-Dagan, V. Frydman, D. Zbaida, M. Morpurgo, and R. Bar-Ziv, *Small* **3**, 500 (2007).  
 [13] A. Buxboim, S. S. Daube, and R. Bar-Ziv, *Mol. Syst. Biol.* **4**, 181 (2008).  
 [14] W. U. Dittmer and F. C. Simmel, *Nano Lett.* **4**, 689 (2004).  
 [15] W. U. Dittmer, S. Kempter, J. O. Rädler, and F. C. Simmel, *Small* **1**, 709 (2005).  
 [16] J. Shin and V. Noireaux, *J. Biol. Eng.* **4**, 8 (2010).  
 [17] J. Shin and V. Noireaux, *J. Biol. Eng.* **4**, 9 (2010).  
 [18] H. Bremer and P. P. Dennis, in *Escherichia Coli and Salmonella: Cellular and Molecular Biology*, edited by F. C. Neidhardt *et al.* (ASM Press, Washington, DC, 1996), pp. 1553–1569, 2nd ed.  
 [19] See supplemental material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.106.048104> for details of the experimental system and procedures.  
 [20] C. Jain and J. G. Belasco, *Genes Dev.* **9**, 84 (1995).  
 [21] S. Proshkin, A. R. Rahmouni, A. Mironov, and E. Nudler, *Science* **328**, 504 (2010).  
 [22] O. L. Miller, Jr., B. A. Hamkalo, and C. A. Thomas, Jr., *Science* **169**, 392 (1970).  
 [23] J. A. Megerle, G. Fritz, U. Gerland, K. Jung, and J. O. Rädler, *Biophys. J.* **95**, 2103 (2008).  
 [24] K. A. Underwood, J. R. Swartz, and J. D. Puglisi, *Biotechnol. Bioeng.* **91**, 425 (2005).  
 [25] E. M. Owens and G. N. Gussin, *Gene* **23**, 157 (1983).  
 [26] W. W. Wong, T. Tsai, and J. C. Liao, *Mol. Syst. Biol.* **3**, 130 (2007).  
 [27] G. L. Hersch, T. A. Baker, and R. T. Sauer, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12 136 (2004).  
 [28] Y. I. Kim, R. E. Burton, B. M. Burton, R. T. Sauer, and T. A. Baker, *Mol. Cell* **5**, 639 (2000).  
 [29] J. A. Bernstein, A. B. Khodursky, P. H. Lin, S. Lin-Chao, and S. N. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9697 (2002).  
 [30] Y. Redko, M. R. Tock, C. J. Adams, V. R. Kabardin, J. A. Grasby, and K. J. McDowall, *J. Biol. Chem.* **278**, 44001 (2003).  
 [31] A. Deana and J. G. Belasco, *Genes Dev.* **19**, 2526 (2005).  
 [32] K. C. Keiler, P. R. H. Waller, and R. T. Sauer, *Science* **271**, 990 (1996).