Single-Molecule Study Reveals a Complex *E. coli* RNA Polymerase

Martin Guthold*[a] and Dorothy A. Erie[b]

KEYWORDS:
nucleic acids · RNA polymerase · single molecules · transcription · transferases

Introduction

RNA polymerase (RNAP) is the enzyme responsible for catalyzing the DNA-directed synthesis of RNA chains and is thus a key player in the expression of the genetic information encoded in DNA. In the past, a great deal has been learned about the dynamics and kinetics of this complicated process, often by analyzing the RNA transcripts using gel electrophoresis. It is known that RNAP is absolutely processive and can transcribe thousands of base pairs (bp) in a single binding event. Yet, along the way, it often pauses,[1] it can "drop" into conformational states that are off the main elongation pathway,[2] and it may move along the DNA in a nonmonotonous manner.[3] At times, studying transcription seems like the Herculean fight against the many heads of Hydra—one question is answered only to be replaced by two new ones.

New techniques can provide a different angle on the investigation of transcription and may provide novel and complementary information to previous experiments and possibly address questions that could not be asked before. One such approach is the study of individual molecules (as opposed to many molecules in a bulk), which has progressed to the point where the dynamics of single enzymes can be investigated. A single-molecule study of the transcription process by Davenport et al. in Carlos Bustamante's lab[4] was recently published in *Science*. This study will be reviewed here and will be used to reflect on the progress of studying the dynamics of individual RNAP molecules.

Single-molecule experiments were on the fringe just a few years ago and many researches doubted whether they could provide useful data that could be incorporated into and compared with the existing body of knowledge. Proponents argued that single-molecule experiments provide additional information that may be averaged out in or inaccessible to bulk experiments. Doubters contended, however, that for this new information to be meaningful, it is crucial that it be compared with bulk experiments. In other words, averaging of data obtained on single molecules must match the results of bulk experiments. Davenport and coworkers[4] accomplish both; they obtain novel and exciting insights into transcription by single-molecule experiments, and they compare the single-molecule results to data obtained from traditional bulk experiments.

Brief history of single-molecule studies on RNAP

Before going more into the details of the study by Davenport et al.,[4] it is appropriate to give a brief history of single-molecule studies of the transcription process. There are four important previous studies that led to Davenport's report. Studying the dynamics of individual RNAP molecules started in the early nineties when Schafer et al. in J. Gelles' lab followed the shortening of the downstream arm of a transcribed piece of DNA by optical microscopy.[5] In these studies, the RNAP of a stalled RNAP – DNA elongation complex was anchored to a glass slide, while the DNA, which had a 40-nm Au colloid attached to the downstream end, was free to float in solution. As nucleoside triphosphates were added, the Brownian motion of the bead, as observed with video microscopy, decreased, indicating that RNAP was pulling the downstream end of the DNA (for more technical details see also ref. [6]). This study was a great technical accomplishment but it did not yet provide much insight into the mechanism of transcription. The next breakthrough followed when Yin et al. measured the force required to stall RNAP by using an optical trap.[7] Interestingly, they found that RNAP exerted a greater force during RNA synthesis than classic motor proteins such as kinesin and myosin. In the third study, Wang et al.[8] measured force vs. distance curves by using an optical trap that had a feedback mechanism to keep the position of the bead in the trap at a predefined displacement (position clamp). They found that forces less than the stall force of 14 pN did not significantly affect the rate of transcription. This result indicates that enzyme translocation along the DNA is not the rate-limiting step to RNA synthesis.[8] Significantly, no other technique, at present, could be used to address this question. These results showed that single-molecule studies could, in fact, provide new insights into the mechanism of transcription. In the fourth study,[9] Yin et al. investigated the mechanism of termination by using the tethered-bead technique of the first study.[5] To assure that their data were consistent with bulk solution experiments, they also compared the termination efficiency found in the single-molecule studies with those determined from bulk solution experiments. The single-molecule termination experi-

[a] *Dr. M. Guthold*
Department of Computer Science, Department of Physics & Astronomy, and Department of Chemistry
University of North Carolina
Chapel Hill, NC 27599 (USA)
Fax: (+1) 919-962-0480
E-mail: guthold@cs.unc.edu

[b] *Prof. Dr. D. A. Erie*
Department of Chemistry
University of North Carolina
Chapel Hill, NC 27599 (USA)

ments revealed that those RNAP molecules that terminated also paused at the terminator, while those that did not terminate did not pause. These data indicate that intrinsic terminators function by a nonequilibrium process in which terminator efficiency is determined by the relative rates of nucleotide addition and pause state entry (see also Figure 3) and that termination is not in direct competition with transcript elongation.[9]

When reviewing the brief history of single-molecule studies on RNAP, the contributions by two Japanese groups, though mainly concerned with the general interactions of RNAP and DNA, should not go unmentioned. In these studies, fluorescence microscopy was used to observe the interactions of fluorescently labeled RNAP with DNA. Kabata et al.[10] employed superintensified fluorescence microscopy to visualize the movement of RNAP over combs of immobilized bacteriophage lambda DNA (λ-DNA). A fraction of the RNAP molecules was seen to deviate from the direction of bulk flow and to move along the extended DNA molecules. This observation suggests that RNAP can slide along nonspecific stretches of DNA. Harada et al.[11] used internal reflection fluorescence microscopy to observe the dissociation and association events of RNAP with different regions of a single λ-DNA molecule, which was suspended in laser tweezers. For AT-rich regions, fast and slow dissociation constants of 3.0 s^{-1} and 0.66 s^{-1} were determined, respectively; for GC-rich regions, a fast dissociation rate of 8.4 s^{-1} was measured.

The set-up of Davenport et al.[4] improved on the technology of the previous studies. In the first and fourth study,[5, 9] the spatial resolution was low because of the unconstrained Brownian motion of the bead. In the other two studies, which employed optical traps,[7, 8] there was significant laser damage to the RNAP because a powerful laser is required to measure the forces exerted by RNAP. To circumvent these problems, Bustamante's laboratory developed the experimental set-up shown in Figure 1. Using this set-up, laser damage is limited because the laser trap is only used for assembling the stalled complex between two polystyrene beads. During the experiment, the laser is switched off, and one bead is held in a static position by suction on a micropipette and the other bead, which is attached to the downstream end of the DNA, floats freely in solution. To improve the resolution, flow was used to minimize Brownian motion. In addition, the flow can be used to exert a controlled force on the DNA. This set-up permitted the researches to follow transcription for up to 3000 bp with a relative position error of about 21 bp and an absolute position error of about 60 bp.

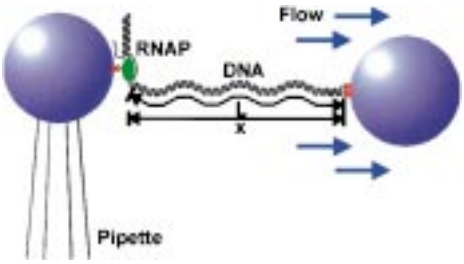

Figure 1. *Experimental set-up used in the study by Davenport et al.[4] Stalled elongation complexes were trapped between two polystyrene beads with biotin–streptavidin linkers. Assembly was aided by holding one bead on a micropipette by applying suction and the other bead with a laser trap. Subsequently, the laser trap was switched off to avoid laser-induced damage and a controlled force was exerted on the polymerase by applying flow. Flow also limited the Brownian motion of the bead, making the measurement more accurate. x = end-to-end distance of the DNA molecule; L = contour length of the DNA molecule. (Reprinted with permission from ref. [4]. Copyright © American Association for the Advancement of Science, 2000.)*

Discussion of the major findings

Transcription is highly irregular

The most interesting discoveries in the studies of Davenport and co-workers concern the dynamics of transcription. In particular, it was found that individual RNAP molecules move along the DNA template in a highly irregular, nonmonotonous manner (Figure 2). Not only does RNAP often pause in the middle of transcription, but it also moves along the DNA at highly variable rates between pauses. As can be seen in the insets of Figure 2, RNAP molecules appear to catalyze RNA synthesis at a given rate for an extended stretch of DNA and then seem to switch rates randomly and slowly. However, it has to be noted that the data were low-pass filtered ($f = 0.067$ Hz; $t = 15$ s), which tends to smooth out edges and makes

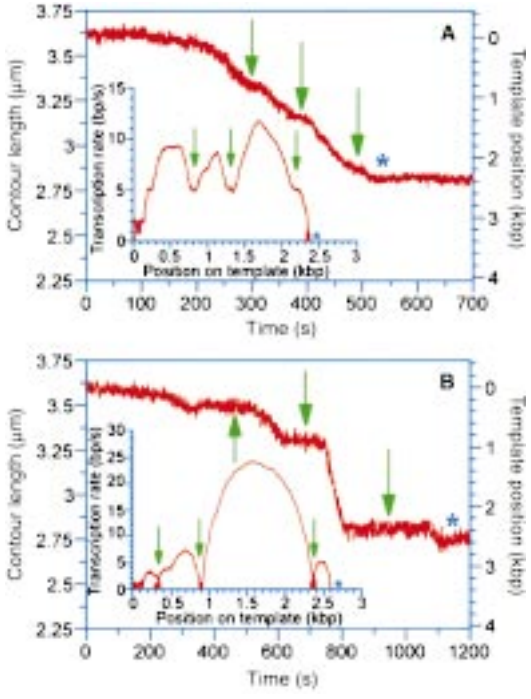

Figure 2. *Transcription by a single RNAP molecule. Plot of the contour length as a function of time and of the transcription rate as a function of the template position (insets). Transcription was carried out against a force of 8 pN (A) and 2 pN (B), respectively. Several pauses and drastic changes in rate can be clearly seen. Note that the distance data were low-pass-filtered (0.067 Hz) before a time derivative was taken to obtain the transcription rate. This implies that pauses shorter than the filtering duration (15 s) are not detected. (Reprinted with permission from ref. [4]. Copyright © American Association for the Advancement of Science, 2000.)*

transitions less rapid. Thus, the apparent slow rate of interconversion may be due to the relatively long filter time (15 s) used in the data analysis and not a property of RNAP.

Surprisingly, the rates did not appear to depend on the DNA sequence, although to fully confirm this statement more molecules at a resolution higher than 21 bp should be analyzed in the future. Such heterogeneity was never seen before in single-molecule studies of RNA polymerase. Furthermore, plotting the *averaged* peak rates of single molecules yields a bimodal distribution with two distinct peaks at 5.5 and 9.1 bp s^{-1} (see Figure 3 A of ref. [4]). A similar bimodal distribution was also obtained by plotting all peak rates without averaging over molecules.[22] These results suggest that RNAP may assume (at least) two distinct conformations during transcription: one that catalyzes synthesis slowly and the other rapidly. A similar suggestion was put forth previously to explain the bulk elongation kinetics of RNA polymerase III.[12]

Analysis of pauses

There were nine locations where the polymerase molecules were found to have a high propensity to pause. Consistent with bulk studies,[13–15] the efficiency of pausing was less than 100%. This finding supports the notion that pauses are off the main pathway, because if pauses were on the main pathway, pause efficiency would necessarily be 100%. In addition, it was observed that the polymerase often entered into an arrested state at these pause sites, indicating that pausing may be an intermediate state between an actively elongating state and an arrested state. This result is consistent with bulk solution studies that found that arrested states were at least two kinetic steps off the main synthesis pathway.[2]

Interestingly, there is a direct correlation between the number of times that an RNA polymerase molecule paused and the inverse of its average peak rate. There are two interpretations for this observation. As the authors suggest, this result could mean that a slower transcribing RNAP molecule is more likely to pause than a faster one. Alternatively, it could mean that a particular RNAP molecule is slower *because* there are more pauses, especially short pauses that may have been filtered out in the analysis. Consistent with this alternative interpretation, the highest peak rates are seen when the pauses are the furthest apart; that is, there appears to be a correlation between the peak synthesis rate and the distance between pauses (Figure 2). Thus, cause and effect of this correlation are not cleanly separated; that is, it is not entirely clear whether the rate appears slower because there are more pauses or that the slower rate causes more pausing. This question should be revisited with an instrument that has higher temporal and spatial resolution.

The effect of force on a transcribing polymerase

While the effect of force on the rate was tested before,[8] the effect of force on pausing was not. Forces up to 15 pN, which is the stall force, had neither an effect on the average transcription rate, as previously observed,[8] nor an effect on pausing. In contrast, it was found that force did have an effect on arrest. Applying a force between 9 – 15 pN increased the probability that RNAP would enter into an arrested state from a pause state. It, thus, appears that external force can increase the rate of conversion of a paused state into an arrested state; whereas, the rate of conversion of an elongating state to a paused state is not affected by force.

Summary of results and tentative model of transcription

By using an integrated optical-trap/flow-control video microscopy system to investigate individual RNAP molecules, Davenport et al. observed the following characteristics of transcription. 1) RNAP can switch randomly between (at least) two different rates, and these rates are *not* correlated to the template position. 2) The pause efficiency is less than 100%, suggesting that paused complexes are off the main pathway (see also refs. [13 – 15]). 3) The pause efficiency is not related to the pause half-life, further corroborating the fact that pauses are off the main pathway. 4) There is a direct correlation between the frequency of pausing and the inverse of the transcription rate. RNAP molecules transcribing at a faster rate seem to be less likely to pause than those transcribing at a slower rate, suggesting a kinetic competition between transcription and pausing. Note, however, the caveat about cause and effect pointed out above. 5) The more often an RNAP molecule pauses, the more likely it is to become arrested. 6) The longer an RNAP molecule remains in a pause, the more likely it is to become arrested. 7) Arrest occurs at pause sites, suggesting that arrest and pausing are related. 8) A force of up to 15 pN neither affects the rate of transcription nor increases the probability or duration of pausing. 9) A force between 9 and 15 pN increases the probability of arrest for molecules that were already paused.

The kinetic model suggested in Figure 3 is consistent with the above results. In this model, RNAP can proceed along the template in a fast or slow state. It is important to note, however, that these studies do not have sufficient resolution to be certain that synthesis occurs on the slow path, because if RNAP catalyzed synthesis only on the fast pathway but spent much of its time in any off-path state, it would appear to be catalyzing synthesis slowly. We suggest synthesis along the slow state as a formal possibility. In this model, RNAP molecules potentially can populate four states (the fast, slow, paused, or arrested state) at each

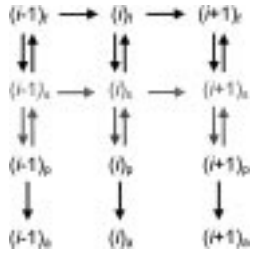

Figure 3. *Proposed kinetic mechanism for transcription. The different states are denoted by subscript letters: f = fast state, s = slow state, p = paused state, a = arrested state.* i − *1,* i, *and* i + *1 denote template positions. The arrows represent rates. The slow state is shown in gray because the evidence for the existence of this state is circumstantial. It is important to note that the rates towards pauses and arrests can greatly vary at different template positions. Moreover, our suggested model is consistent with the available data, but these data do not conclusively prove this mechanism.*

template position (see also refs. [2, 16]). There is strong evidence that the paused state is a kinetic intermediate between the elongation states and the arrested state, and that pausing competes kinetically with elongation (findings 2, 3, and 4 listed above). The model in Figure 3 is also consistent with the single-molecule study on termination[9] that showed that a paused state is an intermediate between an actively elongating complex and termination (findings 5, 6, and 7). For this latter case, the arrested state would be replaced with termination. Finally, it is important to note that the model proposed in Figure 3 is consistent with the data from Davenport et al.[4] and Yin et al.,[9] but the data do not prove the model. For instance, it cannot be ruled out that the fast molecules go directly to a paused state instead of going through a slow state first, giving rise to the question whether the slow state even exists.

Future questions

These studies have just scraped the surface of the possibilities of single-molecule studies on RNAP and other molecules. It is certain that optical microscopy and other single-molecule techniques such as atomic force microscopy (AFM)[17, 18] will continue to expand our understanding of transcription and other biological processes. The data discussed here have opened more new questions than answered old ones, which is to be expected for a young and active field. The most obvious questions that would benefit from single-molecule experiments are raised here. Which factors affect the rates of transcription elongation, pauses, and arrest? Is there a "selection process" in which slow complexes are more likely to pause and become arrested and thus are more exposed to transcriptional regulation and proofreading factors? What happens to arrested complexes in the presence of GreA and GreB, both of which are proteins that have been shown to increase the fidelity of transcription?[2] What are the effects of the RNA transcript length and the presence of hairpins on pausing[1, 14, 19] and termination?[20]

Future instruments may be able to achieve single-base-pair resolution (a factor of 20 better than the current resolution). With such instruments, questions about the single-nucleotide addition cycle, inchworming,[3] and backtracking[21] could be addressed—and most likely are currently being worked on in several laboratories. However, being able to obtain single-base resolution is insufficient. The technique needs to advance to the point where large populations of molecules can be examined with relative ease, so that the obtained results can be averaged and compared with those of bulk experiments. Finally, combining bulk solution studies with statistics on large populations of single molecules in conjunction with computer modeling studies, such as the one done by Matsuzaki et al.,[12] should provide great insights into the mechanism of transcription elongation.[23]

Support for this work came from NIH grants NCRR 5-P41-RR02170 and GM54136. The authors would like to thank Jeff Gelles and Peter Geiduschek for helpful discussions.

[1] L. C. Chan, R. Landick, *New perspectives on RNA chain elongation and termination by E. coli RNA polymerase*, Raven, New York, **1994**.
[2] D. A. Erie, O. Hajiseyedjavadi, M. C. Young, P. H. von Hippel, *Science* **1993**, *262*, 867.
[3] B. Krummel, M. J. Chamberlin, *J. Mol. Biol.* **1992**, *225*, 239.
[4] R. J. Davenport, G. J. L. Wuite, R. Landick, C. Bustamante, *Science* **2000**, *287*, 2497.
[5] D. A. Schafer, J. Gelles, M. P. Sheetz, R. Landick, *Nature* **1991**, *352*, 444.
[6] H. Yin, R. Landick, J. Gelles, *Biophys. J.* **1994**, *67*, 2468.
[7] H. Yin, M. D. Wang, K. Svoboda, R. Landick, S. M. Block, J. Gelles, *Science* **1995**, *270*, 1653.
[8] M. D. Wang, M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, S. M. Block, *Science* **1998**, *282*, 902.
[9] H. Yin, I. Artsimovitch, R. Landick, J. Gelles, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13 124.
[10] H. Kabata, O. Kurosava, I. Arai, M. Washizu, S. A. Marganson, R. E. Glass, N. Shimamoto, *Science* **1993**, *262*, 1561.
[11] Y. Harada, T. Funatsu, K. Murakami, Y. Nonoyama, A. Ishihama, T. Yanagida, *Biophys. J.* **1999**, *76*, 709.
[12] H. Matsuzaki, G. A. Kassavetis, E. P. Geiduschek, *J. Mol. Biol.* **1994**, *235*, 1173.
[13] G. A. Kassavetis, M. J. Chamberlin, *J. Biol. Chem.* **1981**, *256*, 2777.
[14] J. R. Levin, B. Krummel, M. J. Chamberlin, *J. Mol. Biol.* **1987**, *196*, 85.
[15] R. Landick, *Science* **1999**, *284*, 598.
[16] P. H. von Hippel, *Science* **1998**, *281*, 660.
[17] M. Guthold, X. Zhu, C. Rivetti, G. Yang, N. H. Thomson, S. Kasas, H. G. Hansma, B. Smith, P. K. Hansma, C. Bustamante, *Biophys. J.* **1999**, *77*, 2284.
[18] S. Kasas, N. H. Thomson, B. L. Smith, H. G. Hansma, X. Zhu, M. Guthold, C. Bustamante, E. T. Kool, M. Kashlev, P. K. Hansma, *Biochemistry* **1997**, *36*, 461.
[19] I. Artsimovitch, R. Landick, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7090.
[20] K. S. Wilson, P. H. von Hippel, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8793.
[21] N. Komissarova, M. Kashlev, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1755.
[22] G. J. L. Wuite, personal communication.
[23] Note added in proof: DNA rotation during transcription by individual RNAP molecules has just been observed directly, see Y. Harada, O. Ohara, A. Tahatsuki, H. Itoh, N. Shimamoto, K. Kinosita, Jr., *Science* **2001**, *409*, 113.