

Transcriptional Pausing Caught in the Act

Peter H. von Hippel^{1,*}

¹Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, OR 97403, USA

*Contact: petevh@molbio.uoregon.edu

DOI 10.1016/j.cell.2006.06.006

Single-molecule techniques now permit the tracking of a transcription complex along a DNA template in real time and to 1 bp resolution. As reported in this issue of *Cell*, Herbert et al. (2006) exploit this approach to study a central component of transcription regulation, the sequence-dependent pausing of RNA polymerase during transcript elongation.

The elongation of an mRNA transcript by RNA polymerase (RNAP) during transcription is not a straightforward process. Misincorporation of nucleotides, editing, backtracking, and termination reactions can, in principle, all compete with the continuation of elongation at any template position (Greive and von Hippel, 2005). A pause likely comprises the first “offline” state for any of these multistep competing reaction pathways. Whether a given complex enters one of these alternative reaction pathways depends on the relative heights of the transition-state barriers that control access to this first state at that template position (Figure 1). How far down the pathway it proceeds will depend on the heights of the barriers that control access to the subsequent states. Thus, understanding pausing and how it is regulated is central to elucidating how the transcription complex partitions itself between the reaction pathways available at each template position. In this issue of *Cell*, Herbert et al. (2006) use a single-molecule approach to characterize polymerase pausing during transcript elongation at 1 bp resolution. Theoretical (predictive) studies of pausing dynamics as a function of template sequence are also currently under development (for example, see Tadigotla et al., 2006).

Two general classes of sequence-dependent pauses have been defined experimentally (Artsimovitch and Landick, 2000). Class 1 pauses appear to be stabilized by the formation of a “pausing hairpin” in the transcribed RNA. Class

2 pauses depend on the presence of a “weak” (thermodynamically less stable) RNA-DNA hybrid and are stabilized by backtracking of the RNAP from such template positions. The combination of a pausing hairpin immediately upstream of a weak RNA-DNA hybrid containing runs of rU and dA residues can result in both pausing and destabilization of the elongation complex, thereby making transcription termination possible (Figure 1). It is likely that some sort of pausing represents the first step in the progression of elongation complexes down all alternative reaction pathways.

Herbert et al. (2006) use repetitive DNA templates, which permit

the monitoring of individual RNAP molecules transcribing up to eight template repeats. The authors then compare these results with the rates and pausing parameters obtained with different RNAPs. Examples of each class of pauses (*his*, class 1, and *ops*, class 2) have been built into the repeating templates. Pausing occurs at both of these sequences and can be clearly localized to 1 bp resolution. The authors also show that pausing is a stochastic event, as expected from the offline nature of the pausing process and the fact that the activation barriers to pausing and elongation must be comparable at these positions (Figure 1). In sin-

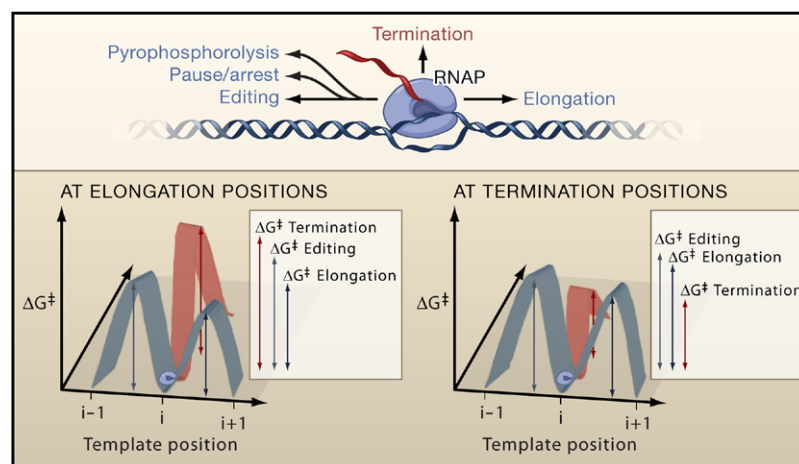


Figure 1. Transcription Complexes Do More than Elongate the Transcript

(Top) Alternative (multistep) reaction pathways that are potentially available to transcription complexes at any given template position. Pausing is the likely first step on any of these alternative pathways. (Bottom) Competing transition-state barriers control the probability of access of the transcription complex (blue sphere at position “i”) to the alternative (offline) pathways; the reaction with the lowest barrier is favored. Elongation is strongly favored over editing and termination at most template positions (bottom left), whereas termination is favored over elongation at termination positions (bottom right). Figure adapted from Greive and von Hippel (2005).

gle-molecule studies, this probabilistic behavior emerges as a yes/no process—that is, an individual transcription complex either goes into a paused state at the pause site or transcribes through it. If these events are measured over enough identical template positions, one can characterize (as Herbert et al. have done) the probability of pausing (the pausing efficiency) as well as the duration of the pauses that occur at any given site. Due to the repeating nature of the templates, one can also ask whether the propensity for an individual elongation complex to pause (or the duration of the pause when it does occur) changes as a function of the prior pausing history of the complex. No systematic change in either of these parameters was seen as individual transcription complexes moved through multiple template repeats. This result is consistent with biochemical studies involving templates with repeating terminators (see Pasman and von Hippel, 2002). Thus, the properties of an individual transcription complex depend on its template position and the presence of other regulatory components (such as proteins or RNA sequences), but not on the prior transcriptional and regulatory history of the complex.

In addition to the two classes of defined pauses, the Herbert et al. (2006) study also reveals what the authors and others (Neuman et al., 2003) have called “ubiquitous” or “elemental” pauses, which the present study shows also appear at defined sequence positions along the template. These elemental pauses appear to be comparable in duration to the *his* and *ops* pauses monitored here and share some features of template sequence. Herbert et al. (2006) show that these elemental pause sites do not involve backtracking; rather, they simply represent template positions at which elongation complexes stop long enough to be defined as a pause (here, ~ 1 s). In fact, perhaps because of a constant small downstream pulling force exerted

on the transcribing polymerase by the experimental setup, the authors also do not see backtracking at the other pauses under study, including the *ops* pause, where biochemical studies have suggested that significant backtracking does occur.

These elemental pauses may represent template positions at which the barrier to movement of the complex into the first of a sequence of offline states is sufficiently low—relative to the barrier to further elongation (Figure 1)—to permit a significant fraction of the complexes to partition transiently into such an initial offline state. The complex may then either return to the elongation-competent (online) state or proceed further down one of the alternative reaction pathways. Given that none of the pauses monitored in the present study appears to involve backtracking, the authors assume that, in these experiments, the 3' end of the transcript remains bound in the pretranslocated substrate binding subsite of the polymerase. (For a more detailed discussion of the geometry of the active site of the RNA polymerase, see Bar-Nahum et al., 2005; Greive and von Hippel, 2005; and references therein.)

The use of repeating templates has allowed Herbert et al. (2006) to tackle another important issue that previous single-molecule studies of transcription complexes have sought to address (e.g., see Tolic-Norrelykke et al., 2004). This is the apparent presence of “microheterogeneity” in preparations of transcription complexes as studied by high-resolution single-molecule techniques. What this means is that individual functional transcription complexes, selected at random within the same experimental field, appear to display somewhat different elongation and pausing characteristics. These differences are not large ($\pm 5\%$ – 10% change in velocity or pausing parameters); however, they fall well outside the limits of error of the measurements. One possible source of such microheterogeneity could be long-lived

(relative to the rate of transcription) “state switching” of an elongation complex within an individual run over a repeating template. This is not observed. Rather, the results of Herbert et al. (2006) indicate that only different complexes appear to demonstrate microheterogeneity, perhaps reflecting a stable heterogeneous distribution of properties within preparations of functioning polymerases. This microheterogeneity could reflect the consequences of minor chemical or conformational differences that are present within the protein preparation and are revealed by these single-molecule measurements but are not easily detected in biochemical studies. Alternatively, these perceived variations may reflect minor differences in how individual functional elongation complexes are “harnessed” to the beads used to tether the RNAP (and the DNA) in these single-molecule studies.

Are these differences real? If indeed such differences reflect a long-lived distribution of conformational states that may have been “trapped” in the process of folding individual polymerase subunits or in the course of assembling these subunits into functioning macromolecular complexes, then single-molecule approaches may enable examination of the distributions, and perhaps also the interconversion rates, of such long-lived conformational states. This has not been possible using more conventional solution techniques. On the other hand, if these differences represent artifacts of transcriptional complex manipulation that are unique to these single-molecule methods and do not reflect basic properties of the preparations of the macromolecular complexes, then we need to know this. Careful mass spectroscopy of preparations used in these studies should reveal whether the observed heterogeneity reflects chemical differences within individual subunits. Conformational heterogeneity at this level may be harder to establish by solution techniques.

This issue is important because concerns about microheterogeneity have a long history. In the 1940s and early 1950s, proteins (and nucleic acids) were considered to be colloids that were unlikely to display the atomic and molecular precision of “real” molecules. This view, of course, began to dissipate with the Watson-Crick structure of DNA, the demonstration that at least some small proteins could be refolded into equilibrium conformations, and the finding that macromolecules could be crystallized and discrete structures obtained by X-ray diffraction. These developments appeared, at the time, to relegate earlier concerns about microheterogeneity into the dustbin of scientific history. On the other hand, modern studies of protein and RNA folding and assembly suggest that there may

be multiple pathways of folding and that individual proteins—and certainly individual macromolecular assemblies—could be trapped into metastable states with long-term stability. Learning more about such distributions of metastable states and their rates of rearrangement is of great intrinsic interest because this may tell us more about folding and assembly processes *in vivo* and *in vitro* and also about how far these processes are driven toward homogeneity by chaperone complexes. In addition, such functional heterogeneity, if it exists, may provide further insight into regulatory events controlled by signal-transduction networks because these regulatory processes could work somewhat differently on “outliers” within conformational distributions of macromolecular assemblies.

REFERENCES

- Artsimovitch, I., and Landick, R. (2000). *Proc. Natl. Acad. Sci. USA* 97, 7090–7095.
- Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., and Nudler, E. (2005). *Cell* 120, 183–193.
- Greive, S.J., and von Hippel, P.H. (2005). *Nat. Rev. Mol. Cell Biol.* 6, 221–232.
- Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). *Cell*, this issue.
- Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J., and Block, S.M. (2003). *Cell* 115, 437–447.
- Pasman, Z., and von Hippel, P.H. (2002). *J. Mol. Biol.* 322, 505–519.
- Tadigotla, V.R., O’Maolaidigh, D., Sengupta, A.M., Epshtein, V., Ebright, R.H., Nudler, E., and Ruckenstein, A.E. (2006). *Proc. Natl. Acad. Sci. USA* 103, 4439–4444.
- Tolic-Norrelykke, S.F., Engh, A.M., Landick, R., and Gelles, J. (2004). *J. Biol. Chem.* 279, 3292–3299.

EGF Receptor Activation: Push Comes to Shove

Stevan R. Hubbard^{1,*}

¹Structural Biology Program, Skirball Institute of Biomolecular Medicine, and Department of Pharmacology, New York University School of Medicine, New York, NY 10016, USA

*Contact: hubbard@saturn.med.nyu.edu

DOI 10.1016/j.cell.2006.05.028

A study by Zhang et al. (2006) in this issue of *Cell* provides compelling evidence that the tyrosine kinase domain of the epidermal growth factor receptor (EGFR) is activated by the formation of an asymmetric dimer, with one kinase domain in the EGF-mediated dimer activating the other through an allosteric mechanism.

The receptor tyrosine kinase (RTK) family of cell-surface receptors includes the epidermal growth factor receptor (EGFR/ErbB1/HER1) and its relatives ErbB2/HER2, ErbB3/HER3, and ErbB4/HER4. Receptors of the EGFR subgroup are essential for embryonic development and adult tissue homeostasis in both vertebrates and invertebrates. Enhanced signaling from these

receptors, due to mutation or over-expression, contributes to several types of human cancer (reviewed in Yarden and Sliwkowski, 2001). Like their RTK brethren, members of the EGFR subgroup feature an extracellular region (ectodomain) that binds polypeptide ligands (EGF, transforming growth factor- α , neuregulins, and several others), a single-pass transmembrane helix,

and a cytoplasmic domain containing intrinsic tyrosine kinase activity (reviewed in Schlessinger, 2002).

RTK activation represents a signal transduction event in which an extracellular cue (such as a growth factor) is converted to a cellular response (for example, cell division) through a series of intracellular steps. The activation process for RTKs is simple yet complex. It