Published in final edited form as:

Methods. 2009 January ; 47(1): 1–5. doi:10.1016/j.ymeth.2008.12.001.

RNA Polymerase: A Nexus of Gene Regulation

John D. Helmann

Department of Microbiology, Cornell University, Ithaca, NY 14853

Abstract

In Bacteria, transcription is catalyzed by a single RNA polymerase (RNAP) whose promoter selectivity and activity is governed by a wide variety of transcription factors. The net effect of these transcriptional regulators is to determine which genes are transcribed, and at what levels, under any specific growth condition. RNAP thus serves as a nexus of gene regulation that integrates the information coming from a variety of sensory systems to appropriately modulate gene expression. The techniques presented in this volume provide a set of tools and approaches for investigating the factors controlling RNAP activity at both individual promoters and on a genomic scale. This introductory chapter provides a brief overview of RNAP and the transcription cycle and introduces general principles of how the fundamental steps of transcription are influenced by both DNA (promoter) sequences and trans-acting factors.

Keywords

RNA polymerase; transcription; promoter; activator; repressor; methods

Introduction

Regulation of gene expression is a fundamental property of living systems that enables individual cells to adapt to their environment and, in multicellular organisms, allows for the differentiation of multiple cell types from a single fertilized zygote. Most commonly, gene expression is regulated at the first step: the copying of the DNA template into an RNA transcript. The task of copying selected regions of DNA into their corresponding RNA transcripts falls to RNA polymerase (RNAP), which has been a focus of intense research in several model systems. Within the Bacteria, the RNAP from *Escherichia coli* is the best understood genetically and physiologically, but many important insights have also emerged from studies in other systems and from recent advances in the determination of three-dimensional structures of RNAP and its subunits from *E. coli* and several thermophilic organisms [1–3]. Ultimately, the activity of RNAP within the cell is governed by a large number of regulatory factors that control where and when transcription occurs. In this introductory chapter, I will provide an overview of the key features that define the process of transcription and highlight those steps that are commonly targeted by the regulatory machinery. The steps in transcription, and a selection of some well-characterized regulators that impinge on these steps, are summarized in Table 1.

Bacterial RNA polymerase as the engine of gene expression

RNAP is an exceptionally complex enzyme that can be thought of as the engine of gene expression. Energy, in the form of nucleoside triphosphates, fuels the synthesis of an RNA

polymer complementary to specific regions of the DNA template. Like all macromolecular synthesis, RNA synthesis can be divided into three general phases: initiation, elongation, and termination [4]. Importantly, each of these phases can be a target of regulation.

Bacterial RNAP is a multisubunit enzyme and consists of a core polymerase (abbreviated as E) containing the beta, beta', and two alpha subunits (together with one or more omega subunits; [5,6]) and a dissociable specificity factor known as sigma (σ) [7]. While the core RNAP (minimally, $\beta\beta'\alpha_2$) is competent for transcription elongation and termination, transcript initiation requires an associated σ subunit. The core+ σ complex is designated holoenzyme. Many bacteria contain multiple types of σ factor and, therefore, multiple holoenzyme species [8,9]. In *E. coli*, the primary σ is called σ^{70} and is required for the transcription of most genes including those for many essential functions. So-called alternative σ factors often control specialized sets of genes that are required for specific cellular functions (e.g. motility and chemotaxis) or under stress conditions (e.g. heat shock). While most σ factors are structurally related to σ^{70} , a second family of σ proteins known as σ^{54} (or σ^N) also mediates promoter recognition of specialized sets of genes [10].

Promoter recognition and transcript initiation

The first step in RNA synthesis, and the most common target for regulation, is initiation. The ability to identify start sites of transcription and initiate chains *de novo* distinguishes RNAP from the mechanistically related enzymes that synthesize DNA (which require a separate primase to generate a primer for elongation; [11]). Since initiation is a frequent target for regulation, and is itself an exceedingly complex reaction, it is worth reviewing the key stages in this process. These can be defined as (i) promoter recognition, (ii) transcript initiation, and (iii) promoter clearance.

The process of transcription begins with binding of RNAP to the promoter. The promoter can be defined as that region of DNA that interacts directly with RNAP during initiation of transcription. Promoter localization likely begins with an RNAP molecule associated non-specifically with DNA that can diffuse to explore the local sequence space prior to promoter recognition or, alternatively, release or transfer to a different DNA segment [12,13]. Once bound at a promoter site, RNAP typically contacts ~80 bp of DNA extending from as much as ~60 bp upstream of the start point of transcription (defined as +1) to ~20 bp downstream [Ross and Gourse, this volume]. Within this extended promoter recognition region reside the key contact points that allow RNAP to distinguish promoter from non-promoter DNA. The precise nature of these sequences varies depending on the promoter and the holoenzyme. For σ^{70} class promoters, the key recognition elements are conserved hexamers located near -35 and -10 (relative to the transcription start point) with consensus sequences of TTGaca and TAtaaT, respectively (where upper case letter denote more highly conserved positions). In the case of alternative σ factors related to σ^{70} , there are often distinct sequence motifs generally centered at these same two positions. In contrast, promoters recognized by the σ^N holoenzyme contain conserved elements near -24 and -12 [10]. Because of their central role in promoter recognition, these elements are often referred to as the core promoter elements.

In addition to the core promoter elements, other sequences throughout the extended promoter recognition region can have a large influence on promoter strength [Ross and Gourse, this volume]. The region upstream of the -35 element can contain phased A- and T-rich sequences which provide a site of interaction for the carboxyl-terminal domain of the two alpha-subunits [14,15]. This upstream promoter (UP) element can increase promoter strength by 100-fold or more and this DNA region, together with the -35 element, is likely the first recognized during promoter engagement. The best characterized UP element is that

of the very strong ribosomal RNA promoter *rrnB* P1 in *E. coli*. However, UP elements are found as part of many different promoters in *E. coli* [16] and sequence comparisons indicate that UP elements are common in other bacteria as well [17]. Since the UP element functions by interaction with the carboxy-terminal domain (CTD) of alpha, UP elements function with many different classes of holoenzyme [18]. In other cases, the region upstream of (or even overlapping) the -35 recognition element may instead serve as a binding site for positively acting transcription factors [19,20]. Interactions between RNAP and proteins bound to upstream regions of DNA can increase the association of RNAP with the promoter by providing favorable protein-protein contacts. The alpha CTD is a frequent target for transcription activators, but contacts with σ [20] and with beta-prime [21] are also documented. Fusion of a protein domain in place of the alpha CTD is one strategy to detect interactions between proteins. This genetic approach, analogous to the yeast two hybrid system, has provided a powerful tool for the dissection of protein-protein interactions both within RNAP and between RNAP and its many regulatory proteins [Nickels, this volume].

The region between the core promoter elements (the spacer sequence) is generally considered relatively unimportant as a determinant of promoter strength but this is perhaps an oversimplification. One class of σ^{70} promoters, for example, contains an "extended"-10 element that contains important recognition sequences for σ upstream of the conventional -10 element (TGnTATAaT). The presence of an extended -10 element can compensate or even substitute for a poor -35 region [22]. The region between the -10 element and +1 is known as the discriminator sequence and can also influence promoter properties and strength [23]. This region, like much of the -10 element itself, is converted to a single-stranded form during open complex formation. Finally, the sequences from +1 to +20 (designated the initial transcribed sequence or ITS) can also greatly influence the efficiency of RNA production by affecting, primarily, the promoter escape process [24].

The process of promoter recognition begins with the formation of a recognition complex between RNAP and the promoter region DNA [1,25]. This typically involves the recognition of one or both core elements by the associated σ subunits and may include interactions of alpha-CTD with the upstream DNA. In this initial closed complex, RNAP has not yet melted the two DNA strands and thereby engaged the template strand [26]. In subsequent steps, the σ subunit mediates the strand separation event that initiates at or near the -11 position and ultimately extends past +1. This promoter melting process involves sequence specific recognition of the non-template strand by σ and positioning of the template strand in the active site of the enzyme [27–29]. The isomerization of the closed- to the open-complex is temperature-dependent and, once formed, the open complex is often quite stable. This isomerization leads to an extension of the footprint of RNAP on the DNA from near +1 to near +20.

Once RNAP has formed an open complex, the second step, transcript initiation, can occur. In this reaction, RNAP binds two NTPs, complementary to the +1 and +2 positions on the template strand. The +1 NTP binding site is preferentially occupied by a purine (most RNA transcripts begin with either an A or a G) and this site (the *i* site) has a relatively high K_m (low affinity) for NTPs. The +2 NTP is bound in the (*i*+1) site. Catalysis of phosphodiester bond formation then occurs leading to release of pyrophosphate (PP_i) as product and translocation of the terminal NMP (in the pppNpN dinucleotide product) from the (*i*+1) to the (*i*) site. This cycle of NTP binding, catalysis, and transcript translocation continues with each subsequently incorporated NTP [30–32].

In general, RNAP prefers to initiate transcription within a narrow window located between 6 and 9 base-pairs downstream of the -10 element. If there is a suitably positioned T or C (on the template strand), initiation will begin with the preferred substrate: a purine residue. The

process of start-site selection reflects the balance between the optimal position of the start site and the preference for purines in the *i* site. Commonly, RNAP will initiate at 2 or more closely spaced positions and, in some systems, the different transcripts have different fates in the cell [33]. Bacteria have taken advantage of the process of NTP-dependent start site selection to regulate gene expression. One notable example is the CTP-dependent switching of start sites for the *E. coli pyrC* gene encoding a pyrimidine biosynthetic enzyme. When CTP levels are low, precluding initiation at the preferred distance from the -10 element (7 bp), RNAP will initiate with a GTP (a preferred NTP) at the non-preferred distance (9 bp downstream). The resulting GTP-initiated transcript is efficiently translated whereas that initiated with CTP is not [33].

Promoter clearance and elongation

The early stages of transcription often do not lead to a productive transcript [34]. The initial transcribing complex (ITC) retains the σ subunit and maintains contacts with the promoter elements. Short transcripts (typically between 2 and ~ 10 nt in length) are repeatedly synthesized and released from RNAP in a process termed abortive initiation [Hsu, this volume]. Often, RNAP will synthesize numerous abortive transcripts with, after each cycle of NTP addition, a certain probability of transcript release. Eventually, RNAP will successfully transition from an ITC to a stable elongation complex (EC). The EC has significantly reduced affinity for σ (and is generally assumed to have lost σ , although this has been a point of debate; [35]) and is highly stable and processive.

The early stages of transcription are also sensitive, at some promoters, to a phenomenon known as slippage or reiterative synthesis [33]. This occurs most prominently when the early transcribed region encodes three or more consecutive U residues. Reiterative synthesis will generally lead to non-productive transcription and is favored by high concentrations of UTP. Conversely, when UTP concentrations are lower, productive transcription is favored. *E. coli* has adapted this as a regulatory mechanism to modulate the rate of productive transcription of the *pyrBI* pyrimidine biosynthetic gene [33].

Once formed, the EC can synthesize RNA transcripts of many thousands of nucleotides without dissociation. However, the process of transcript elongation is not uniform but often includes sites where RNAP tentatively halts (pause sites) [36]. In some cases, pauses are observed proximal to the promoter suggesting that trapping of RNAP during the ITC to EC transition [37,38]. Promoter proximal pauses may result from interactions of σ factor, retained within the transcription complex, with -10 like sequences in the exposed non-template strand [39,40]. Pause sites may also result from a back-tracking reaction in which the 3'-end of the nascent transcript is removed from the active center of the enzyme [41]. Alternatively, secondary structures in the nascent RNA or encounters between the EC and DNA-bound proteins may also contribute to pausing [42,43]. *In vitro*, pausing is often dramatically enhanced when one or more NTP substrates is limiting for RNAP. There is evidence that pauses induced by lowered NTP pools may also be important in regulating gene expression in the cell. The best understood example is provided by attenuation control of the *E. coli pyrBI* gene [33]. In this case, pausing of RNAP, when UTP levels are low, allows a closer coupling between the elongating RNAP and a ribosome translating a leader peptide. Active translation, in turn, precludes formation of an intrinsic terminator that would otherwise lead to termination within the leader region of the *pyrBI* gene.

The interaction of elongation factors with RNAP, such as NusG and RfaH, affects the frequency and duration of pausing during transcription [44,45]. Pause sites during transcription can be visualized *in vitro* using polyacrylamide gel electrophoresis from single-cycle transcription reactions (initiated either from a promoter or from an EC generated by

NTP omission) as those transcripts that accumulate transiently during a time course experiment [e.g. Hsu, this volume; Artsimovitch and Henkin, this volume].

Termination and recycling

The final step of transcription is termination. At specific terminator sequences, or in response to the Rho termination factor, RNA synthesis ceases and the completed transcript is released [46]. In many cases, the structures that elicit transcript termination are G:C rich stem loops followed by a run of U residues. The ability of a given transcript to form a terminator structure can, in turn, be controlled by proteins or small molecules that interact with the nascent transcript to control transcription of downstream genes [Artsimovitch and Henkin, this volume]. These mechanisms include well characterized attenuation and anti-termination control pathways [33,47,48].

The events occurring at bacterial transcription termination sites are not well understood. In general, it is clear that the core RNAP is released and recycled by re-binding to σ factor to regenerate holoenzyme. The newly formed holoenzyme (with the same or a different σ factor) is now poised to begin the cycle anew. However, it is unclear whether the disassembly of the elongating ternary complex (so-called since it has three components: DNA, RNA, and protein) is an ordered or a random event. If the RNA transcript is released first, it is possible that the core RNAP remains bound to the template DNA prior to release. Release of stably bound core enzyme has been suggested as one activity of the RNA polymerase-associated protein, RapA [49,50]. Alternatively, if the DNA is released first, the resulting core-RNA complex may be slow to dissociate and this may impede efficient recycling. RNAP is known to be product inhibited, at least *in vitro*, and displacement of RNA from binary core:RNA complexes has been suggested to be the function of the RNAP delta subunit commonly found in Gram positive bacteria [51,52]. The process of efficient release and recycling of RNAP may also be directly facilitated by σ factor [53].

In order to re-enter the transcription cycle, the released core RNAP must engage productively with a σ factor and begin again the process of promoter localization. Formation of holoenzyme is governed, in part, by the abundance and availability of σ factors [9]. Since free core is generally limiting in the cell, it is assumed that σ factors compete to access the core RNAP [54]. This process can be further regulated by factors such as anti- σ factors, which can sequester σ factors and prevent core association, and possibly by factors that enhance σ -core association. Recently, the *E. coli* Crl protein has been described as such a positively-acting factor that facilitates the association of σ^S with core RNAP [55].

Determinants of promoter strength

Numerous studies have sought to understand the relationship between promoter sequence and intrinsic promoter strength. As a first approximation, the closer the core promoter elements (i.e. the -35 and -10 elements) are to consensus, the stronger the promoter is likely to be [25]. However, this is clearly an oversimplification. For example, if the core elements are too strong (within the context of the other promoter recognition elements), RNAP will bind very tightly but transcript initiation will be inefficient since RNAP will be impaired in promoter clearance [56]. Indeed, one study suggests that the correct correlation is between closeness of match to consensus of the core elements and the association rate constant for the promoter, rather than overall promoter strength [57]. Depending on the promoter sequence, and availability of the cognate holoenzyme in the cell, the promoter may be present in the largely un-occupied state (for promoters limited by RNAP binding), bound with RNAP in a closed or intermediate complex (for promoters limited by open complex formation), or associated with RNAP in an open complex or ITC (for promoters limited at

the clearance stage). Recent results suggest that the frequency of post-initiation pausing proximal to the promoter is likely to be much greater than previously appreciated [37,38].

Imposing regulation on the process of transcription

As the above discussion has made clear, the process of forming a transcript involves numerous steps that can be divided for convenience into the initiation, elongation, and termination but phases. Most often, the overall rate of transcript production is determined at the initiation phase. The sequence and context of the promoter recognition region, and the availability of the cognate holoenzyme, will determine the intrinsic (or basal) promoter strength which can vary over at least 4 orders of magnitude. Some promoters (e.g. ribosomal RNA operons) initiate synthesis extremely rapidly with RNAPs essentially lining up along the transcription unit limited in their proximity only by their physical size. Others may initiate productive synthesis on the same time scale as the bacterial generation time.

Imposed on this intrinsic strength is the influence of the various global and gene-specific regulators [e.g. Minchin and Busby, this volume]. Regulatory proteins can function to either enhance the rate of initiation (transcription activation) or diminish initiation (transcription repression). Repression is clearly the simplest to understand and, mechanistically, the easiest to accomplish [58,59]. Any protein that binds to sites overlapping the promoter recognition elements can sterically occlude the promoter and prevent transcription initiation. This mechanism accounts for the majority of repression, but there are numerous examples where repressors act at stages subsequent to promoter engagement by, for example, preventing promoter isomerization or clearance.

Activation is more difficult to define and often requires more sophisticated approaches to understand at the mechanistic level. The ability of a protein to enhance the rate of productive initiation from a particular promoter requires that the activator increase the rate of the slowest (or rate-limiting) step. In some cases, this is RNAP-binding, but in other examples it might be isomerization or clearance [60,61]. Thus, understanding the mechanisms of transcription activator proteins, even in the simplest cases, requires that one first define the overall kinetic parameters of the promoter in both the basal and activated states. At the simplest levels, these include those steps that are dependent on RNAP concentration (and therefore loosely correlated with binding) and those that are concentration-independent (e.g. isomerization, at least at some promoters, and clearance). One common technique for distinguishing between these two fundamental steps in initiation is a "tau plot" as described by Ross and Gourse [this volume].

In addition to "simple" activation and repression mechanisms, many systems have evolved to integrate signals from multiple regulatory pathways to control expression of a particular gene or operon. This can involve multiple regulators working together to form an activation complex [see Minchin and Busby, this volume, for an example], or multiple promoters controlling expression of a single operon. In some cases, multiple holoenzyme species can activate closely spaced or even superimposed promoter elements thereby allowing generation of similar (or even identical) transcripts by different forms of holoenzyme [62,63]. Finally, operons can be subject to multiple levels of control including regulatory inputs both at the level of transcript initiation and transcript elongation/termination.

Finally, there are now numerous examples where the cell has exploited the unique properties of RNAP as an NTP-dependent enzyme to regulate transcription. Notably, the growth rate control of many operons in *E. coli* has been shown to depend on promoters such as the *rrnB* P1 promoter where the affinity of the *i* site for NTP is unusually low [64,65]. As a result, these low affinity (high K_m) promoters are highly sensitive to fluctuations in cellular NTP pools which, of course, correlate with energy availability and therefore growth rate. In other

cases, RNAP detects the levels of one or another NTP by its effect on start site selection and/or reiterative synthesis with the consequence that mRNAs initiating at slightly different positions downstream of a common -10 element have different fates in the cell [33]. The process of start site selection can be governed by the availability of either the $+1$ or the $+2$ NTP, depending on the promoter, although the former is more common since the $+1$ NTP binds with the weakest affinity to the open complex.

Regulation within the context of the cell

Most of above considerations apply to the interactions of RNAP with a single regulatory region preceding a single transcription unit. Frequently, these relatively simple regulatory interactions can be recapitulated, at least qualitatively, in a reconstituted *in vitro* system with purified RNAP and appropriate transcription factors. However, in the context of the cell additional factors can also be important in determining both absolute and relative transcription levels. These include the presence of numerous global regulators and nucleoid-associated proteins, negative-supercoiling of the chromosomal DNA, the large concentration of non-specific DNA (which effectively buffers the action of all DNA-binding proteins; [66]), and competition amongst the numerous promoters and regulatory factors for limiting pools of RNAP holoenzyme. Moreover, for those factors that interact with the elongation complex the kinetics of elongation, and the process of co-transcriptional RNA folding, may be affected by factors (elongation factors; translation of the nascent transcript, etc.) missing from a simplified *in vitro* system. Given these complexities, it is often challenging to recapitulate all aspects of a regulatory system with purely biochemical approaches and complementary *in vivo* analyses provide an important reality check on inferences made using such reconstituted systems.

Overview of approaches to the investigation of bacterial transcription regulation

The purpose of this volume is to provide the experimenter with a concise overview of some of the most important methods that have enabled researchers to dissect the more common types of bacterial regulatory mechanisms. Since these mechanisms are often best understood in *E. coli*, most of the relevant procedures have been developed with *E. coli* RNAP but they are, in many cases, applicable to other organisms and to holoenzymes other than the primary, σ^{70} -holoenzyme. This is particularly true for the biochemical approaches highlighted here since DNA, RNA, and proteins are, for most practical purposes, chemically identical regardless of source organism. Naturally, the genetic approaches are much more organism-specific.

We begin this volume with a survey of the philosophy and approaches that were used to define one complex regulatory system: the *melAB* operon in *E. coli* [Minchin and Busby, this volume]. The following chapters describe the most useful biochemical approaches for monitoring RNAP promoter association and engagement [Ross and Gourse, this volume], promoter clearance [Hsu, this volume], and the elongation and termination reactions [Artsimovitch and Henkin, this volume]. The biochemical characterization of RNAP and its activities of course requires sources of purified enzyme. RNAP and its purification is well described in other volumes [67–69], but we here include a summary of recent advances for the reconstitution of holoenzymes from overproduced and purified subunits [Kuznedelov and Severinov, this volume]. As noted above, many activator proteins function by directly contacting one or more subunits of RNAP to enhance either RNAP binding or subsequent stages in the initiation process. The use of bacterial two-hybrid systems [Nickels, this volume] has provided a convenient and accessible genetic approach for defining these interactions and, importantly, enables the selection of mutants altered in these precise

contacts. Finally, we conclude with two complementary global approaches. The use of DNA microarrays in which one or more probes for each gene in a genome are spotted (typically on glass slides) provides a powerful tool for the analysis of transcription (or more accurately, RNA levels) on a genome-wide scale (transcriptomics). An additional powerful approach is to correlate the observed *in vivo* effects of a particular transcription factor (as judged by changes in the transcriptome) with actual protein occupancy on the DNA (as judged by chromatin immunoprecipitation-DNA chip, also known as ChIP-to-chip analysis) [Rhodius and Wade, this volume]. One complexity with conventional transcriptomic approaches is that it is sometimes difficult to distinguish direct and indirect effects of genetically or chemically perturbing the activity of a regulatory protein. The use of an alternative *in vitro* method, run-off transcription/microarray analysis (ROMA), provides one solution and allows the activity of transcription factors to be monitored on a genome-wide scale in the absence of downstream effects [MacLellan *et al.*, this volume]. In sum, these approaches should provide a powerful set of tools for dissecting the complex and fascinating molecular mechanisms that underlie the ability of cells to synthesis the proper set of proteins in response to an ever-changing environment.

Acknowledgments

I would like to thank Dr. M.J. Chamberlin for introducing me to the wonders of the bacterial transcription cycle, and my many colleagues in the bacterial transcription community for sharing this passion. Work on transcriptional control mechanisms in the Helmann laboratory is supported by grants from the NIH (GM-047446 and GM-059323) and the NSF (MCB-0640616).

References

1. Borukhov S, Nudler E. Trends Microbiol 2008;16:126–134. [PubMed: 18280161]
2. Murakami KS, Darst SA. Curr Opin Struct Biol 2003;13:31–39. [PubMed: 12581657]
3. Darst SA. Curr Opin Struct Biol 2001;11:155–162. [PubMed: 11297923]
4. Chamberlin MJ, Nierman WC, Wiggs J, Neff N. J Biol Chem 1979;254:10061–10069. [PubMed: 114520]
5. Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K. Proc Natl Acad Sci U S A 2001;98:892–897. [PubMed: 11158566]
6. Mathew R, Chatterji D. Trends Microbiol 2006;14:450–455. [PubMed: 16908155]
7. Paget MS, Helmann JD. Genome Biol 2003;4:203. [PubMed: 12540296]
8. Helmann JD. Adv Microb Physiol 2002;46:47–110. [PubMed: 12073657]
9. Gruber TM, Gross CA. Annu Rev Microbiol 2003;57:441–466. [PubMed: 14527287]
10. Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, Buck M. Mol Microbiol 2008;68:538–546. [PubMed: 18331472]
11. Zenkin N, Severinov K. Cell Mol Life Sci 2008;65:2280–2288. [PubMed: 18373065]
12. von Hippel PH, Berg OG. J Biol Chem 1989;264:675–678. [PubMed: 2642903]
13. Bustamante C, Guthold M, Zhu X, Yang G. J Biol Chem 1999;274:16665–16668. [PubMed: 10358002]
14. Estrem ST, Gaal T, Ross W, Gourse RL. Proc Natl Acad Sci U S A 1998;95:9761–9766. [PubMed: 9707549]
15. Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL. Science 1993;262:1407–1413. [PubMed: 8248780]
16. Ross W, Aiyar SE, Salomon J, Gourse RL. J Bacteriol 1998;180:5375–5383. [PubMed: 9765569]
17. Helmann JD. Nucleic Acids Res 1995;23:2351–2360. [PubMed: 7630711]
18. Fredrick K, Caramori T, Chen YF, Galizzi A, Helmann JD. Proc Natl Acad Sci U S A 1995;92:2582–2586. [PubMed: 7708689]
19. Busby S, Ebright RH. J Mol Biol 1999;293:199–213. [PubMed: 10550204]
20. Dove SL, Darst SA, Hochschild A. Mol Microbiol 2003;48:863–874. [PubMed: 12753182]

21. Miller A, Wood D, Ebright RH, Rothman-Denes LB. *Science* 1997;275:1655–1657. [PubMed: 9054361]
22. Mitchell JE, Zheng D, Busby SJ, Minchin SD. *Nucleic Acids Res* 2003;31:4689–4695. [PubMed: 12907708]
23. Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, Gourse RL. *Cell* 2006;125:1069–1082. [PubMed: 16777598]
24. Hsu LM, Cobb IM, Ozmore JR, Khoo M, Nahm G, Xia L, Bao Y, Ahn C. *Biochemistry* 2006;45:8841–8854. [PubMed: 16846227]
25. Haugen SP, Ross W, Gourse RL. *Nat Rev Microbiol* 2008;6:507–519. [PubMed: 18521075]
26. deHaseth PL, Zupancic ML, Record MT Jr. *J Bacteriol* 1998;180:3019–3025. [PubMed: 9620948]
27. Helmann JD, deHaseth PL. *Biochemistry* 1999;38:5959–5967. [PubMed: 10320321]
28. Young BA, Gruber TM, Gross CA. *Science* 2004;303:1382–1384. [PubMed: 14988563]
29. Borukhov S, Nudler E. *Curr Opin Microbiol* 2003;6:93–100. [PubMed: 12732296]
30. Erie DA, Yager TD, von Hippel PH. *Annu Rev Biophys Biomol Struct* 1992;21:379–415. [PubMed: 1381976]
31. Vassilyev DG, Artsimovitch I. *Cell* 2005;123:977–979. [PubMed: 16360025]
32. Greive SJ, von Hippel PH. *Nat Rev Mol Cell Biol* 2005;6:221–232. [PubMed: 15714199]
33. Turnbough CL Jr, Switzer RL. *Microbiol Mol Biol Rev* 2008;72:266–300. table of contents. [PubMed: 18535147]
34. Hsu LM. *Biochim Biophys Acta* 2002;1577:191–207. [PubMed: 12213652]
35. Mooney RA, Darst SA, Landick R. *Mol Cell* 2005;20:335–345. [PubMed: 16285916]
36. Landick R. *Biochem Soc Trans* 2006;34:1062–1066. [PubMed: 17073751]
37. Artsimovitch I. *Mol Microbiol* 2008;68:1–3. [PubMed: 18312277]
38. Hatoum A, Roberts J. *Mol Microbiol* 2008;68:17–28. [PubMed: 18333883]
39. Ring BZ, Yarnell WS, Roberts JW. *Cell* 1996;86:485–493. [PubMed: 8756730]
40. Brodolin K, Zenkin N, Mustaev A, Mamaeva D, Heumann H. *Nat Struct Mol Biol* 2004;11:551–557. [PubMed: 15122346]
41. Toulme F, Mosrin-Huaman C, Artsimovitch I, Rahmouni AR. *J Mol Biol* 2005;351:39–51. [PubMed: 15993420]
42. Neuman KC, Abbondanzieri EA, Landick R, Gelles J, Block SM. *Cell* 2003;115:437–447. [PubMed: 14622598]
43. Komissarova N, Velikodvorskaya T, Sen R, King RA, Banik-Maiti S, Weisberg RA. *Mol Cell* 2008;31:683–694. [PubMed: 18775328]
44. Svetlov V, Belogurov GA, Shabrova E, Vassilyev DG, Artsimovitch I. *Nucleic Acids Res* 2007;35:5694–5705. [PubMed: 17711918]
45. Artsimovitch I, Landick R. *Proc Natl Acad Sci U S A* 2000;97:7090–7095. [PubMed: 10860976]
46. Ciampi MS. *Microbiology* 2006;152:2515–2528. [PubMed: 16946247]
47. Henkin TM, Yanofsky C. *Bioessays* 2002;24:700–707. [PubMed: 12210530]
48. Winkler WC, Breaker RR. *Annu Rev Microbiol* 2005;59:487–517. [PubMed: 16153177]
49. Sukhodolets MV, Cabrera JE, Zhi H, Jin DJ. *Genes Dev* 2001;15:3330–3341. [PubMed: 11751638]
50. Nechaev S, Severinov K. *Structure* 2008;16:1294–1295. [PubMed: 18786392]
51. Juang YL, Helmann JD. *J Mol Biol* 1994;239:1–14. [PubMed: 7515111]
52. Lopez de Saro FJ, Woody AY, Helmann JD. *J Mol Biol* 1995;252:189–202. [PubMed: 7545758]
53. Arndt KM, Chamberlin MJ. *J Mol Biol* 1988;202:271–285. [PubMed: 2459392]
54. Grigorova IL, Phleger NJ, Mutalik VK, Gross CA. *Proc Natl Acad Sci U S A* 2006;103:5332–5337. [PubMed: 16567622]
55. England P, Westblade LF, Karimova G, Robbe-Saule V, Norel F, Kolb A. *J Biol Chem*. 2008
56. Vo NV, Hsu LM, Kane CM, Chamberlin MJ. *Biochemistry* 2003;42:3798–3811. [PubMed: 12667071]
57. Brunner M, Bujard H. *Embo J* 1987;6:3139–3144. [PubMed: 2961560]

58. Rojo F. *J Bacteriol* 1999;181:2987–2991. [PubMed: 10321997]
59. Rojo F. *Curr Opin Microbiol* 2001;4:145–151. [PubMed: 11282469]
60. Rhodius VA, Busby SJ. *Curr Opin Microbiol* 1998;1:152–159. [PubMed: 10066477]
61. Gralla JD. *Curr Opin Genet Dev* 1996;6:526–530. [PubMed: 8939721]
62. Gaal T, Ross W, Estrem ST, Nguyen LH, Burgess RR, Gourse RL. *Mol Microbiol* 2001;42:939–954. [PubMed: 11737638]
63. Qiu J, Helmann JD. *J Bacteriol* 2001;183:1921–1927. [PubMed: 11222589]
64. Gaal T, Bartlett MS, Ross W, Turnbough CL Jr, Gourse RL. *Science* 1997;278:2092–2097. [PubMed: 9405339]
65. Krasny L, Gourse RL. *Embo J* 2004;23:4473–4483. [PubMed: 15496987]
66. Berg OG, von Hippel PH. *Trends Biochem Sci* 1988;13:207–211. [PubMed: 3079537]
67. Helmann JD. *Methods Enzymol* 2003;370:10–24. [PubMed: 14712630]
68. Anthony JR, Green HA, Donohue TJ. *Methods Enzymol* 2003;370:54–65. [PubMed: 14712633]
69. Kuznedelov K, Minakhin L, Severinov K. *Methods Enzymol* 2003;370:94–108. [PubMed: 14712637]
70. Wassarman KM. *Mol Microbiol* 2007;65:1425–1431. [PubMed: 17714443]
71. Rhodius VA, West DM, Webster CL, Busby SJ, Savery NJ. *Nucleic Acids Res* 1997;25:326–332. [PubMed: 9016561]
72. Paul BJ, Ross W, Gaal T, Gourse RL. *Annu Rev Genet* 2004;38:749–770. [PubMed: 15568992]
73. Smith TL, Sauer RT. *Proc Natl Acad Sci U S A* 1996;93:8868–8872. [PubMed: 8799119]
74. Menendez M, Kolb A, Buc H. *Embo J* 1987;6:4227–4234. [PubMed: 2832158]
75. Yakhnin AV, Yakhnin H, Babitzke P. *Proc Natl Acad Sci U S A*. 2008
76. Nickels BE, Hochschild A. *Cell* 2004;118:281–284. [PubMed: 15294154]
77. Grundy FJ, Henkin TM. *Crit Rev Biochem Mol Biol* 2006;41:329–338. [PubMed: 17092822]
78. Campbell EA, Westblade LF, Darst SA. *Curr Opin Microbiol* 2008;11:121–127. [PubMed: 18375176]

Table 1

An Overview of the Transcription Cycle and Selected Examples of Regulatory Factors

Stage of Transcription	Intermediate ^I	Examples of Regulation	Reference(s)
Promoter recognition			
(i) RNA polymerase (E σ) binds to DNA	RD _{NS}	6S RNA	[70]
(ii) Promoter binding	RP _C	Alternative σ factors, CAP (class I sites)	[9,19]
(iii) Nucleation	RP _N		
(iv) Melting of promoter DNA	RP _O	CAP (class II sites)	[71]
Transcript Initiation and Clearance			
(i) NTP/RNA binding at (<i>i</i>) site	RP _O	Growth-rate control (ATP as effector); NTP-sensitive start site selection	[33,72]
(ii) NTP binding at (<i>i</i> +1) site	RP _O	<i>E. coli codBA</i> (sensitive to UTP-binding at the <i>i</i> +1 site for one of two alternate transcripts)	[33]
(iii) Bond formation, PP _i release	ITC		
(iv) Translocation – repeat cycle; OR Release abortive transcript; start again	ITC		
(v) Release core promoter contacts (escape)	ITC / EC	CAP at <i>malT</i> , P22 Arc	[44,73,74]
Elongation		NusG, RfaH	[44,75]
NTP binding at (<i>i</i> +1) site, bond formation, translocation (back-tracking)	EC	GreA, GreB; UTP-sensitive attenuation	[33,76]
Termination		Rho, Riboswitches	[48,77]
Pausing of RNAP, Transcript and core release	PTC	RapA; proposed to mediate core release, <i>B. subtilis</i> δ subunit and RNA displacement	[50–52]
Recycling			
Binding of σ to core to regenerate holoenzyme	E σ	Crl (for σ^S), anti- σ factors	[55,78]

^IThe abbreviations for the various intermediates are RD_{NS} (RNAP bound to DNA, non-specifically), RP_C, RP_N and RP_O (closed, nucleated, and open complex, respectively), ITC (initial transcribing complex), EC (elongation complex), PTC (post-termination complex), E σ (holoenzyme).