

# Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters

India G. Hook-Barnard and Deborah M. Hinton

Gene Expression and Regulation Section, Laboratory of Molecular and Cellular Biology, National Institute of Diabetes Digestive and Kidney Diseases, National Institutes of Health, Bldg. 8 Room 2A-13, Bethesda, MD 20892-0830.

**Abstract:** Bacterial RNA polymerase is composed of a core of subunits ( $\beta$ ,  $\beta'$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\omega$ ), which have RNA synthesizing activity, and a specificity factor ( $\sigma$ ), which identifies the start of transcription by recognizing and binding to sequence elements within promoter DNA. Four core promoter consensus sequences, the  $-10$  element, the extended  $-10$  (TGn) element, the  $-35$  element, and the UP elements, have been known for many years; the importance of a nontemplate G at position  $-5$  has been recognized more recently. However, the functions of these elements are not the same. The AT-rich UP elements, the  $-35$  elements ( $^{-35}\text{TTGACA}^{-30}$ ), and the extended  $-10$  ( $^{-15}\text{TGn}^{-13}$ ) are recognized as double-stranded binding elements, whereas the  $-5$  nontemplate G is recognized in the context of single-stranded DNA at the transcription bubble. Furthermore, the  $-10$  element ( $^{-12}\text{TATAAT}^{-7}$ ) is recognized as both double-stranded DNA for the T:A bp at position  $-12$  and as nontemplate, single-stranded DNA from positions  $-11$  to  $-7$ . The single-stranded sequences at positions  $-11$  to  $-7$  as well as the  $-5$  contribute to later steps in transcription initiation that involve isomerization of polymerase and separation of the promoter DNA around the transcription start site. Recent work has demonstrated that the double-stranded elements may be used in various combinations to yield an effective promoter. Thus, while some minimal number of contacts is required for promoter function, polymerase allows the elements to be mixed and matched. Interestingly, which particular elements are used does not appear to fundamentally alter the transcription bubble generated in the stable complex. In this review, we discuss the multiple steps involved in forming a transcriptionally competent polymerase/promoter complex, and we examine what is known about polymerase recognition of core promoter elements. We suggest that considering promoter elements according to their involvement in early (polymerase binding) or later (polymerase isomerization) steps in transcription initiation rather than simply from their match to conventional promoter consensus sequences is a more instructive form of promoter classification.

**Abbreviations:** ss: single-stranded; ds: double-stranded; EMSA: electrophoretic mobility shift assay; bp: base pair; n: any nucleotide; r: purine; y: pyrimidine; w: A or T; NTP: ribonucleoside triphosphate.

**Keywords:** polymerase, sigma70, promoter, transcription

## The Multi-Step Process of Transcription Initiation

The process of transcription begins when RNA polymerase recognizes and binds to DNA elements within a promoter sequence (reviewed in (Browning and Busby, 2004; Murakami and Darst, 2003; Young et al. 2002)). In bacteria, RNA polymerase is composed of a core of multiple subunits ( $\beta$ ,  $\beta'$ ,  $\alpha_1$ ,  $\alpha_2$ , and  $\omega$ ) that is tightly associated with a  $\sigma$  specificity factor. While core contains the active site for polymerase and thus is capable of synthesizing RNA, the sigma factor controls when and where transcription is initiated. Prokaryotes have multiple sigma factors; a primary  $\sigma$  that is needed for the expression of housekeeping genes during exponential growth and alternate  $\sigma$  factors that are used under certain conditions of growth or stress (Gruber and Gross, 2003; Paget and Helmann, 2003). The primary  $\sigma$  factor in *E. coli* is  $\sigma^{70}$ . Like other primary  $\sigma$  factors,  $\sigma^{70}$  can be divided into four major domains (regions 1, 2, 3, and 4) and subdomains (2.1, 2.2, 2.3, etc), based on function, structure, and sequence conservation (Gruber and Gross, 2003; Lonetto et al. 1992) (Fig. 1A and B).

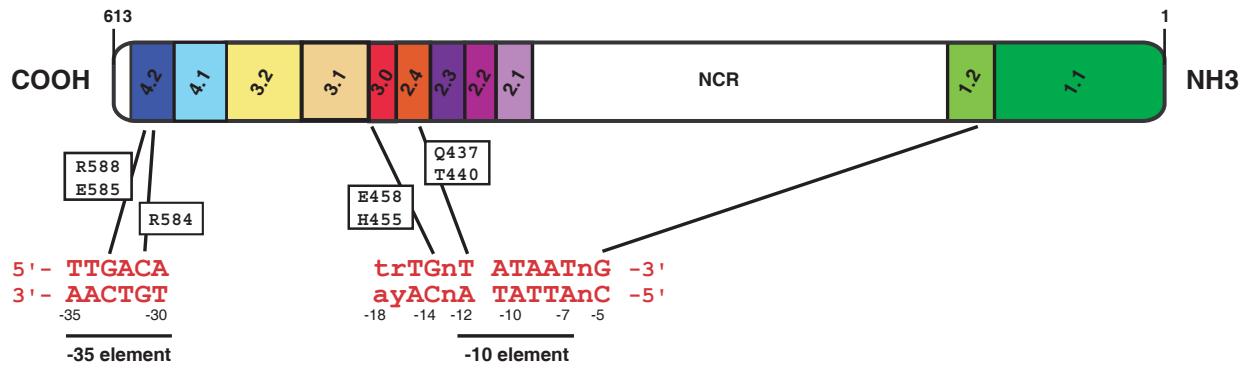
Detailed studies of transcription initiation at a few well-characterized promoters have demonstrated that initiation by *E. coli* RNA polymerase is a multi-step process (Fig. 2) that requires multiple contacts between the polymerase and the DNA (Buc and McClure, 1985; Craig et al. 1998; Davis et al. 2007a; Kontur et al. 2006; Schickor et al. 1990; Sclavi et al. 2005; Spassky et al. 1985). Each step on the path from free RNA polymerase and promoter to the final transcriptionally competent complex is an opportunity to regulate the initiation process. In addition to biochemical and kinetic studies, structural

**Correspondence:** India G. Hook-Barnard, Tel: (301) 435-3374; Fax: (301) 402-0053; Email: indiah@intra.niddk.nih.gov

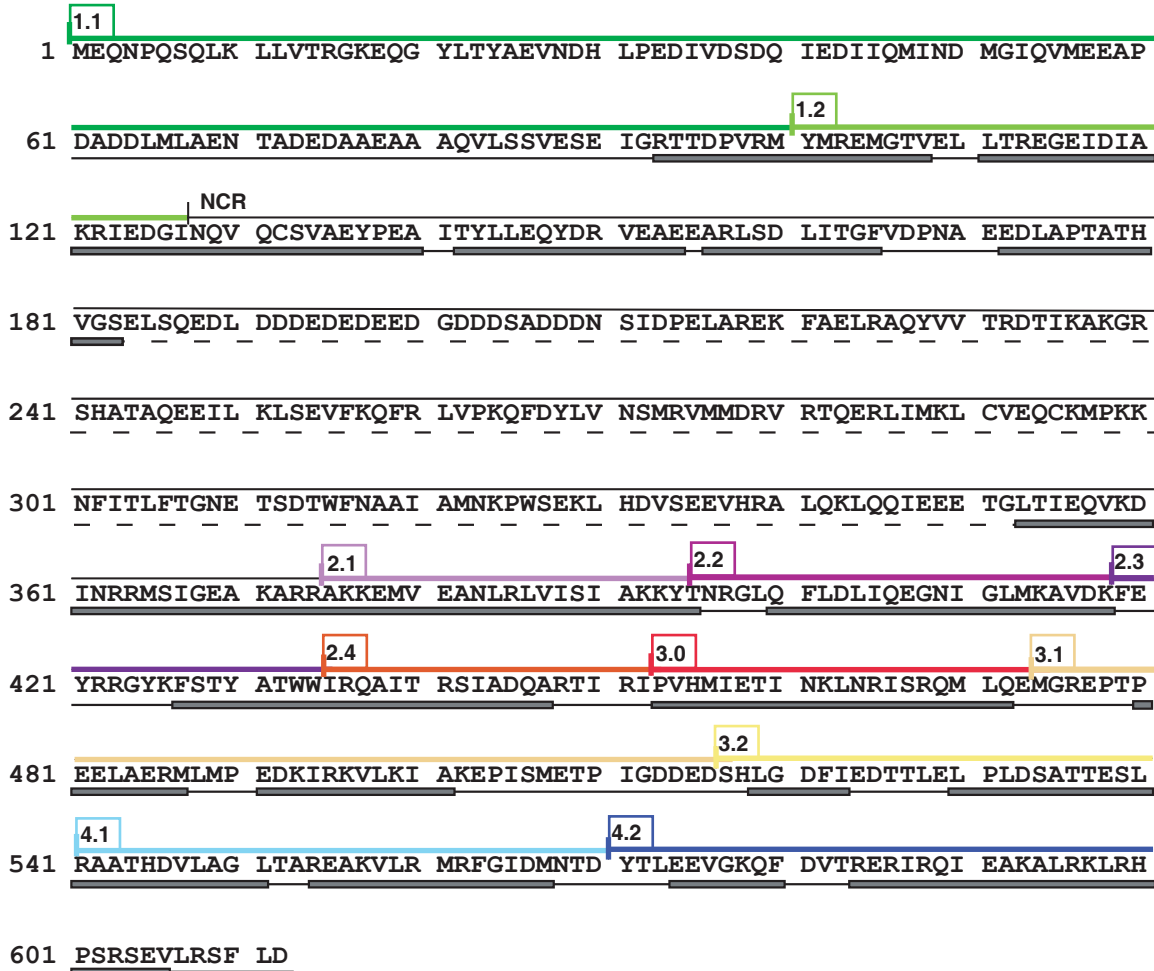


Copyright in this article, its metadata, and any supplementary data is held by its author or authors. It is published under the Creative Commons Attribution By licence. For further information go to: <http://creativecommons.org/licenses/by/3.0/>.

# A. Regions of $\sigma^{70}$



# B.

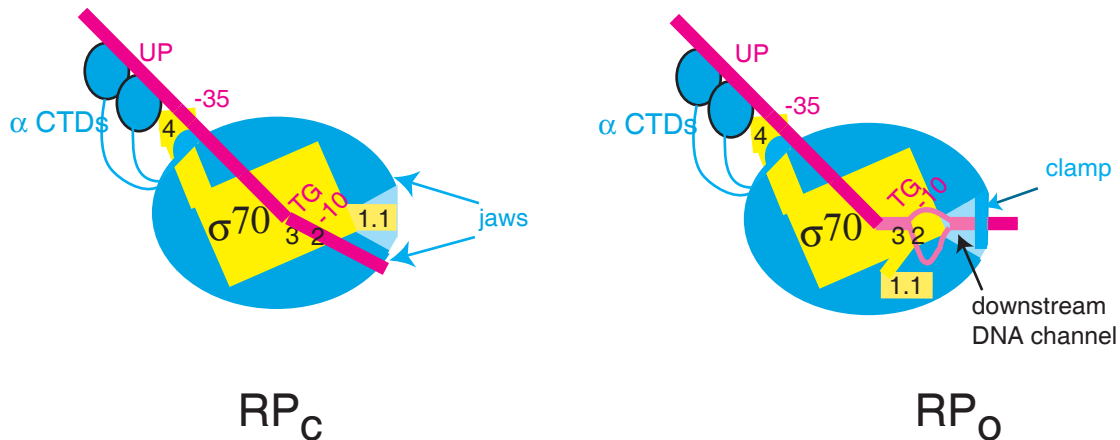


**Figure 1. (A)** Schematic of *E. coli*  $\sigma^{70}$ . Subregions (1.1, 1.2, etc) and non-conserved region (NCR) are based on function, structure, and sequence conservation (Gruber and Gross, 2003). Canonical promoter sequence elements are shown below sigma. Specific  $\sigma^{70}$  residues, which are thought to interact with base determinants, are indicated (see text for details). **(B)** Amino acid sequence of *E. coli*  $\sigma^{70}$ . Subregions of  $\sigma^{70}$  are indicated above the sequence and are colored as in A. Below the sequence, secondary structure is represented as gray boxes ( $\alpha$ -helices), black lines (coils) and dashed lines (disordered structure). (See (Gruber and Gross, 2003; Vassilyev et al. 2002) and text for details.)

A.



B.



**Figure 2.** (A) Steps in transcription initiation. R (polymerase) and P (promoter) first interact to form a closed complex ( $RP_C$ ). This complex then proceeds through intermediates ( $I_1$  and  $I_2$ ) to form the open complex ( $RP_O$ ). (B) Cartoon depicting polymerase promoter contacts in  $RP_C$  and  $RP_O$ . Core polymerase (subunits  $\beta$ ,  $\beta'$ ,  $\alpha_2$ , and  $\omega$ ) is shown in teal,  $\sigma^{70}$  is shown in yellow, and the DNA is shown in magenta. Interactions between the C-terminal domains of the  $\alpha$  subunits ( $\alpha$  CTDs) and the UP element,  $\sigma^{70}$  region 4 and the  $-35$  element,  $\sigma^{70}$  region 3 and the  $^{-15}\text{TGN}^{-13}$  element, and  $\sigma^{70}$  region 2 and the  $-10$  element are indicated.  $RP_C$  is a closed complex in which the DNA has not yet entered the primary channel. Full entry of DNA into the channel is blocked by the presence of  $\sigma^{70}$  region 1.1. In  $RP_O$ ,  $\sigma^{70}$  region 1.1 has moved, the DNA from around  $-11$  to  $+3$  is unwound, the template strand has descended into the active site of polymerase, and a portion of  $\beta'$ , called the clamp, has secured the downstream DNA (Kontur et al. 2006). See text for details and additional references.

analyses using polymerase, core, and the primary  $\sigma$  of thermophilic bacteria with and without DNA have provided a detailed model upon which to conceptualize these steps (Campbell et al. 2002; Murakami et al. 2002a; Murakami et al. 2002b; Murakami and Darst, 2003; Vassylyev et al. 2002; Young et al. 2002). The structure of a portion of *E. coli*  $\sigma^{70}$  has also been obtained (Malhotra et al. 1996). Thus, we can now generally chart the changes in DNA and polymerase that occur during initiation. In this review, we summarize data that addresses the contribution of individual promoter elements to the process of transcription initiation. However, to understand these interactions, we must first discuss the identified steps in transcription initiation that lead to the start of transcription.

### The closed ( $RP_C$ ) polymerase/promoter complex

Initial recognition and binding of promoter DNA (P) by RNA polymerase (R) is referred to as closed complex,  $RP_C$  (Fig. 2), because the DNA is

fully double-stranded (ds) or closed. In  $RP_C$  the polymerase specifically recognizes ds binding elements and is able to form a complex with the DNA detectable by electrophoretic mobility shift assay (EMSA). As we detail in sections below, specific promoter elements that may be used for recognition include: UP elements (positioned between  $-60$  and  $-40$ ) that are bound by the C-terminal domains of the  $\alpha$  subunits of core ( $\alpha$ -CTDs), a  $-35$  element that is recognized by residues in  $\sigma^{70}$  region 4, a TGn motif that is recognized by residues in  $\sigma^{70}$  region 3, and sequences within a  $-10$  element, primarily the base pair at  $-12$ , that is recognized by residues in  $\sigma^{70}$  region 2 (Fig. 1 and 2). Although the  $RP_C$  complex is specific, it is also readily reversible and therefore, easily challenged by the presence of competitor such as nonspecific DNA or the polyanion heparin. Footprinting data demonstrate that the polymerase initially protects the DNA from cleavage from at least  $-55$  to  $+5$ . However, this short footprint is only seen for the earliest closed complex when footprints are carried out at

lower temperatures (e.g. 8 °C) (Schickor et al. 1990) or with polymerase mutants (Cook and deHaseth, 2007) that trap the complex at this stage of the process. This protection pattern is consistent with the idea that in R<sub>Pc</sub> the promoter DNA lies outside of the polymerase primary channel and is mostly undistorted (Kontur et al. 2006; Mekler et al. 2002; Murakami and Darst, 2003).

### Kinetic intermediates on the path from R<sub>Pc</sub> to R<sub>Po</sub>

For transcription initiation to occur, R<sub>Pc</sub> must transition into a species, R<sub>Po</sub>, in which the DNA is both bent and unwound (open), and polymerase has undergone major conformational changes, called isomerization. Kinetic analyses (Buc and McClure, 1985; Craig et al. 1998; Davis et al. 2007a; Kontur et al. 2006; McKane et al. 2001; Saecker et al. 2002; Schickor et al. 1990; Sclavi et al. 2005; Spassky et al. 1985; Straney and Crothers, 1985; Straney and Crothers, 1987a; Straney and Crothers, 1987b) have revealed two significant intermediates in the pathway from R<sub>Pc</sub> to R<sub>Po</sub>. For the lambda promoter P<sub>R</sub>, the first intermediate I<sub>1</sub> is characterized by a bending of the DNA and the extension of the footprint to +20. Recent work suggests that this early step may be facilitated by the presence of far upstream DNA via interaction with other components of core (Davis et al. 2005; Davis et al. 2007a; Ross and Gourse, 2005). Wrapping of upstream DNA has been proposed to reposition portions of the β and β' subunits called the jaws, widening the primary channel, and allowing the DNA to move into the downstream end of this channel (Saecker et al. 2002) (Davis et al. 2007b). Alternatively, (or in addition) contacts between the far upstream DNA and the α-CTD of RNAP may aid in this transition (Ross and Gourse, 2005). However, despite these changes, in I<sub>1</sub> the polymerase has not yet fully undergone its isomerization process, and, thus, I<sub>1</sub> is competitor sensitive.

A later intermediate for lambda P<sub>R</sub>, I<sub>2</sub>, requires a large conformational change in polymerase. Because σ<sup>70</sup> region 1.1 (the N-terminal 100 residues of σ<sup>70</sup>) is thought to reside within the primary channel in the absence of DNA (Kontur et al. 2006; Mekler et al. 2002; Murakami and Darst, 2003), region 1.1 must relocate to make the channel fully accessible for the DNA (Fig. 2B). The movement of region 1.1 out of the channel is proposed to be coupled to the late folding of portions of the β' subunit, designated the 'clamp'. Movement of

region 1.1 can affect the rate of the transition from R<sub>Pc</sub> to R<sub>Po</sub>, but this modulation varies with the promoter context (Hook-Barnard and Hinton, unpublished) (Vuthoori et al. 2001; Wilson and Dombroski, 1997). Also, in I<sub>2</sub>, DNA downstream of -12 begins to unwind and the template strand begins its decent in the active site (Kontur et al. 2006). Evidence indicates that the -11A in the -10 element is crucial for instigating the melting process, which then propagates downstream through the start site (Fenton and Gralla, 2001; Lim et al. 2001; Tsujikawa et al. 2002). It is currently unclear which begins first, movement of region 1.1 out of the channel or DNA melting. However, the movement of region 1.1, together with DNA melting appears to be the rate-limiting step, and the order may be determined by the promoter context.

### The open (R<sub>Po</sub>) polymerase/promoter complex

The stable R<sub>Po</sub> species is achieved when the DNA around the transcriptional start site is fully separated, creating a transcription bubble from around -11 to +3, the template strand is located in the active site with the +1 nucleotide ready for base pairing with incoming NTPs, and the polymerase 'clamp' fully closes onto the DNA that is lying in the DNA channel (Fig. 2B) (Brodolin et al. 2005; Davis et al. 2007a; Kainz and Roberts, 1992; Lim et al. 2001; Nguyen and Burgess, 1997; Sasse-Dwight and Gralla, 1989; Sasse-Dwight and Gralla, 1991). In R<sub>Po</sub>, the nontemplate strand nucleotides of the -10 element interact with σ region 2.3, stabilizing the polymerase-promoter complex. Additionally, the nucleotide at -5 may be recognized by σ<sup>70</sup> region 1.2 (Feklistov et al. 2006; Haugen et al. 2006; Zenkin et al. 2007). R<sub>Po</sub> is the primary species at 37 °C and the DNase footprint at R<sub>Po</sub> typically extends from -55 to +25 (Davis et al. 2007a; Schickor et al. 1990). R<sub>Po</sub> is usually resistant to competition, although the stability of this complex varies with the promoter context (Haugen et al. 2006; Vuthoori et al. 2001).

### Double-Stranded Promoter Elements Involved in Transcription Initiation

For more than 25 years, it has been known that σ<sup>70</sup>-dependent promoters require specific sequences for efficient transcription. Early work revealed consensus sequences at the -35 (<sup>-35</sup>TTGACA<sup>-30</sup>) and -10 (<sup>-12</sup>TATAAT<sup>-7</sup>) regions of σ<sup>70</sup>-dependent

## A.

	UP elements				-10 element			+1
	distal	proximal	-35 element		trTGnT	ATAAT	G	
	<b>NNAWWWWTTTT</b>	<b>AAAAARNR</b>	<b>TTGACA</b>		<b>trTGnT</b>	<b>ATAAT</b>	<b>G</b>	<b>A</b>
			<sup>-33 -31</sup>		<sup>-14 -12</sup>	<sup>-11</sup>	<sup>-7 -5</sup>	
Plac	GCTCACTCATTAGGCACCCAGGC	<b>TTTACA</b>	CTTTATGCTTCCG	GCTCGT	<b>ATGTTGT</b>	GTGGAATTGTGAG		
lacUV5		<b>TTTACA</b>	CTTTATGCTTCCG	GCTCGT	<b>ATAATGT</b>	GTGGAATTGTGAG		
tac		<b>TTGACA</b>	ATTAATCATCCG	GCTCGT	<b>ATAATGT</b>	GTGGAATTGTGAG		
$\lambda P_R$	GATAAATATCTAACACCGTGCGTG	<b>TTGACT</b>	ATTTTACCTCTG	GCGGTG	<b>ATAATGG</b>	TTGCA TGTACTAA		
T7A1	ATTTAAAATTTATCAAAAAGAGTA	<b>TTGACT</b>	TAAAGTCTAACC	<b>TATAGG</b>	<b>ATACTTA</b>	CAGCCATCGAGAG		
rrbB P1	CAGAAAATTTTAAAAATTCCTC	<b>TTGTCA</b>	GGCCGAATAA	C <sup>T</sup> CCCT	<b>ATAATGC</b>	GCCACCACTGACA		
$\lambda$ Pre	TCTCGATTCTAGAGCCTCGTTGC	<b>GTTTGT</b>	TTGCACGAACCA	<b>TATGTA</b>	<b>AGTATTT</b>	CCTTAGATAACAA		
galP1	ATTCCACTAATTTATTCCATGTCA	CAC <sup>T</sup> TT	TCGCATCTTTGT	<b>TATGCT</b>	<b>ATGGTTA</b>	TTTCATACCATAA		
PcysG	TGGGTAATAC <sup>T</sup> TTATCAGGTGCCG	<b>TATTC</b>	TGGGATGGGTT	<b>ATGGT</b>	<b>ATGCTAC</b>	GCCGAAGCGAAT		
PompF	TTTATCTTTGTAGCACTTTCACGG	<b>TAGCGA</b>	AACGTTAGTTTG	<b>AATGGA</b>	<b>AAGATGC</b>	CTGCAGACACATA		
PproU	TCATGCCACATTTGCCATCAGGGG	<b>TTGCC</b>	CAGATTCTCAG	<b>TATGTT</b>	<b>AGGGTAG</b>	AAAAAAGTGACTA		
PtyrT	AGTTCA <sup>T</sup> TTTTCTCAACGTAACAC	<b>TTTACA</b>	GCGGCGCGTCA	<b>TTTGAT</b>	<b>ATGATGC</b>	GCCCCGCTTCCCG		
PuvrA	CGGGTAATGCAT <sup>T</sup> CCAATACTGTA	<b>TATTC</b>	TTCAGGTCAATT	<b>TGTGTC</b>	<b>ATAATTA</b>	ACCGTTGTGTGATC		
Pminor	GTGACCCAATAATGTGGGATAACA	<b>TTGAAA</b>	AGATTAAAGAAA	<b>TATGGG</b>	<b>AAACTC</b>	TGGAAATCCGGG		
gapAP1	TCGCAATGATTGACACGATTCCGC	<b>TTGACG</b>	CTGCCTAAGGTT	<b>TTGTGA</b>	<b>ATTTTAC</b>	AGGCACCTTTTTA		
PcspA	TCATAAATATGAAAAATAATTGTG	<b>TTGCAT</b>	CACCCGCCAATG	<b>CGTGCC</b>	<b>TTAATGC</b>	ACATCAACGGTTT		

Binding

Isomerization

**Figure 3.** Sequences of various  $\sigma^{70}$ -dependent promoters. Consensus sequences of DNA elements are shown at the top in red. (Sequences are given 5' 3' of the nontemplate (top) strand; W = A or T; R = A or G; Y = C or T; N = any nucleotide) *E. coli* promoters are from the *E. coli* genome (accession number NC000913): Plac, 365627-365559; rrbP1, 4164330-4164396; galP1, 791363-791296; PcysG, 3495716-3495783; PompF, 986374-986307; PproU, 2802718-2802784; PtyrT, 1286947-1286881; PuvrA, 4272014-4271947; gapAP1, 1860700-1860767; PcspA, 3717854-3717919. The bp changes present in the Plac derivatives lacUV5 (Stefano et al. 1980) and Ptac (Mulligan et al. 1985) are indicated.  $\lambda$  promoters are from the  $\lambda$  genome (accession number J02459.1):  $P_R$ , 37964-38031; Pre, 38402-38335. Pminor is from the T4 genome (accession number AF158101.6; positions 23736-23670). T7A1 is from the T7 genome (accession number AY264774.1; positions 438-505). (It should be noted that among various references, promoter sequences with the same name can differ at the farthest upstream and downstream portions of the sequences because of differences in the length of DNA cloned for the particular study.) Matches to consensus sequences within the various promoter sequences are given in red. For the -5 contact, the preferred base is a G (shown in red) and the highly nonpreferred base is a C (shown in blue). The involvement of portions of promoter DNA with earlier steps of transcription initiation (polymerase binding) or later steps (polymerase isomerization) is indicated at the bottom.

promoters as well as a weaker consensus of  $^{-17}\text{trTGn}^{-13}$  (Fig. 3) ((Harley and Reynolds, 1987; Hawley and McClure, 1983; Lissner and Margalit, 1993; Lissner and Margalit, 1994; Siebenlist and Gilbert, 1980; Siebenlist et al. 1980); reviewed in (Campbell et al. 2002; Gruber and Gross, 2003; Murakami et al. 2002a)). Subsequent genetic, biochemical, and structural studies have demonstrated that specific residues within  $\sigma^{70}$  interact both directly and indirectly with bases within these consensus elements (Fig. 1).

### Recognition of the -35 element

For the -35 element, specific substitutions within  $\sigma^{70}$  region 4.2 have been shown to relax the specificity of polymerase for bps at positions -33

(E585Q, R588H) (Gardella et al. 1989; Keener and Nomura, 1993) and -31 (R584C) (Gregory et al. 2005; Siegele et al. 1989) (Fig. 1A). In addition, *in vitro*, polymerase containing the  $\sigma^{70}$  R584C or R588H mutation shows the same relaxed specificity (Makela and Hinton, unpublished), and polymerase with an  $\sigma^{70}$  R584A substitution prefers a -31 A:T bp rather than the wild type C:G (Gregory et al. 2005). These results have suggested that  $\sigma^{70}$  residues R584, E585, R588 interact directly with base determinants at these positions. The structure of *Thermus aquaticus*  $\sigma$  region 4 in a complex with -35 DNA element DNA has revealed an extensive interaction between 11 residues in  $\sigma^{70}$  region 4 and 9 bp of the -35 element DNA, either via specific contact of base determinants, water mediated contact,

or contact with the DNA backbone (Campbell et al. 2002). Consistent with the previous work, the specific contacts with template base determinants are seen at positions  $-31$  and  $-33$  with residues R584 and E585, respectively. These residues lie within the second helix of a classic DNA binding helix-turn-helix motif. The structure does not show a specific contact between R588 and a  $-35$  bp element, but suggests that the position of R588 could affect E585, (Campbell et al. 2002) or perhaps it makes direct contact in the context of the entire polymerase.

### Recognition of the $-10$ element

Although the  $-10$  element ( $^{-12}\text{TATAAT}^{-7}$ ) is widely accepted as being crucial for promoter recognition, the  $-12$  T:A position is the only base pair for which suppressor amino acids have been found. Both Q437H (Waldburger et al. 1990) and T440I (Siegele et al. 1989) within  $\sigma^{70}$  region 2.4 suppress a T:A to C:G mutation at position  $-12$  (Fig. 1A). The  $6.5 \text{ \AA}$  resolution structure of *T. aquaticus* polymerase with DNA (Murakami et al. 2002a) is consistent with this work. It shows residues that correspond to Q437 and T440 in *E. coli*  $\sigma^{70}$  positioned on the same side of an  $\alpha$  helix facing the major groove at position  $-12$ . Another study has shown that mutations at  $\sigma^{70}$  residues W434, R436, R441, or R451 can impair binding of polymerase to a duplex promoter sequence from  $-41$  to  $-12$  (Fenton et al. 2000), suggesting that residues both in  $\sigma^{70}$  region 2.4 and in region 2.3 can influence binding to position  $-12$ . Recent work using polymerase with quadruple mutations in region 2.3 also argues that this region of  $\sigma^{70}$  may contact portions of the  $-10$  element while it is in ds form (Cook and deHaseth, 2007).

Further information about  $\sigma^{70}$  recognition of ds DNA within the  $-10$  element has been obtained using a portion of free  $\sigma^{70}$ . Although full length  $\sigma^{70}$  does not bind DNA, a weak interaction of  $\sigma^{70}$  lacking region 1.1 with ds DNA can be detected using a nitrocellulose filter binding assay (Dombroski et al. 1992). With this assay, Dombroski (Dombroski, 1997) demonstrated that within the  $-10$  element, only bp mutations at positions  $-12$ ,  $-11$ , or  $-10$  were deleterious for binding by a  $\sigma^{70}$  peptide containing regions 2 through 4, suggesting again that it is the upstream portion of this element that is specifically recognized as ds DNA.

### Recognition of the $^{-15}\text{TGn}^{-13}$ element

A third consensus element recognized by  $\sigma^{70}$  is the  $^{-15}\text{TGn}^{-13}$  motif, also called the extended  $-10$  sequence (Keilty and Rosenberg, 1987). Optimal activity requires the presence of both the  $-15$  T:A and  $-14$  G:C; although depending on the promoter, the presence of just one of these bp can be advantageous for promoter activity even when the other is mutated (Burr et al. 2000; Grana et al. 1985; Keilty and Rosenberg, 1987; Mitchell et al. 2003). Biochemistry and suppressor genetics, including comprehensive alanine mutagenesis throughout region 3 and *in vitro* transcription assays using mutant  $\sigma^{70}$  proteins, have indicated that residues E458 and H455 within  $\sigma^{70}$  region 3.0 interact with the  $-14$  G:C base pair (Barne et al. 1997; Bown et al. 1999; Sanderson et al. 2003) (Fig. 1A). Glutamic acid residues are known to interact with a cytosine determinant, such as the interaction of  $\sigma^{70}$  region 4 E585 with the C of the G:C bp at position  $-33$ , suggesting that E458 may contact the C base determinant of the G:C bp at position  $-14$  (Barne et al. 1997). Furthermore, the *T. aquaticus* polymerase/promoter structure (Murakami et al. 2002a) shows the residues analogous to H455 and E458 lying along the surface of the region 3  $\alpha$ -helix and toward the major groove of the extended  $-10$  element. Mutation analyses have also revealed that residues 455 and 458 in region 2.4 influence interaction of polymerase with the TGn motif. Thus, it has been suggested that the two  $\alpha$ -helices,  $\sigma^{70}$  regions 2.4 and 3, comprise a pincer for the ds binding element of the TGn motif (Sanderson et al. 2003). As such, regions 2.4 and 3 together would constitute a recognition domain for a ds binding promoter element  $^{-15}\text{TgnT}^{-12}$  that includes both the TGn and the T at position  $-12$ .

In gram-positive bacterial promoters, the extended  $-10$  motif extends to  $-17$ , with a consensus sequence of  $^{-17}\text{TRTG}^{-14}$  (Helmann, 1995; Moran et al. 1982; Voskuil and Chambliss, 2002). For *E. coli* polymerase, the specific sequence at  $-17$  and  $-16$  can affect the level of transcription (Burr et al. 2000; Mitchell et al. 2003), but there is no evidence that these positions are contacted directly.

### Recognition of the UP elements by the C-terminal domains of the $\alpha$ subunits

An early list of  $\sigma^{70}$ -dependent promoter sequences revealed an A/T rich sequence upstream of the  $-35$

element that was present enough to seem significant (Hawley and McClure, 1983). In addition, sequences within this region were found to significantly enhance transcription from some promoters (Chan and Busby, 1989) (also (Ross et al. 1993) and references therein). Using the ribosomal promoter *rrnB* P1, Ross and co-workers demonstrated that the region from  $-40$  to  $-60$ , termed an UP element, was specifically bound by the C-terminal domain of the  $\alpha$ -subunits of polymerase (Gaal et al. 1996; Ross et al. 1993). Subsequent work has shown that 2 subsites of A/T rich sequences, a proximal UP element ( $-45$  to  $-37$ ) and a distal UP element ( $-58$  to  $-45$ ) (Estrem et al. 1998) are contacted through minor groove interactions with the DNA (Naryshkin et al. 2000; Ross et al. 2001) (Fig. 3, numbering is relative to the  $-35$  element of  $^{-35}\text{TGACA}^{-30}$ ). Each element can be contacted by one of the two  $\alpha$  subunits present in polymerase, and whether the proximal, distal or both elements are contacted can affect promoter recognition and activity (Cellai et al. 2007; Typas and Hengge, 2005). Although a specific sequence is preferred for binding, the  $\alpha$  subunits will also interact nonspecifically with this region of DNA (Burns et al. 1999; Ross and Gourse, 2005).

## Interaction of Polymerase with Single-Stranded DNA Elements

### Interaction of $\sigma^{70}$ region 2.4 with an unpaired base at position $-12$

Although the T:A bp at position  $-12$  is recognized by  $\sigma^{70}$  region 2.4 within the context of ds DNA, several lines of evidence indicate that the nontemplate  $-12$  T may also be recognized when it is ss. Marr and Roberts (Marr and Roberts, 1997) demonstrated that RNA polymerase specifically recognizes and binds to a ss oligomer containing the nontemplate sequence of the  $-10$  element and that a  $-12$  T to C mutation within this oligomer is suppressed by the same  $\sigma^{70}$  substitution, Q437H, that suppresses the  $-12$  T:A to C:G bp change in ds promoter DNA. Recognition of the  $-12$  nucleotide within ss DNA has also been inferred from competition binding experiments which have demonstrated the specificity of polymerase for a  $-12$  T (Qiu and Helmann, 1999). Finally, the sensitivity of the nontemplate T at  $-12$  to  $\text{KMnO}_4$ , which will react with unpaired thymines and to a lesser extent cytosines, seems to depend on the particular

promoter and conditions (Davis et al. 2007a; Hook-Barnard et al. 2006; Kainz and Roberts, 1992; Lim et al. 2001; Nguyen and Burgess, 1997; Sasse-Dwight and Gralla, 1989; Sasse-Dwight and Gralla, 1991; Thouvenot et al. 2004). Thus, whether an interaction between  $\sigma^{70}$  and position  $-12$  in the context of ss DNA is crucial for the formation of a productive polymerase/promoter complex is not known.

### Interaction of $\sigma^{70}$ region 2.3 with positions $-11$ through $-7$ of the nontemplate strand

A large body of work, detailed below, has demonstrated that polymerase recognizes the  $-10$  element positions  $^{-11}\text{ATAAT}^{-7}$  as ss nontemplate DNA formed at the transcription bubble surrounding the start of transcription. In most of these studies, the A at  $-11$ , thought to nucleate strand opening, and the nearly invariant T at  $-7$  have been shown to be particularly important, with lesser and varying contributions from the bases at positions  $-10$  through  $-8$ . In many cases, these conclusions have been driven by EMSAs of polymerase and DNA performed in the absence and presence of a challenge, such as heparin. This analysis is meant to distinguish between two types of binding: 1) unstable binding of polymerase to fully duplex DNA, which should only be seen in the absence of the competitor challenge, and 2) the stable, competitor resistant complex. For most promoters, the unstable form is thought to be at an early step in initiation, like  $\text{RPc}$  or  $\text{I}_1$ . The stable form is presumed to be at a late step, such as  $\text{RPo}$ , in which the DNA is unwound from position  $-11$  downstream and conformational changes have resulted in the isomerization of the polymerase. In some cases, the presence of the transcription bubble has been confirmed by  $\text{KMnO}_4$  footprinting.

The first work to demonstrate that polymerase specifically recognizes only the nontemplate strand of the  $-10$  element was reported by Roberts and Roberts (Roberts and Roberts, 1996). These authors investigated the ability of polymerase to form an open complex with promoter DNA containing mismatched nucleotides at the crucial  $-12$ ,  $-11$ , and  $-7$  positions on either the non-template or template strand of the  $-10$  element. To do this analysis they performed single round transcription assays after allowing the polymerase and DNA to form heparin-resistant complexes. Mutations away from

consensus on the nontemplate strand at each of these three positions were deleterious. In contrast, mutations of template strand bases at these positions were inconsequential. EMSA analysis demonstrated that polymerase will bind a ss oligomer that contains the nontemplate sequence of the  $-10$  element (Marr and Roberts, 1997; Qiu and Helmann, 1999). In these cases, binding specificity, which was determined by competition of the complex with a mutant ss oligomer, showed that mutations away from the  $-10$  consensus sequence impaired binding. These studies confirmed the importance of the nucleotide identity on the nontemplate strand.

Contact between polymerase and the nontemplate strand positions  $-11$  to  $-7$  has also been inferred in studies using fork template DNA substrates. Fork junction templates are ds DNAs with a  $3'$  extension on the nontemplate strand that includes part or all of the  $-10$  element. One set of experiments has used  $\sigma^{70}$  labeled with (Eu<sup>3+</sup>)DTPA-AMCA-maleimide to assay luminescence resonance energy transfer to fork junction DNA having a  $3'$  nontemplate extension of  $-11$  to  $-4$  and labeled with the fluorophore Cy5 (Matlock and Heyduk, 2000). Binding, as assayed by the energy transfer, was measured in the presence of a non-labeled DNA containing a specific mutation. This work again indicated that the nontemplate  $-11$  A and  $-7$  T are particularly important for binding. In addition, substitutions of the  $-11$  A with a series of adenine analogs argued that the N1 position of the  $-11$  adenine is crucial for polymerase contact. This suggests that the interaction of polymerase with this base determinant could help disrupt the  $-11$ A:T bp, nucleating the strand separation that is needed for the open complex.

A series of studies from the Gralla lab has investigated the binding of polymerase to fork junction DNAs using EMSAs with or without a heparin challenge. Experiments with duplex promoter DNA from  $-41$  to  $-12$  with a nontemplate ss extension from  $-11$  to  $-7$  showed that specific base mutations within the nontemplate extension impair the formation of the stable, isomerized complex and again demonstrated the importance of the sequence of the nontemplate strand, especially at positions  $-11$  and  $-7$ , for stable binding (Fenton and Gralla, 2001). Furthermore, even a fork junction substrate containing a single, nontemplate base  $3'$  extension at position  $-11$  can form a heparin resistant complex if the  $-11$  nucleotide is the consensus A (Guo and Gralla, 1998) or the adenine analog 2-aminopurine, which has an amino group at C2 rather than C6 (Tsujikawa

et al. 2002). These results suggested that the interaction of polymerase with the nontemplate  $-11$  A is sufficient for generating a stable polymerase/promoter complex, in which the polymerase has presumably isomerized.

Interestingly, the effect of a 2-aminopurine at  $-11$  in duplex promoter DNA depends on the particular promoter context. In one experiment using the natural Pgal promoters, which have less than ideal promoter elements, this substitution completely inhibited strand separation (Lim et al. 2001), while in another study using an ideal promoter (a perfect  $-35$  element, a TGn motif, and a perfect  $-10$  element) strand separation was unaffected (Tsujikawa et al. 2002). It has been proposed that a flipping out of the  $-11$  A nucleates strand separation, which then propagates downstream (Helmann and Chamberlin, 1988; Helmann and deHaseh, 1999; Young et al. 2004). In this model, polymerase could facilitate the flip or capture the  $-11$  A after the flip. Thus, the ability of polymerase to engage in this process when a 2-aminopurine is present at position  $-11$  may be influenced by the strength of the polymerase interactions with other promoter elements.

Clusters of aromatic residues can signify a ss DNA or RNA binding motif (Bochkarev et al. 1995; Burd and Dreyfuss, 1994; Helmann and Chamberlin, 1988; Shamoo et al. 1995). Structural analysis of  $\sigma^{70}$  region 2 (Malhotra et al. 1996) suggested that several aromatic residues lying along an  $\alpha$ -helix face in region 2.3 might interact with ss DNA in the transcription bubble. Multiple alanine substitutions at Y425, F427, Y430, W433, and/or W434 in various combinations render polymerase significantly impaired in open, but not closed, complex formation at the lambda promoters P<sub>R</sub> and P<sub>RM</sub> (Cook and deHaseh, 2007; Panaghie et al. 2000). Earlier work has shown that substitutions at corresponding residues within region 2.3 of the primary  $\sigma$  of *B. subtilis* are defective for binding to an oligonucleotide containing the nontemplate  $-10$  sequence (Juang and Helmann, 1994). However, it should be noted that despite the extensive analyses indicating that the  $^{-11}$ ATAAT $^{-7}$  sequence is highly conserved, that the  $-11$  A appears to be crucial for starting the process of strand separation, and that various aromatic residues in region 2.3 are important for ss DNA binding and polymerase function, definitive residue/base interactions have not been identified [(Fenton and Gralla, 2003; Schroeder et al. 2007) and references therein].



## Recognition of nontemplate G at position -5

Studies have indicated that nucleotides downstream of the -10 element are contacted by polymerase, and the role of  $\sigma^{70}$  region 1.2 in these interactions has been demonstrated (Feklistov et al. 2006; Haugen et al. 2006; Zenkin et al. 2007). At the ribosomal promoter *rrnBP1*, RPo is unstable, and this short half-life is a crucial factor in transcriptional regulation in response to growth conditions and nutrient availability (Paul et al. 2004). Experiments investigating this instability have revealed that the nucleotide on the nontemplate strand at position -5 can directly interact with  $\sigma^{70}$  region 1.2, stabilizing RPo (Haugen et al. 2006) (Fig. 1A). The identity of the nontemplate base two bases downstream of the -10 element (-5 at most promoters, but -7 at *rrnBP1*) is crucial in determining RPo half-life. For example, changing the nontemplate -5 nucleotide from a C to a G results in a dramatic increase in half-life at *rrnBP1*. Conversely, substituting the base at -5 for a C at other promoters (*Pgal*,  $\lambda P_R$ ,  $\lambda P_L$ ) significantly reduces RPo half-life. The -5G was shown to crosslink to  $\sigma^{70}$  region 1.2, indicating a direct contact between this region and downstream DNA. Interestingly, the crosslink between region 1.2 and the -5 base was strongest when the  $\sigma^{70}$  region 1.1 was missing. This result suggests that either region 1.1 blocks access to the crosslinking site, or that in the absence of region 1.1 the contact between region 1.2 and -5G is longer lived. The latter interpretation implies that the region 1.2/DNA interaction is transient and may be modulated by the movement of 1.1.

The strength of the region 1.2/-5 nucleotide interaction appears to play an important role in determining the inherent stability of RPo. The complex half-life can then be further influenced by cellular factors, such as NTP concentration, the small effector molecule ppGpp, and the protein factor DksA (Gralla, 2005; Rutherford et al. 2007). Thus, the half-life of RPo can be a major determinant of transcriptional activity and is the central conveyor of regulation at *rrnBP1*.

## Importance of Spacers between Elements

### The spacer between the -35 and -10 elements

In addition to the sequence determinants, RNA polymerase also has specific spacer length

requirements for promoter recognition. An early analysis (Hawley and McClure, 1983) of 112 characterized bacterial and phage promoters found that 50% had a spacer of 17 bp, while 20% had a spacer of 16 bp and another 20% had a spacer of 18 bp. A more recent analysis indicated that in *E. coli* 44% of promoters have a spacer length of 17 bp, although promoters with a TGn motif are more likely to have a spacer of 18 or more bp (Mitchell et al. 2003). It is well established that promoters with a 17 bp spacer yield higher levels of transcription than otherwise identical promoters that have spacers of 16 or 18 bp (Aoyama et al. 1983; Mulligan et al. 1985; Stefano and Gralla, 1982).

The preferred spacer length between the -10 and -35 promoter elements is dictated by the distance between recognition domains within the polymerase structure. This distance is set in part, by the interaction of  $\sigma^{70}$  region 4 with a structure in core called the  $\beta$ -flap, which positions region 4 to correctly make contact with the -35 sequence while region 2 interacts with the -10 sequences (Kuznedelov et al. 2002; Murakami et al. 2002b; Vassylyev et al. 2002). The role of the sigma subunit in perceiving the spacer length was first demonstrated using free  $\sigma^{70}$  polypeptides and the *tac* promoter with spacer lengths ranging from 11 to 26 bp (Dombroski et al. 1996).  $\sigma^{70}$  peptides containing either region 2 or region 4 alone bound the promoter, unaffected by spacer length. However, when the polypeptide included regions 2 through 4, the length of the spacer became so important that a change of +/-1 bp rendered the promoter no more recognizable than non-promoter DNA. This work indicated that the sigma regions 4 and 2 simultaneously contact the -35 and the -10 elements. Luminescence Resonance Energy Transfer measurements have also showed that  $\sigma^{70}$  undergoes a conformational change upon binding to core, which alters the distance between regions 2.4 and 4.2 from 34Å apart in free  $\sigma^{70}$  to 50Å in holoenzyme (Callaci et al. 1998; Callaci et al. 1999). The 50Å distance between the recognition domains is conducive to interacting with -10 and -35 elements separated by the preferred 17 bp spacer.

As discussed above, polymerase is known to make contacts just upstream of the -10 element at the  $^{-15}\text{TGn}^{-13}$  motif. Although polymerase does not appear to make specific contacts with the spacer nucleotides upstream of position -15, the sequence of this spacer, in addition to its length, can affect

promoter function (Chan and Busby, 1989; Chan et al. 1990; Liu et al. 2004b; Mellies et al. 1994; Repoila and Gottesman, 2003; Thouvenot et al. 2004; Warne and deHaseh, 1993). For example, changing the spacer from GC- to AT-rich sequences increases promoter activity and also affects regulation at Plac (Liu et al. 2004b) and the *dsrA* promoter, (Repoila and Gottesman, 2003), while individual base substitutions in the spacer affect transcription from *galP1* (Chan and Busby, 1989), *gapAP1* (Thouvenot et al. 2004), and the *proU* promoter (Mellies et al. 1994). Due to the lack of any known specific contacts, it has been assumed that the spacer sequences may affect the structure or trajectory of the promoter DNA.

### The spacer between the $-10$ element and position $+1$ , the start site of transcription

The distance between the  $-10$  element and transcriptional start site, which is constrained by the polymerase structure (Murakami et al. 2002b; Vassilyev et al. 2002), can also affect transcription levels. The preference of starting nucleotide and location relative to the  $-10$  has been investigated in detail at the *pyrC* promoter (Liu and Turnbough, 1994). This study found that the preferred  $+1$  nucleotide is  $A \geq G > T >> C$  and is typically located  $7 > 8 > 9$  nucleotides downstream of the  $-7$  base of the  $-10$  element (Fig. 3). Similar results have been reported for PlacUV5 (Jeong and Kang, 1994) and the *gal* promoters *galP1* and *galP2* (Lewis and Adhya, 2004). The conclusions of these studies are consistent with those derived from a computer analysis (O'Neill, 1989) of a previous promoter set (Hawley and McClure, 1983).

### Mix and Match Elements for Binding of ds DNA Elements

#### The $-35/10$ promoter

Promoters containing good matches to the  $-35$  and  $-10$  elements represent the classic group of *E. coli* promoters. Although conservation of each base pair varies somewhat within these elements, depending upon the promoters included in the data set, the relative importance of the bases is fairly consistent. The  $-12$  T:A (79%),  $-11$  A:T (87%), and  $-7$  T:A (90%) bases are the most highly conserved within the  $-10$  element (Lisser and Margalit, 1993; Mitchell

et al. 2003). Within the  $-35$  region, the  $-35$  T:A,  $-34$  T:A, and  $-33$  G:C are each found at  $>69\%$  of promoters (Lisser and Margalit, 1993). Interestingly, the  $-31$  C:G bp, which is contacted directly on the template strand, is less conserved ( $\sim 50\%$ ) (Lisser and Margalit, 1993). As expected, a very strong correlation between the consensus sequence and promoter activity has been demonstrated [reviewed in (Hawley and McClure, 1983)].

Several natural promoters have been studied as models for the  $-10/-35$  class including T7A1,  $\lambda$  P<sub>R</sub>, and Plac (Fig. 3). Plac is perhaps the best characterized of all promoters and has served as a paradigm for the  $-10/-35$  promoter class (Borukhov and Lee, 2005; Reznikoff, 1992; Silverstone et al. 1970; Stefano and Gralla, 1982). However, Plac deviates from consensus in several ways. The  $-35$  element is TTTACA, the  $-10$  element is TATGTT, and there is an 18 bp spacer (Fig. 3). As a result, Plac transcription is at fairly low levels until activated by CRP (c-AMP receptor protein), in response to glucose level. Several mutants of Plac, which make the  $-10$  and/or the  $-35$  elements more consensus, are CRP-independent (Arditti et al. 1968; Silverstone et al. 1970). Subsequent studies have demonstrated the importance of the  $-10$ , the  $-35$ , and the spacer regions for these promoters (Ackerson and Gralla, 1983; Chan and Lebowitz, 1990; Liu et al. 2004b; Makoff and Ozer, 1991; Mandrecki and Reznikoff, 1982). One Plac derivative, lacUV5 (Fig. 3), has been studied extensively by genetic, biochemical, and structural methods (Stefano et al. 1980; Stefano and Gralla, 1982) [reviewed in (Borukhov and Lee, 2005; Siebenlist et al. 1980)]. The lacUV5 mutation, which changes the Plac  $-10$  element (TATGTT) to the consensus sequence (TATAAT), increases transcription and renders lacUV5 independent of the CRP activator (Arditti et al. 1968; Silverstone et al. 1970). Comparative studies have demonstrated that, although lacUV5 is a stronger promoter, whose kinetics of RPo formation differ from Plac, the final RPo complex is quite similar to that of Plac (Meiklejohn and Gralla, 1989) (Fenton and Gralla, 2001). Footprinting of RPo at these promoters indicates that the transcription bubble is from  $-12/-11$  to  $+4$ , and the DNA is protected from cleavage from  $-55$  to  $+20$  for both promoters (Sasse-Dwight and Gralla, 1989; Spassky et al. 1984; Spassky et al. 1985).

Despite their classification as  $-35/-10$  promoters, many promoters in this class have other

sequences that contribute to promoter activity. For instance, AT-rich sequences upstream of  $\lambda P_R$  and T7A1 significantly increase open complex formation (Cellai et al. 2007; Davis et al. 2005; Sclavi et al. 2005). Plac (and therefore lacUV5) has a T:A bp at position  $-15$  that influences transcription levels (Liu et al. 2004a; Munson et al. 1984). Furthermore, the T7A1 promoter has a  $-15$  T:A and  $\lambda P_R$  has a  $-14$  G:C, which may compensate for their noncanonical bp at the highly conserved  $-12$  position. It is well established that the effect of promoter mutations is dependent upon context and that strong polymerase-DNA contacts can compensate for weaker sites of interaction (Grana et al. 1988; Michalowski et al. 2004; Miroslavova and Busby, 2006; Moyle et al. 1991), and simply assigning a promoter by its matches to the  $-10$  and  $-35$  elements can fail to identify active promoters *in vivo* (Kawano et al. 2005). Thus, polymerase-promoter contacts outside of the  $-10$  and  $-35$  elements influence the activity of classic  $-10/-35$  promoters and these effects should be given consideration when interpreting data.

### The TGn/-10 promoter (extended $-10$ promoter)

An early compilation and analysis of promoters indicated that there was some preference for the TG sequence upstream of  $-10$  (Hawley and McClure, 1983). Genetic and biochemical data also suggested the importance of polymerase-DNA contact in this region of DNA (Berman and Landy, 1979; Busby et al. 1984; Grana et al. 1985; Ponnambalam et al. 1986; Siebenlist and Gilbert, 1980). More recently, 554 promoters of *E. coli*, whose  $+1$  start sites had been determined and whose  $-10$  elements had been identified, were analyzed (Burr et al. 2000; Mitchell et al. 2003). This work revealed that  $^{-15}\text{TGn}^{-13}$  is present in 20% of *E. coli* promoters. In addition, 43% of the 554 promoters have a G:C at position  $-14$ . Another 246 promoters were eliminated from the analysis because they were not well defined, leaving the possibility that the occurrence of  $-15\text{T}$  or  $-14\text{G}$  is even higher. This could be the case since promoters that deviate significantly from consensus in the  $-10$  and  $-35$  elements are more likely to have a TGn motif and may be less well-characterized.

Investigations into the importance of the  $^{-15}\text{TGn}^{-13}$  motif have shown that this sequence can compensate for a poor or missing  $-35$  element.

One such TGn/ $-10$  promoter is  $\lambda\text{Pre}$  (Fig. 3), which has a very poor match to the  $\sigma^{70}$   $-35$  element, a noncanonical  $-10$  element ( $^{-12}\text{AAGTAT}^{-7}$ ), and requires an activator, CII, for detectable promoter function (Keilty and Rosenberg, 1987). The mutant derivative  $\lambda\text{Pre}^*$  was created with a consensus  $-10$  element, and although it still has no recognizable  $-35$  element,  $\lambda\text{Pre}^*$  is independent of CII (Keilty and Rosenberg, 1987). Both  $\lambda\text{Pre}$  and  $\lambda\text{Pre}^*$  require the TGn motif, and neither promoter requires specific sequences within the  $-35$  region (Keilty and Rosenberg, 1987). In fact, region 4.2 of  $\sigma^{70}$ , which contacts the  $-35$  element, can be removed entirely without eliminating transcription from  $\lambda\text{Pre}^*$  (Kumar et al. 1993). In contrast, a  $\lambda\text{Pre}^*$  mutant, in which the  $^{-15}\text{TG}^{-14}$  was changed to  $^{-15}\text{CC}^{-14}$ , and the promoter Pcons, which has consensus  $-35$  and  $-10$ , but no TGn, have little activity without  $\sigma^{70}$  region 4.2 (Kumar et al. 1993).

Another well-studied member of the TGn/ $-10$  class is galP1 (Fig. 3). galP1 has an imperfect extended  $-10$  sequence (TGnTATGGT) and no recognizable homology to the  $\sigma^{70}$   $-35$  element. As with  $\lambda\text{Pre}$ , the  $-35$  region of galP1 is inessential for transcription (Ponnambalam et al. 1986), whereas the TGn is required (Busby et al. 1984). In addition, galP1 transcription is dependent on A/T rich upstream sequences or CRP for activation (Chan and Busby, 1989). Perfecting the  $-10$  element makes galP1 independent of CRP; however, this derivative still requires TGn (Kuhnke et al. 1987; Kumar et al. 1993; Kumar et al. 1994). In contrast, creation of an ideal  $-35$  element allows mutation of the TGn, yielding a functional and CRP independent promoter (Chan and Busby, 1989; Chan et al. 1990).

Footprinting analyses of the open complex at galP1 have indicated that the promoter region is protected from  $-55$  to  $+20$ , similar to that observed with the  $-35/-10$  promoters, yet there is less protection of the  $-35$  region (Grimes et al. 1991). The protection upstream from the  $-35$  element to  $-55$  is dependent on the presence of the  $\alpha$ -CTD (Burns et al. 1999). At another TGn promoter, PcysG (Fig. 3), the DNA is well protected from  $-55$  to  $+20$  with an additional weak protection extending upstream to  $-80$ , presumably due to interactions with the UP elements (Belyaeva et al. 1993). These results are consistent with the model that at a TGn/ $-10$  promoter, polymerase uses other contacts to compensate for the lack of the  $\sigma^{70}$  region 4/ $-35$

interaction (Busby et al. 1987). However,  $\text{KMnO}_4$  footprints are the same (-11 to +3) whether using galP1 or galPcon, a derivative of galP1 in which the sequences upstream of -12 have been replaced, mutating the TGn and inserting a perfect -35 element (Grimes et al. 1991). Thus, a change in upstream contacts does not appear to alter the final transcription bubble.

Unlike  $\lambda\text{Pre}$ , galP1, and PcysG, other promoters, such as those for proU (Mellies et al. 1994) (Fig. 3) and tyrT (Berman and Landy, 1979) (Fig. 3) have been characterized as extended -10 promoters, but they have reasonable matches to the -35 element, and for the tyrT promoter, this element is essential for activity (Lamond and Travers, 1983). In addition, the TGn promoter for uvrA (Fig. 3) has a relatively poor -35 element, yet this element still enhances transcription (Backendorf et al. 1983). This is also the case with the ompF promoter (Fig. 3), which has a weakly conserved -10 region (AAAGAT) and requires the TGn motif for function (Mitchell et al. 2003; Taylor et al. 1985). ompF is activated by OmpR, whose binding site overlaps the -35 region, yet transcription is also influenced by the -35 element in the absence of OmpR (Inokuchi et al. 1984; Mitchell et al. 2003; Taylor et al. 1985). Mutation of the -12 A:T to T:A makes transcription independent of the -35 region and OmpR (Dairi et al. 1985; Ozawa et al. 1987). In contrast, a ompF-tet hybrid, which has the -12 A:T to T:A change, but also has a mutated TGn, is dependent upon OmpR activation (Dairi et al. 1985). Thus, it appears that below a minimum number of polymerase contacts, i.e. a threshold, the promoter becomes activator dependent.

In summary, although extended -10 promoters have frequently been characterized as not requiring sequences upstream of position -15, this is a misleading generalization. Most of the examples discussed above require or benefit from -35 and/or UP element contacts. In addition, the majority of TGn promoters have a -35 region with at least a 3 out of 6 match to the -35 consensus sequence (Mitchell et al. 2003). As observed by Kumar et al. (Kumar et al. 1993), "there are no clear examples of "pure" extended -10 promoters lacking any -35 consensus or activator protein-binding site" (Kumar et al. 1993 p. 415). Furthermore, TGn promoters are as likely to have an imperfect -10 element as a nonconsensus -35 element (Mitchell et al. 2003). Thus, the TGn motif can compensate

for poor or missing -35 contacts, but, as discussed below, it may also compensate for weak -10 interactions.

## The -35/TGn promoter

Recently, a new class of promoters has been defined, which demonstrates the role of TGn contacts in compensating for a weak -10 element. As discussed above, many promoters identified as extended -10 promoters require upstream contacts for full promoter function. It is also clear that the TGn can compensate for a -10 element with a poor match to consensus. In fact, as determined by Mitchell et al., 38% of TGn promoters have 3 or less matches within the -10 element consensus sequence, compared to 24% of non-TGn promoters (Mitchell et al. 2003).

One well-characterized -35/TGn promoter is gapAP1 (Thouvenot et al. 2004) (Fig. 3). Mutagenesis and *in vitro* transcription have demonstrated that both the -35 and TGn sequences are required for function at this promoter. These elements compensate for weak -10 interactions; when the noncanonical -10 element ( $^{-12}\text{AATTTT}^{-7}$ ) was perfected, creating the sequence  $^{-15}\text{TGnTATAAT}^{-7}$ , the -35 element was no longer necessary (Thouvenot et al. 2004). This was also the case for another -35/TGn promoter, Pminor (Hook-Barnard et al. 2006; Vuthoori et al. 2001) (Fig. 3). Both the -35 element and the TGn of Pminor are required and compensate for the poor -10 element (GAAAAC) (Hook-Barnard et al. 2006). Mutation of the -14 G:C to A:T eliminates expression, but the combination of this mutation with a -12 G:C to T:A change results in the same level of transcription as wild type Pminor. Thus the -14 G:C contact compensates for the poor -10 element and specifically for the lack of contact at -12 G:C.

Chemical probing of gapAP1 demonstrated that the polymerase-promoter interactions occur throughout the promoter region including the TGn and -35 elements. The transcription bubble was determined by  $\text{KMnO}_4$  footprinting to extend from -12 to +3. (The observation that the -12 is assessable for  $\text{KMnO}_4$  reactivity in RPo at gapAP1 is not unique to this promoter class, since this observation was also made at the -35/-10 promoter  $\lambda\text{P}_R$  (Suh et al. 1993)). In addition, both the Pminor promoter and a Pminor derivative with -12 T:A yielded a transcription bubble from -11 to +3. DNase I footprinting of Pminor shows protection from -55 to +25;

this is also observed with a Pminor derivative in which the  $-10$  element has been perfected (Hook-Barnard, unpublished). Thus, the final open complex is essentially the same for all promoter types.

In addition to these well-defined  $-35/\text{TGn}$  promoters, there are other examples that have not been recognized as such. For instance, P<sub>cspA</sub> (Tanabe et al. 1992) (Fig. 3) has been described as an extended  $-10$  promoter, because the TGn is required for transcription (Phadtare and Severinov, 2005). However, the  $-35$  sequence (TTGCAT) is a good match to consensus and may affect promoter activity. Moreover, the  $-10$  element (CTTAAT) is nonconsensus at the crucial  $-12$  and  $-11$  positions. Thus, the required TGn is as likely to compensate for the poor  $-10$  element as the  $-35$  element. As discussed in the previous section, many promoters, which are considered part of TGn/ $-10$  class are dependent upon  $-35$  sequences and have poorly functional  $-10$  elements. Thus, assumptions should not be made about promoter/polymerase contacts with other regions simply because a promoter has or even requires a TGn sequence.

### An UP/ $-10$ promoter (?)

Although a UP/ $-10$  promoter has not been identified, the possibility of such a promoter has been revealed by experiments using the bacteriophage T4 protein, AsiA. AsiA binds tightly to  $\sigma^{70}$  region 4, dramatically changing the structure of region 4 and preventing its interaction with the  $-35$  DNA element ((Lambert et al. 2004) reviewed in (Hinton et al. 2005)). Consequently, AsiA significantly inhibits transcription from promoters requiring a  $-35$  element such as the  $-35/-10$  promoter lacUV5 or the  $-35/\text{TGn}$  promoter Pminor. TGn/ $-10$  promoters are not inhibited by AsiA (Colland et al. 1998; Pahari and Chatterji, 1997; Severinova et al. 1998). Many T4 early promoters are very strong, having portions of all four recognition elements (UP/ $-35/\text{TGn}/-10$ ) and as expected, these promoters are much less susceptible to AsiA inhibition (Pene and Uzan, 2000). An analysis of what is needed to impart resistance to AsiA inhibition for one of these promoters has revealed that even if the TGn element is mutated, contact between the  $\alpha$ -CTDs and the UP elements is still sufficient to provide significant resistance to AsiA inhibition, suggesting that in the correct context, an UP/ $-10$  promoter is acceptable for recognition (Orsini et al. 2004). Furthermore, an extended incubation of

AsiA-associated polymerase with lacUV5 or Pminor eventually results in transcriptionally competent RPo complexes at these promoters (Orsini et al. 2001; Pal et al. 2003). Given that the  $\alpha$ -CTDs can make non-specific contacts in the  $-40$  to  $-60$  region, even when recognizable UP element sequences are missing, these results suggest that the extended period of incubation provides the time needed to form open complexes using these less specific contacts in the upstream region.

Another example of promoters functioning with only a  $-10$  element and upstream contacts was demonstrated recently using a set of synthetic promoters (Miroslavova and Busby, 2006). In that study, the authors began with a promoter, which had a  $-10$  element, but none of the other consensus sequences. As expected, expression was essentially zero. In the presence of CRP and an upstream CRP binding site, the promoter became functional even without the addition of other contact sequences, suggesting that upstream contacts plus a  $-10$  element can be sufficient for promoter activity. Thus, although an UP element/ $-10$  promoter has not been described, these examples indicate that such promoters may be present within the repertoire of the *E. coli* genome.

### Another Way to Consider Promoter Types: Binding/Isomerization Elements

As detailed in the first section, the earliest step in promoter recognition and binding is the formation of an unstable polymerase/promoter complex through interaction of polymerase with recognition elements present in ds DNA. Recent analyses of the earliest intermediates leading up to stable complex formation at the strong T7A1 promoter (Sclavi et al. 2005) and the  $\lambda P_R$  promoter (Davis et al. 2007a) have been performed using hydroxyl radical footprinting. In these analyses, contact with the promoter initiates with the farthest upstream regions and proceeds toward the start site of transcription. Thus, the recognition of the specific ds DNA binding elements starts the process that can eventually result in the stable open complex.

Typically, a promoter has been classified depending on how well its sequence matches the consensus sequences that have been observed in  $\sigma^{70}$ -dependent promoters (Harley and Reynolds, 1987; Hawley and McClure, 1983; Shultzaberger et al. 2007) Thus, as described above, promoters

have been identified as  $-35/-10$ , TGn/ $-10$ , or more recently,  $-35$ /TGn promoters. For some promoters, such as the  $-35/-10$  promoter lacUV5, the TGn/ $-10$  promoter galP1, and the  $-35$ /TGn promoter Pminor, these classifications seem reasonable (Fig. 3). Matches to the identified consensus sequences are clear, and biochemical analyses have indicated that the designations actually reflect the regions that are important for the function of the promoter. However, for most promoters this seems like an arbitrary system of classification for several reasons. First, many promoter sequences have a mixture of various potential ds binding elements that could contribute to binding (Mitchell et al. 2003) (Fig. 3). TGn/ $-10$  promoters often have a recognizable  $-35$  element, and this  $-35$  element can contribute to transcriptional activity (Miroslavova and Busby, 2006; Mitchell et al. 2003). In addition, UP elements can improve the activity of a TGn/ $-10$  promoter in the absence of a good  $-35$  element (Miroslavova and Busby, 2006), and improvement of the  $-10$  element of the  $-35$ /TGn promoter Pminor increases its activity (Hook-Barnard et al. 2006). Many promoters that are placed in the  $-35/-10$  class have either a  $-15$  T:A or  $-14$  G:C and it has been shown that having one match to the TGn element can improve activity in some cases (Mitchell et al. 2003). The second reason that the present promoter classification is somewhat misleading is because the  $-10$  element, which is used as one of the classifications, is fundamentally different from the other elements. Accumulated evidence suggests that the  $-12$  bp is primarily a ds recognition element for the  $-10$  region and the  $-11$  to  $-7$  base pairs contribute little to specific duplex binding. Recognition of base determinants in the sequences downstream of position  $-12$  occurs during and after polymerase isomerization. Thus, the  $-10$  element is not a single recognition element. Rather it should be considered as two elements: a  $-10$  binding element (primarily position  $-12$ ) and a  $-10$  melting element (positions  $-11$  to  $-7$ ). Finally, regardless of the promoter classification, the final transcription bubble in RPo appears to be the same. KMnO<sub>4</sub> analyses have indicated that once the open complex is formed, the unpaired thymines always extend from  $-12$  or  $-11$  to around  $+3$  (Davis et al. 2007a; Grimes et al. 1991; Hook-Barnard et al. 2006; Kainz and Roberts, 1992; Lim et al. 2001; Nguyen and Burgess, 1997; Sasse-Dwight and Gralla, 1989;

Sasse-Dwight and Gralla, 1991; Thouvenot et al. 2004). In fact, even a minimal promoter, a short DNA duplex from  $-18$  to  $-5$  that only contains a consensus  $-10$  element (without a TGn motif), will eventually produce a transcription bubble after a long incubation with polymerase (Niedziela-Majka and Heyduk, 2005). Although this is an inefficient process, it illustrates the point that the initial binding contacts do not qualitatively influence the final stable complex. Instead, all the contacts needed for isomerization are contained in the  $-10$  element. Furthermore, other work has shown that a minimal polymerase, containing only  $\sigma^{70}$  regions 2 and 3 and just a portion of core, is capable of forming an open complex with an extended  $-10$  promoter (Young et al. 2004). This finding is also consistent with the idea that minimal polymerase/DNA contacts are acceptable for isomerization. Thus, the strength and number of binding contacts appear to affect the kinetics of recognition and isomerization without substantively affecting the salient features of the final open complex.

Taken together, the data argue that there are 3 promoter locations that can be used for ds binding: UP elements, the  $-35$  element and a  $-15$  element,  $^{-17}\text{tgTGnT}^{-12}$ . (The  $-15$  element is a combination of the TGn element with the  $-12$ T.) Certainly a minimal number of contacts within these ds binding elements are needed for sufficient recognition and binding, but various combinations are absolutely permitted. Thus, the ds binding elements represent a set of mix and match, or as designated by the Busby lab: modular (Miroslavova and Busby, 2006) elements, which appear to work interchangeably for the early steps in transcription initiation. The subsequent steps then rely on interactions with ss elements, the  $^{-11}\text{ATAAT}^{-7}$  and the  $-5$  G, which interact with  $\sigma^{70}$  regions 2.3 and 1.2 respectively.

Interestingly, a fully consensus promoter is undesirable in several ways. First, a number of studies have indicated that having too many contacts actually reduces transcriptional activity for a promoter, presumably because they impede the transition from the open complex to promoter clearance and elongation (Grana et al. 1988; Miroslavova and Busby, 2006). Thus, promoters have evolved to contain the optimal number of contacts such that the promoter not only can be recognized, but will also be released as transcription proceeds. Second, a less than canonical sequence provides the oppor-

tunity for regulation. Promoters that are appropriately responsive to cellular signals provide a selective advantage, and are favored over promoters that are intrinsically high functioning, but unregulated. As we learn more about the steps leading to transcription, we see that a promoter not only supplies information via its consensus sequences, but that the nature of noncanonical sequences may provide insight into the form of regulation required to trigger activity. For instance, promoters with relatively few matches to ds binding elements may remain latent until an activator supplies the additional contacts that strengthen initial binding. In contrast, promoters with mismatches in the  $-11$  to  $-7$  region may require activators, or conditions (temperature, salt, supercoiling) that facilitate DNA melting, or polymerase isomerization. Furthermore, a nonpreferred nucleotide at  $-7$  or  $-5$  may indicate that the RPO is unstable and therefore responsive to modulating factors such as DksA, ppGpp, or NTP concentration (Rutherford et al. 2007). Thus, the ability of a promoter to respond to cellular conditions is absolutely dependent upon its flaws, and the mix and match nature of promoter elements allows a wide variety of ways to arrive at the desired outcome: effective transcription. Consequently, perfect promoters are not biologically relevant.

## Acknowledgements

We are grateful to R. Martin, R. Bonocora, L. Knipling, K. Baxter, and P. Decker for helpful discussions. This research was supported by the Intramural Research Program of the NIH, National Institute of Diabetes and Digestive and Kidney Diseases.

## References

Ackerson, J.W. and Gralla, J.D. 1983. In vivo expression of lac promoter variants with altered  $-10$ ,  $-35$ , and spacer sequences. *Cold Spring Harb. Symp. Quant. Biol.*, 47 Pt 1:473–6.

Aoyama, T., Takanami, M., Ohtsuka, E., Taniyama, Y., Marumoto, R., Sato, H. and Ikehara, M. 1983. Essential structure of *E. coli* promoter: effect of spacer length between the two consensus sequences on promoter function. *Nucleic Acids Res.*, 11:5855–64.

Arditti, R.R., Scaife, J.G. and Beckwith, J.R. 1968. The nature of mutants in the lac promoter region. *J. Mol. Biol.*, 38:421–6.

Backendorf, C., Brandsma, J.A., Kartasova, T. and van de Putte, P. (1983). In vivo regulation of the *uvrA* gene: role of the “ $-10$ ” and “ $-35$ ” promoter regions. *Nucleic Acids Res.*, 11:5795–810.

Barne, K.A., Bown, J.A., Busby, S.J. and Minchin, S.D. 1997. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the ‘extended-10’ motif at promoters. *Embo J.*, 16:4034–40.

Belyaeva, T., Griffiths, L., Minchin, S., Cole, J. and Busby, S. 1993. The *Escherichia coli* *cysG* promoter belongs to the ‘extended-10’ class of bacterial promoters. *Biochem. J.*, 296:851–7.

Berman, M.L. and Landy, A. 1979. Promoter mutations in the transfer RNA gene *tyrT* of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, 76:4303–7.

Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Furey, W. Jr., Edwards, A.M. and Frappier, L. 1995. Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA 1. *Cell.*, 83:39–46.

Borukhov, S. and Lee, J. 2005. RNA polymerase structure and function at lac operon. *C.R. Biol.*, 328:576–87.

Bown, J.A., Owens, J.T., Meares, C.F., Fujita, N., Ishihama, A., Busby, S. J. and Minchin, S.D. (1999). Organization of open complexes at *Escherichia coli* promoters. Location of promoter DNA sites close to region 2.5 of the sigma70 subunit of RNA polymerase. *J. Biol. Chem.*, 274:2263–70.

Brodolin, K., Zenkin, N. and Severinov, K. 2005. Remodeling of the sigma70 subunit non-template DNA strand contacts during the final step of transcription initiation. *J. Mol. Biol.*, 350:930–7.

Browning, D.F. and Busby, S.J. 2004. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, 2:57–65.

Buc, H. and McClure, W.R. 1985. Kinetics of open complex formation between *Escherichia coli* RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry*, 24:2712–23.

Burd, C.G. and Dreyfuss, G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science*, 265:615–21.

Burns, H.D., Ishihama, A. and Minchin, S.D. 1999. Open complex formation during transcription initiation at the *Escherichia coli* galP1 promoter: the role of the RNA polymerase alpha subunit at promoters lacking an UP-element. *Nucleic Acids Res.*, 27:2051–6.

Burr, T., Mitchell, J., Kolb, A., Minchin, S. and Busby, S. 2000. DNA sequence elements located immediately upstream of the  $-10$  hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Res.*, 28:1864–70.

Busby, S., Truelle, N., Spassky, A., Dreyfus, M. and Buc, H. 1984. The selection and characterisation of two novel mutations in the overlapping promoters of the *Escherichia coli* galactose operon. *Gene*, 28:201–9.

Busby, S., Spassky, A. and Chan, B. 1987. RNA polymerase makes important contacts upstream from base pair  $-49$  at the *Escherichia coli* galactose operon P1 promoter. *Gene*, 53:145–52.

Callaci, S., Heyduk, E. and Heyduk, T. 1998. Conformational changes of *Escherichia coli* RNA polymerase sigma70 factor induced by binding to the core enzyme. *J. Biol. Chem.*, 273:32995–33001.

Callaci, S., Heyduk, E. and Heyduk, T. 1999. Core RNA polymerase from *E. coli* induces a major change in the domain arrangement of the sigma70 subunit. *Mol. Cell.*, 3:229–38.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L. and Darst, S.A. 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell.*, 9:527–39.

Cellai, S., Mangiarotti, L., Vannini, N., Naryshkin, N., Kortkhonjia, E., Ebricht, R.H. and Rivetti, C. 2007. Upstream promoter sequences and alphaCTD mediate stable DNA wrapping within the RNA polymerase-promoter open complex. *EMBO Rep.*, 8:271–8.

Chan, B. and Busby, S. 1989. Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA polymerase. *Gene*, 84:227–36.

Chan, B., Spassky, A. and Busby, S. 1990. The organization of open complexes between *Escherichia coli* RNA polymerase and DNA fragments carrying promoters either with or without consensus  $-35$  region sequences. *Biochem. J.*, 270:141–8.

Chan, P.T. and Lebowitz, J. 1990. Site-directed mutagenesis of the  $-10$  region of the lacUV5 promoter. Introduction of dA4.dT4 tract suppresses open complex formation. *J. Biol. Chem.*, 265:4091–7.

- Colland, F., Orsini, G., Brody, E.N., Buc, H. and Kolb, A. 1998. The bacteriophage T4 AsiA protein: a molecular switch for sigma70-dependent promoters. *Mol. Microbiol.*, 27:819–29.
- Cook, V.M. and deHaseth, P.L. 2007. Strand opening-deficient E. coli RNA polymerase facilitates investigation of closed complexes with promoter DNA: effects of DNA sequence and temperature. *J. Biol. Chem.*
- Craig, M.L., Tsodikov, O.V., McQuade, K.L., Schlax, P.E., Jr, Capp, M.W., Saecker, R.M. and Record, M.T. Jr. 1998. DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA. *J. Mol. Biol.*, 283:741–56.
- Dairi, T., Inokuchi, K., Mizuno, T. and Mizushima, S. 1985. Positive control of transcription initiation in Escherichia coli. A base substitution at the Pribnow box renders ompF expression independent of a positive regulator. *J. Mol. Biol.*, 184:1–6.
- Davis, C.A., Capp, M.W., Record, M.T., Jr. and Saecker, R.M. 2005. The effects of upstream DNA on open complex formation by Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, 102:285–90.
- Davis, C.A., Bingman, C.A., Landick, R., Record, M.T., Jr and Saecker, R.M. 2007a. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*
- Davis, C.A., Bingman, C.A., Landick, R., Record, M.T., Jr. and Saecker, R.M. 2007b. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, 104:7833–8.
- Dombroski, A.J., Walter, W.A., Record, M.T. Jr., Siegele, D.A. and Gross, C.A. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma70 exhibit specificity of binding to promoter DNA. *Cell.*, 70:501–12.
- Dombroski, A.J., Johnson, B.D., Lonetto, M. and Gross, C.A. 1996. The sigma subunit of Escherichia coli RNA polymerase senses promoter spacing. *Proc. Natl. Acad. Sci. U.S.A.*, 93:8858–62.
- Dombroski, A.J. 1997. Recognition of the –10 promoter sequence by a partial polypeptide of sigma70 in vitro. *J. Biol. Chem.*, 272:3487–94.
- Estrem, S.T., Gaal, T., Ross, W. and Gourse, R.L. 1998. Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. U.S.A.*, 95:9761–6.
- Feklistov, A., Barinova, N., Sevostyanova, A. and other authors 2006. A Basal Promoter Element Recognized by Free RNA Polymerase sigma Subunit Determines Promoter Recognition by RNA Polymerase Holoenzyme. *Mol. Cell.*
- Fenton, M.S., Lee, S.J. and Gralla, J.D. 2000. Escherichia coli promoter opening and –10 recognition: mutational analysis of sigma70. *Embo J.*, 19:1130–7.
- Fenton, M.S. and Gralla, J.D. 2001. Function of the bacterial TATAAT –10 element as single-stranded DNA during RNA polymerase isomerization. *Proc. Natl. Acad. Sci. U.S.A.*, 98:9020–5.
- Fenton, M.S. and Gralla, J.D. 2003. Roles for inhibitory interactions in the use of the –10 promoter element by sigma70 holoenzyme. *J. Biol. Chem.*, 278:39669–74.
- Gaal, T., Ross, W., Blatter, E.E., Tang, H., Jia, X., Krishnan, V.V., Assamunt, N., Ebright, R.H. and Gourse, R.L. 1996. DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes Dev.*, 10:16–26.
- Gardella, T., Moyle, H. and Susskind, M.M. 1989. A mutant Escherichia coli sigma70 subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.*, 206:579–90.
- Gralla, J.D. 2005. Escherichia coli ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. *Mol. Microbiol.*, 55:973–7.
- Grana, D., Youderian, P. and Susskind, M.M. 1985. Mutations that improve the ant promoter of Salmonella phage P22. *Genetics*, 110:1–16.
- Grana, D., Gardella, T. and Susskind, M.M. 1988. The effects of mutations in the ant promoter of phage P22 depend on context. *Genetics*, 120:319–27.
- Gregory, B.D., Nickels, B.E., Darst, S.A. and Hochschild, A. 2005. An altered-specificity DNA-binding mutant of E. coli sigma70 facilitates the analysis of sigma70 function in vivo. *Molec Microbiol*, in press
- Grimes, E., Busby, S. and Minchin, S. 1991. Different thermal energy requirement for open complex formation by Escherichia coli RNA polymerase at two related promoters. *Nucleic Acids Res.*, 19:6113–18.
- Gruber, T.M. and Gross, C.A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev. Microbiol.*, 57:441–66.
- Guo, Y. and Gralla, J.D. 1998. Promoter opening via a DNA fork junction binding activity. *Proc. Natl. Acad. Sci. U.S.A.*, 95:11655–60.
- Harley, C.B. and Reynolds, R.P. 1987. Analysis of E. coli promoter sequences. *Nucleic Acids Res.*, 15:2343–61.
- Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C. and Gourse, R.L. 2006. rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell.*, 125:1069–82.
- Hawley, D.K. and McClure, W.R. 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. *Nucleic Acids Res.*, 11:2237–55.
- Helmann, J.D. and Chamberlin, M.J. 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.*, 57:839–72.
- Helmann, J.D. 1995. Compilation and analysis of Bacillus subtilis sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.*, 23:2351–60.
- Helmann, J.D. and deHaseth, P.L. 1999. Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry*, 38:5959–67.
- Hinton, D.M., Pande, S., Wais, N., Johnson, X.B., Vuthoori, M., Makela, A. and Hook-Barnard, I. 2005. Transcriptional takeover by sigma appropriation: remodelling of the sigma70 subunit of Escherichia coli RNA polymerase by the bacteriophage T4 activator MotA and co-activator AsiA. *Microbiology*, 151:1729–40.
- Hook-Barnard, I., Johnson, X.B. and Hinton, D.M. 2006. Escherichia coli RNA polymerase recognition of a sigma70-dependent promoter requiring a –35 DNA element and an extended –10 TGn motif. *J. Bacteriol.*, 188:8352–9.
- Inokuchi, K., Furukawa, H., Nakamura, K. and Mizushima, S. 1984. Characterization by deletion mutagenesis in vitro of the promoter region of ompF, a positively regulated gene of Escherichia coli. *J. Mol. Biol.*, 178:653–68.
- Jeong, W. and Kang, C. 1994. Start site selection at lacUV5 promoter affected by the sequence context around the initiation sites. *Nucleic Acids Res.*, 22:4667–72.
- Juang, Y.L. and Helmann, J.D. 1994. A promoter melting region in the primary sigma factor of Bacillus subtilis. Identification of functionally important aromatic amino acids. *J. Mol. Biol.*, 235:1470–88.
- Kainz, M. and Roberts, J. 1992. Structure of transcription elongation complexes in vivo. *Science*, 255:838–41.
- Kawano, M., Storz, G., Rao, B.S., Rosner, J.L. and Martin, R.G. 2005. Detection of low-level promoter activity within open reading frame sequences of Escherichia coli. *Nucleic Acids Res.*, 33:6268–76.
- Keener, J. and Nomura, M. 1993. Dominant lethal phenotype of a mutation in the –35 recognition region of Escherichia coli sigma70. *Proc. Natl. Acad. Sci. U.S.A.*, 90:1751–5.
- Keilty, S. and Rosenberg, M. 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.*, 262:6389–95.
- Kontur, W.S., Saecker, R.M., Davis, C.A., Capp, M.W. and Record, M.T., Jr. 2006. Solute probes of conformational changes in open complex (RPO) formation by Escherichia coli RNA polymerase at the lambda-daPR promoter: evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RPO. *Biochemistry*, 45:2161–77.



- Kuhnke, G., Fritz, H.J. and Ehring, R. 1987. Unusual properties of promoter-up mutations in the Escherichia coli galactose operon and evidence suggesting RNA polymerase-induced DNA bending. *Embo J.*, 6:507–13.
- Kumar, A., Malloch, R.A., Fujita, N., Smillie, D.A., Ishihama, A. and Hayward, R.S. 1993. The minus 35-recognition region of Escherichia coli sigma70 is inessential for initiation of transcription at an “extended minus 10” promoter. *J. Mol. Biol.*, 232:406–18.
- Kumar, A., Grimes, B., Fujita, N., Makino, K., Malloch, R.A., Hayward, R.S. and Ishihama, A. 1994. Role of the sigma70 subunit of Escherichia coli RNA polymerase in transcription activation. *J. Mol. Biol.*, 235:405–13.
- Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S.L., Rogulja, D., Nickels, B.E., Hochschild, A., Heyduk, T. and Severinov, K. 2002. A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science*, 295:855–7.
- Lambert, L.J., Wei, Y., Schirf, V., Demeler, B. and Werner, M.H. 2004. T4 AsiA blocks DNA recognition by remodeling sigma(70) region 4. *Embo J.*, 23:2952–62.
- Lamond, A.I. and Travers, A.A. 1983. Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. *Nature*, 305:248–50.
- Lewis, D.E. and Adhya, S. 2004. Axiom of determining transcription start points by RNA polymerase in Escherichia coli. *Mol. Microbiol.*, 54:692–701.
- Lim, H.M., Lee, H.J., Roy, S. and Adhya, S. 2001. A “master” in base unpairing during isomerization of a promoter upon RNA polymerase binding. *Proc. Natl. Acad. Sci. U.S.A.*, 98:14849–52.
- Lisser, S. and Margalit, H. 1993. Compilation of E. coli mRNA promoter sequences. *Nucleic Acids Res.*, 21:1507–16.
- Lisser, S. and Margalit, H. 1994. Determination of common structural features in Escherichia coli promoters by computer analysis. *Eur. J. Biochem.*, 223:823–30.
- Liu, J. and Turnbough, C.L., Jr. 1994. Effects of transcriptional start site sequence and position on nucleotide-sensitive selection of alternative start sites at the pyrC promoter in Escherichia coli. *J. Bacteriol.*, 176:2938–45.
- Liu, M., Garges, S. and Adhya, S. 2004a. lacP1 promoter with an extended –10 motif. Pleiotropic effects of cyclic AMP protein at different steps of transcription initiation. *J. Biol. Chem.*, 279:54552–7.
- Liu, M., Tolstorukov, M., Zhurkin, V., Garges, S. and Adhya, S. 2004b. A mutant spacer sequence between –35 and –10 elements makes the Plac promoter hyperactive and cAMP receptor protein-independent. *Proc. Natl. Acad. Sci. U.S.A.*, 101:6911–6.
- Lonetto, M., Gribskov, M. and Gross, C.A. 1992. The sigma70 family: sequence conservation and evolutionary relationships. *J. Bacteriol.*, 174:3843.
- Makoff, A.J. and Oxer, M.D. 1991. High level heterologous expression in E. coli using mutant forms of the lac promoter. *Nucleic Acids Res.*, 19:2417–21.
- Malhotra, A., Severinova, E. and Darst, S.A. 1996. Crystal structure of a sigma70 subunit fragment from E. coli RNA polymerase. *Cell.*, 87:127–36.
- Mandecki, W. and Reznikoff, W.S. 1982. A lac promoter with a changed distance between –10 and –35 regions. *Nucleic Acids Res.*, 10:903–12.
- Marr, M.T. and Roberts, J.W. 1997. Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, 276:1258–60.
- Matlock, D.L. and Heyduk, T. 2000. Sequence determinants for the recognition of the fork junction DNA containing the –10 region of promoter DNA by E. coli RNA polymerase. *Biochemistry*, 39:12274–83.
- McKane, M., Malone, C. and Gussin, G.N. 2001. Mutations at position –10 in the lambda PR promoter primarily affect conversion of the initial closed complex (R<sub>PC</sub>) to a stable, closed intermediate (R<sub>PI</sub>). *Biochemistry*, 40:2023–31.
- Meiklejohn, A.L. and Gralla, J.D. 1989. Activation of the lac promoter and its variants. Synergistic effects of catabolite activator protein and supercoiling in vitro. *J. Mol. Biol.*, 207:661–73.
- Mekler, V., Kortkhonja, E., Mukhopadhyay, J. and other authors. 2002. Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell.*, 108:599–614.
- Mellies, J., Brems, R. and Villarejo, M. 1994. The Escherichia coli proU promoter element and its contribution to osmotically signaled transcription activation. *J. Bacteriol.*, 176:3638–45.
- Michalowski, C.B., Short, M.D. and Little, J.W. 2004. Sequence tolerance of the phage lambda PRM promoter: implications for evolution of gene regulatory circuitry. *J. Bacteriol.*, 186:7988–99.
- Miroslavova, N.S. and Busby, S.J. 2006. Investigations of the modular structure of bacterial promoters. *Biochem. Soc. Symp.*, 1–10.
- Mitchell, J.E., Zheng, D., Busby, S.J. and Minchin, S.D. 2003. Identification and analysis of ‘extended –10’ promoters in Escherichia coli. *Nucleic Acids Res.*, 31:4689–95.
- Moran, C.P., Jr, Lang, N., LeGrice, S.F., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J. and Losick, R. 1982. Nucleotide sequences that signal the initiation of transcription and translation in Bacillus subtilis. *Mol. Gen. Genet.*, 186:339–46.
- Moyle, H., Waldburger, C. and Susskind, M.M. 1991. Hierarchies of base pair preferences in the P22 ant promoter. *J. Bacteriol.*, 173:1944–50.
- Mulligan, M.E., Brosius, J. and McClure, W.R. 1985. Characterization in vitro of the effect of spacer length on the activity of Escherichia coli RNA polymerase at the TAC promoter. *J. Biol. Chem.*, 260:3529–38.
- Munson, L.M., Mandecki, W., Caruthers, M.H. and Reznikoff, W.S. 1984. Oligonucleotide mutagenesis of the lacPUV5 promoter. *Nucleic Acids Res.*, 12:4011–17.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A. 2002a. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, 296:1285–90.
- Murakami, K.S., Masuda, S. and Darst, S.A. 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*, 296:1280–4.
- Murakami, K.S. and Darst, S.A. 2003. Bacterial RNA polymerases: the whole story. *Curr. Opin. Struct. Biol.*, 13:31–9.
- Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V. and Ebright, R.H. 2000. Structural organization of the RNA polymerase-promoter open complex. *Cell.*, 101:601–11.
- Nguyen, L.H. and Burgess, R.R. 1997. Comparative analysis of the interactions of Escherichia coli sigma S and sigma70 RNA polymerase holoenzyme with the stationary-phase-specific bolAp1 promoter. *Biochemistry*, 36:1748–54.
- Niedziela-Majka, A. and Heyduk, T. 2005. Escherichia coli RNA polymerase contacts outside the –10 promoter element are not essential for promoter melting. *J. Biol. Chem.*, 280:38219–27.
- O’Neill, M.C. 1989. Escherichia coli promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. *J. Biol. Chem.*, 264:5522–30.
- Orsini, G., Kolb, A. and Buc, H. 2001. The Escherichia coli RNA polymerase-anti-sigma70 AsiA complex utilizes alpha-carboxyl-terminal domain upstream promoter contacts to transcribe from a –10/–35 promoter. *J. Biol. Chem.*, 276:19812–19.
- Orsini, G., Igonet, S., Pene, C., Sclavi, B., Buckle, M., Uzan, M. and Kolb, A. 2004. Phage T4 early promoters are resistant to inhibition by the anti-sigma factor AsiA. *Mol. Microbiol.*, 52:1013–28.
- Ozawa, Y., Mizuno, T. and Mizushima, S. 1987. Roles of the Pribnow box in positive regulation of the ompC and ompF genes in Escherichia coli. *J. Bacteriol.*, 169:1331–4.
- Paget, M.S. and Helmann, J.D. 2003. The sigma70 family of sigma factors. *Genome Biol.*, 4:203.
- Pahari, S. and Chatterji, D. 1997. Interaction of bacteriophage T4 AsiA protein with Escherichia coli sigma70 and its variant. *FEBS Lett.*, 411:60–2.
- Pal, D., Vuthoori, M., Pande, S., Wheeler, D. and Hinton, D.M. 2003. Analysis of regions within the bacteriophage T4 AsiA protein involved in its binding to the sigma70 subunit of E. coli RNA polymerase and its role as a transcriptional inhibitor and co-activator. *J. Mol. Biol.*, 325:827–41.

- Panaghie, G., Aiyar, S.E., Bobb, K.L., Hayward, R.S. and de Haseth, P.L. 2000. Aromatic amino acids in region 2.3 of Escherichia coli sigma70 participate collectively in the formation of an RNA polymerase-promoter open complex. *J. Mol. Biol.*, 299:1217–30.
- Paul, B.J., Ross, W., Gaal, T. and Gourse, R.L. 2004. rRNA transcription in Escherichia coli. *Annu Rev. Genet.*, 38:749–70.
- Pene, C. and Uzan, M. 2000. The bacteriophage T4 anti-sigma factor AsiA is not necessary for the inhibition of early promoters in vivo. *Mol. Microbiol.*, 35:1180–91.
- Phadtare, S. and Severinov, K. 2005. Extended -10 motif is critical for activity of the cspA promoter but does not contribute to low-temperature transcription. *J. Bacteriol.*, 187:6584–9.
- Ponnambalam, S., Webster, C., Bingham, A. and Busby, S. 1986. Transcription initiation at the Escherichia coli galactose operon promoters in the absence of the normal -35 region sequences. *J. Biol. Chem.*, 261:16043–8.
- Qiu, J. and Helmann, J.D. 1999. Adenines at -11, -9 and -8 play a key role in the binding of Bacillus subtilis Esigma(A) RNA polymerase to -10 region single-stranded DNA. *Nucleic Acids Res.*, 27:4541–6.
- Repoila, F. and Gottesman, S. 2003. Temperature sensing by the dsrA promoter. *J. Bacteriol.*, 185:6609–14.
- Reznikoff, W.S. 1992. The lactose operon-controlling elements: a complex paradigm. *Mol. Microbiol.*, 6:2419–22.
- Roberts, C.W. and Roberts, J.W. 1996. Base-specific recognition of the nontemplate strand of promoter DNA by E. coli RNA polymerase. *Cell.*, 86:495–501.
- Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, 262:1407–13.
- Ross, W., Ernst, A. and Gourse, R.L. 2001. Fine structure of E. coli RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev.*, 15:491–506.
- Ross, W. and Gourse, R.L. 2005. Sequence-independent upstream DNA-alphaCTD interactions strongly stimulate Escherichia coli RNA polymerase-lacUV5 promoter association. *Proc. Natl. Acad. Sci. U.S.A.*, 102:291–6.
- Rutherford, S.T., Lemke, J.J., Vrentas, C.E., Gaal, T., Ross, W. and Gourse, R.L. 2007. Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase. *J. Mol. Biol.*, 366:1243–57.
- Saecker, R.M., Tsodikov, O.V., McQuade, K.L., Schlax, P.E., Jr, Capp, M.W. and Record, M.T., Jr. 2002. Kinetic studies and structural models of the association of E. coli sigma(70) RNA polymerase with the lambdaP(R) promoter: large scale conformational changes in forming the kinetically significant intermediates. *J. Mol. Biol.*, 319:649–71.
- Sanderson, A., Mitchell, J.E., Minchin, S.D. and Busby, S.J. 2003. Substitutions in the Escherichia coli RNA polymerase sigma70 factor that affect recognition of extended -10 elements at promoters. *FEBS Lett.*, 544:199–205.
- Sasse-Dwight, S. and Gralla, J.D. 1989. KMnO4 as a probe for lac promoter DNA melting and mechanism in vivo. *J. Biol. Chem.*, 264:8074–81.
- Sasse-Dwight, S. and Gralla, J.D. 1991. Footprinting protein-DNA complexes in vivo. *Methods Enzymol.*, 208:146–68.
- Schickor, P., Metzger, W., Werel, W., Lederer, H. and Heumann, H. 1990. Topography of intermediates in transcription initiation of E. coli. *Embo J.*, 9:2215–20.
- Schroeder, L.A., Choi, A.J. and DeHaseth, P.L. 2007. The -11A of promoter DNA and two conserved amino acids in the melting region of sigma70 both directly affect the rate limiting step in formation of the stable RNA polymerase-promoter complex, but they do not necessarily interact. *Nucleic Acids Res.*, 35:4141–53.
- Sclavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M. and Heumann, H. 2005. Real-time characterization of intermediates in the pathway to open complex formation by Escherichia coli RNA polymerase at the T7A1 promoter. *Proc. Natl. Acad. Sci. U.S.A.*, 102:4706–11.
- Severinova, E., Severinov, K. and Darst, S.A. 1998. Inhibition of Escherichia coli RNA polymerase by bacteriophage T4 AsiA. *J. Mol. Biol.*, 279:9–18.
- Shamoo, Y., Friedman, A.M., Parsons, M.R., Konigsberg, W.H. and Steitz, T.A. 1995. Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature*, 376:362–6.
- Shultzaberger, R.K., Chen, Z., Lewis, K.A. and Schneider, T.D. 2007. Anatomy of Escherichia coli sigma70 promoters. *Nucleic Acids Res.*, 35:771–88.
- Siebenlist, U. and Gilbert, W. 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. *Proc. Natl. Acad. Sci. U.S.A.*, 77:122–6.
- Siebenlist, U., Simpson, R.B. and Gilbert, W. 1980. E. coli RNA polymerase interacts homologously with two different promoters. *Cell.*, 20:269–81.
- Siegele, D.A., Hu, J.C., Walter, W.A. and Gross, C.A. 1989. Altered promoter recognition by mutant forms of the sigma70 subunit of Escherichia coli RNA polymerase. *J. Mol. Biol.*, 206:591–603.
- Silverstone, A.E., Arditti, R.R. and Magasanik, B. 1970. Catabolite-insensitive revertants of lac promoter mutants. *Proc. Natl. Acad. Sci. U.S.A.*, 66:773–9.
- Spassky, A., Busby, S. and Buc, H. 1984. On the action of the cyclic AMP-cyclic AMP receptor protein complex at the Escherichia coli lactose and galactose promoter regions. *Embo J.*, 3:43–50.
- Spassky, A., Kirkegaard, K. and Buc, H. 1985. Changes in the DNA structure of the lac UV5 promoter during formation of an open complex with Escherichia coli RNA polymerase. *Biochemistry*, 24:2723–31.
- Stefano, J.E., Ackerson, J.W. and Gralla, J.D. 1980. Alterations in two conserved regions of promoter sequence lead to altered rates of polymerase binding and levels of gene expression. *Nucleic Acids Res.*, 8:2709–23.
- Stefano, J.E. and Gralla, J.D. 1982. Spacer mutations in the lac ps promoter. *Proc. Natl. Acad. Sci. U.S.A.*, 79:1069–72.
- Straney, D.C. and Crothers, D.M. 1985. Intermediates in transcription initiation from the E. coli lac UV5 promoter. *Cell.*, 43:449–59.
- Straney, D.C. and Crothers, D.M. 1987a. Comparison of the open complexes formed by RNA polymerase at the Escherichia coli lac UV5 promoter. *J. Mol. Biol.*, 193:279–92.
- Straney, S.B. and Crothers, D.M. 1987b. Kinetics of the stages of transcription initiation at the Escherichia coli lac UV5 promoter. *Biochemistry*, 26:5063–70.
- Suh, W.C., Ross, W. and Record, M.T. Jr. 1993. Two open complexes and a requirement for Mg2+ to open the lambda PR transcription start site. *Science*, 259:358–61.
- Tanabe, H., Goldstein, J., Yang, M. and Inouye, M. 1992. Identification of the promoter region of the Escherichia coli major cold shock gene, cspA. *J. Bacteriol.*, 174:3867–73.
- Taylor, R.K., Garrett, S., Sodergren, E. and Silhavy, T.J. 1985. Mutations that define the promoter of ompF, a gene specifying a major outer membrane porin protein. *J. Bacteriol.*, 162:1054–60.
- Thouvenot, B., Charpentier, B. and Branlant, C. 2004. The strong efficiency of the Escherichia coli gapA P1 promoter depends on a complex combination of functional determinants. *Biochem. J.*, 383:371–82.
- Tsujikawa, L., Strainic, M.G., Watrob, H., Barkley, M.D. and DeHaseth, P.L. 2002. RNA polymerase alters the mobility of an A-residue crucial to polymerase-induced melting of promoter DNA. *Biochemistry*, 41:15334–41.
- Typas, A. and Hengge, R. 2005. Differential ability of sigma(s) and sigma70 of Escherichia coli to utilize promoters containing half or full UP-element sites. *Mol. Microbiol.*, 55:250–60.
- Vassilyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassilyeva, M.N., Borukhov, S. and Yokoyama, S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, 417:712–19.
- Voskuil, M.I. and Chambliss, G.H. 2002. The TRTGn motif stabilizes the transcription initiation open complex. *J. Mol. Biol.*, 322:521–32.

- Vuthoori, S., Bowers, C.W., McCracken, A., Dombroski, A.J. and Hinton, D.M. 2001. Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. *J. Mol. Biol.*, 309:561–72.
- Waldburger, C., Gardella, T., Wong, R. and Susskind, M.M. 1990. Changes in conserved region 2 of Escherichia coli sigma70 affecting promoter recognition. *J. Mol. Biol.*, 215:267–76.
- Warne, S.E. and deHaseth, P.L. 1993. Promoter recognition by Escherichia coli RNA polymerase. Effects of single base pair deletions and insertions in the spacer DNA separating the -10 and -35 regions are dependent on spacer DNA sequence. *Biochemistry*, 32:6134–40.
- Wilson, C. and Dombroski, A.J. 1997. Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. *J. Mol. Biol.*, 267:60–74.
- Young, B.A., Gruber, T.M. and Gross, C.A. 2002. Views of transcription initiation. *Cell.*, 109:417–20.
- Young, B.A., Gruber, T.M. and Gross, C.A. 2004. Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science*, 303:1382–4.
- Zenkin, N., Kulbachinskiy, A., Yuzenkova, Y., Mustaev, A., Bass, I., Severinov, K. and Brodolin, K. 2007. Region 1.2 of the RNA polymerase sigma subunit controls recognition of the -10 promoter element. *Embo J.*, 26:955–64.