doi:10.1016/j.jmb.2011.01.018 J. Mol. Biol. (2011) 412, 754–771

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

REVIEW

Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis

Ruth M. Saecker[1], M. Thomas Record Jr.[1,2] and Pieter L. deHaseth[3,4]*

[1]*Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*
[2]*Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*
[3]*Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, OH 44106-4973, USA*
[4]*Department of Biochemistry, Case Western Reserve University, Cleveland, OH 44106, USA*

Received 25 November 2010; received in revised form 7 January 2011; accepted 8 January 2011
Available online 1 March 2011

Edited by M. Gottesman

Keywords:
RNA polymerase; Promoter; Sigma factor; Transcription initiation; Intermediate complexes

Initiation of RNA synthesis from DNA templates by RNA polymerase (RNAP) is a multi-step process, in which initial recognition of promoter DNA by RNAP triggers a series of conformational changes in both RNAP and promoter DNA. The bacterial RNAP functions as a molecular isomerization machine, using binding free energy to remodel the initial recognition complex, placing downstream duplex DNA in the active site cleft and then separating the nontemplate and template strands in the region surrounding the start site of RNA synthesis. In this initial unstable "open" complex the template strand appears correctly positioned in the active site. Subsequently, the nontemplate strand is repositioned and a clamp is assembled on duplex DNA downstream of the open region to form the highly stable open complex, RP_o. The transcription initiation factor, σ^{70}, plays critical roles in promoter recognition and RP_o formation as well as in early steps of RNA synthesis.

Introduction

We begin this perspective with a brief overview of transcription initiation by bacterial RNA polymerase (RNAP), summarizing the players and the major steps in the process. Excellent review articles provide a more detailed coverage of many aspects of transcription initiation.[1–9] Here we focus on current advances in understanding the process of isomerization of the initial closed complex to form the stable open complex RP_o and the many crucial roles of the specificity subunit σ^{70} in all steps of initiation.

**Corresponding author.* Center for RNA Molecular Biology, School of Medicine, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-4973, USA. E-mail address: pld2@case.edu.

Abbreviations used: RNAP, RNA polymerase; NT, nontemplate strand; T, template strand; αCTD, α-subunit C-terminal domain; FRET, fluorescence resonance energy transfer; myxo, myxopyronin; lpm, lipiarmycin; SI, sequence insertion.

Initiation of RNA Synthesis in Bacteria

The essential players

The bacterial RNAP "core enzyme" (E) consists of five subunits, $\beta\beta'\alpha_2\omega$ (see Fig. 1). The core enzyme is capable of nonspecific DNA binding and initiation of RNA synthesis from DNA ends or nicks, but requires a sigma factor to initiate specific transcription from promoter DNA. Sigma assembles with

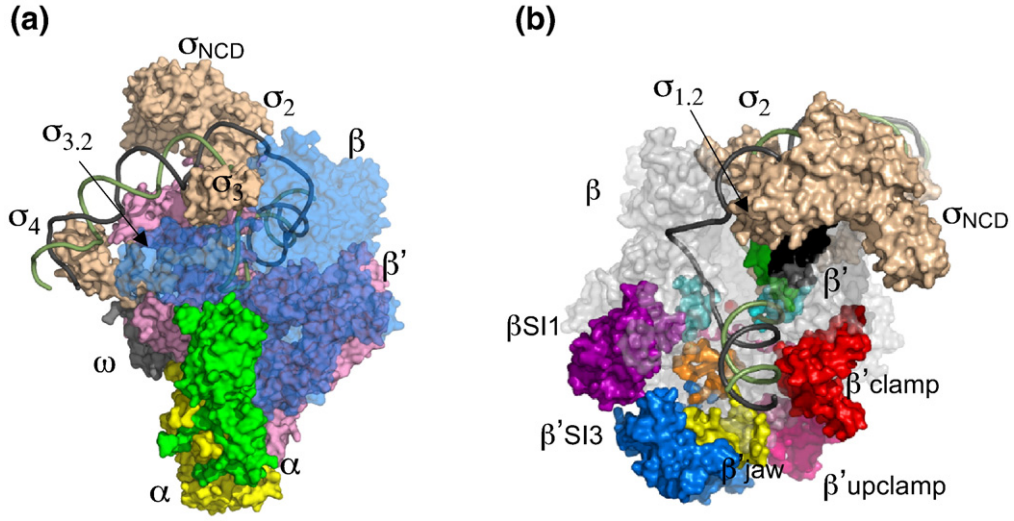

Fig. 1. Model of the *E. coli* RNAP (σ^{70} $\alpha_2\beta\beta'\omega$) open complex RP_o based on Protein Data Bank IDs 3IYD[10] and 3LU0.[11] (a) View of RP_o illustrating the interactions between promoter DNA [nontemplate strand (NT), black; template (T), dark green] and σ_2, σ_3, and σ_4 (wheat). Linker $\sigma_{3.2}$ is buried in the RNA exit channel. The N-terminal domains of α (bright green, yellow) form a hinge at the bottom of the cleft. σ_{NCD} is a folded nonconserved domain connecting $\sigma_{1.2}$ and σ_2. ω is shown in light gray. Missing from the figure are $\sigma_{1.1}$ and the flexibly tethered αCTD (not resolved in any holoenzyme structure to date). (b) View down into the active-site channel highlighting mobile regions on the periphery of the cleft and in the cleft. At the upstream entrance to the cleft, β′ clamp helices (black) tightly interact with $\sigma_{1.2}$ and σ_2. The open transcription bubble (−11 to +3 in this model) binds in the cleft, with the template strand start site (+1) next to the active site Mg^{2+} (red sphere) at the bottom. βSI1 (magenta) and β′SI3 (blue) are species-specific sequence insertions (SIs) present in *E. coli*. The remaining colored regions are highly conserved in bacteria.[12] Along with βSI and β′SI3, β′ jaw (yellow) and β′ clamp (red) appear positioned to clamp on the downstream duplex DNA after the bubble has opened.[13–15] Flexible elements in the cleft that likely bind and stabilize the DNA single strands in RP_o include the bridge helix (visible under the double-stranded–single-stranded boundary of the downstream DNA; pink), rudder (green), fork loop 2 (teal), and switch 2 (light blue). Other mobile elements shown are the β′ upclamp (hot pink; see Supplementary Fig. 1), which is proposed to interact with upstream DNA in forming I_1 (the first kinetically-significant intermediate at the λP_R promoter),[16] and the trigger loop (orange), which is known to be critically involved in the RNA synthesis steps.[17]

core to form the "holoenzyme" (or Eσ).[18,19] Sigma factors recognize specific promoter DNA sequences, interact with transcription activators, participate in promoter DNA opening, and influence the early phases of transcription (e.g., Gruber and Gross;[5] the latter two roles of sigma are further discussed in this review). The vast majority of studies of bacterial initiation have been carried out using *Escherichia coli* as model system. A model of the structure of the open complex formed by *E. coli* $E\sigma^{70}$ RNAP (shown in Fig. 1) highlights: (i) the positioning of σ^{70} on the core enzyme (Fig. 1a); (ii) the deep, wide cleft formed by β and β′ that binds the transcription bubble (Fig. 1b); and (iii) the flexible domains of β and β′ at the downstream end of the cleft proposed to assemble on the downstream duplex DNA to stabilize the open complex(es) (Fig. 1b).

All bacteria have a primary sigma factor that suffices for growth under nutrient-rich conditions. In *E. coli*, the primary sigma factor is σ^{70} (also called σ^{D}), reflecting its molecular mass of approximately 70,000 Da. In many other bacteria, the analogous primary sigma factor is designated σ^{A}. Most bacteria also have a complement of "alternative" or "minor" sigma factors (six in *E. coli*). Holoenzymes containing minor sigma factors recognize promoters of genes that can mitigate the effects of various adverse conditions.[5,6] Most bacterial sigma factors exhibit significant homology to *E. coli* σ^{70} [6] and, as such, belong the σ^{70} class. Others belong to the σ^{54} class due to their similarity to *E. coli* σ^{54} (also called σ^{N}, which is responsible for the expression of genes involved in nitrogen utilization), which has little sequence similarity with σ^{70}.[3] The evolution of these two distinct lineages of sigma factors is not understood.

The structure of σ^{70} is shown in Fig. 1a. The four regions of sequence conservation common to the σ^{70} class sigma factors[20] and the architecture of promoter DNA sequences that they recognize are shown in Fig. 2. (Regions of σ^{70} are designated in this review as subscripts; i.e., σ_2 refers to region 2 of σ^{70}.) In addition to the −10 and −35 hexameric recognition sequences[22] (Fig. 2), σ^{70} factors also recognize a TG sequence upstream of −10 (together called the extended −10)[23–26] and guanines in the discriminator region (see the text below) at −6 and −5.[27,28] The spacer length (i.e., the number of base pairs separating the −10 and −35 elements,

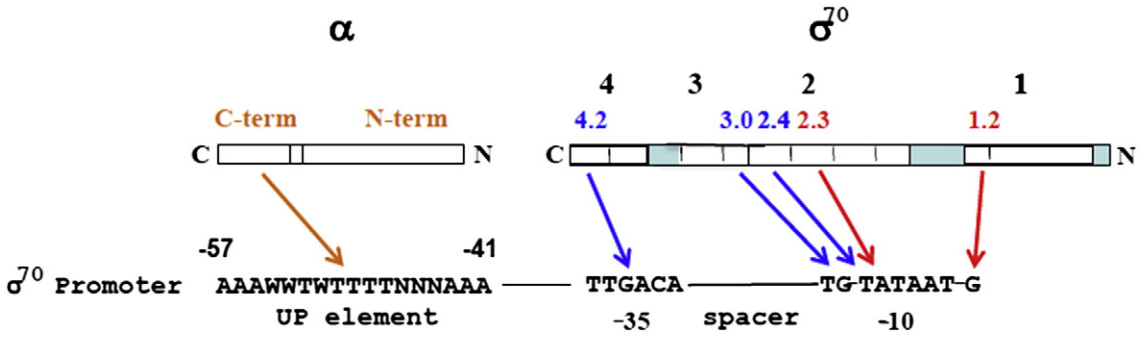

Fig. 2. Promoter recognition by amino acids of the α subunit and σ^{70}. Orange and blue arrows indicate recognition of promoter regions as double-stranded DNA elements by the α and σ^{70} subunits, respectively. The two red arrows delineate a region of the nontemplate strand DNA recognized by σ^{70} subsequent to strand separation. In the linear representations (not drawn to scale) of both σ^{70} and α, the N-termini are on the right. Only the sequence of the nontemplate strand is shown (5′ end on the left). A typical *E. coli* promoter does not have all elements shown and exhibits deviations from the consensus sequences shown here for the −10, −35, and UP[21] elements, as well as the consensus spacer length (17 bp).

optimally 17 bp) and the number of base pairs separating the −10 element from the transcription start site (optimally 7 bp)[29–31] both modulate the interactions of $E\sigma^{70}$ with the promoter. Some promoters also include ~20 bp of A/T-rich sequence upstream of the −35 element, referred to as an "UP element" (see Fig. 2). The UP element is recognized by the flexibly tethered α-subunit C-terminal domain (αCTD).[21] The αCTD also can bind nonspecifically to upstream DNA,[32–35] making contacts up to ~−90.

Steps of transcription initiation

Specific binding of $E\sigma^{70}$ RNAP to promoter DNA, forming an initial closed complex RP_c, triggers a series of conformational changes in both biomolecules. This series of events, often collectively called "isomerization," opens ~13 bp from the −10 element to just beyond the transcription start site, creating the initiation "bubble" and an unstable open complex.[36] In this step or in subsequent steps of forming the final stable open complex (RP_o), the +1 template (T) strand base is placed in the active site of the polymerase, and the nontemplate (NT) strand is placed in its binding track. RP_o is stabilized by the assembly and DNA binding of a downstream jaw/clamp,[13–15,36] which presumably is important for processive transcription (see Fig. 1b and the text below).

During isomerization, contacts between σ_2 and the duplex form of the −10 region in the closed complex are replaced by interactions between conserved aromatic residues in σ_2 and NT strand bases from −11 to −7 during or after DNA opening (see the text below). Work to date indicates that the −10 element (with the exception of −12, which remains base-paired) is primarily recognized as single-stranded DNA.[37] Open bases at positions −6 and −5 on the NT strand (the discriminator region; see Fig. 2) interact with $\sigma_{1.2}$,[27,28] as judged by cross-linking experiments.[38] Base identity at these positions has very large effects on the rate of dissociation of the open complex at the ribosomal *rrnB* P1 promoter, but only small effects on the binding and isomerization steps that determine the association kinetics.[28] Thus, bases on the NT strand from −11 to −5 appear to be largely recognized in the single-stranded state[28] after the opening of the initiation bubble. Importantly, no such recognition occurs on the T strand. The difference in interactions with the T strand *versus* the NT strand has consequences for the later steps of NTP addition and promoter escape. σ_3 (also called $\sigma_{2.5}$) interacts with the extended −10 TG sequences;[23] it appears to also play a role in the steps after promoter binding (Fig. 2). Although the extended −10 element remains duplex throughout initiation, changes in this sequence primarily affect the rate of isomerization and not closed complex formation.[39,40]

Templated RNA synthesis (transcription) involves covalent bond formation between the 3′ OH end of the nascent RNA and the α phosphate of the incoming NTP (nucleoside triphosphate). Phosphodiester bond synthesis results in the extension of the chain by one residue and in the release of pyrophosphate. At most $E\sigma^{70}$ promoters, transcription is initiated with ATP or GTP, but promoters at which initiation occurs with CTP or UTP have also been characterized. The initial transcribing complex may go down either productive or abortive paths. As a short RNA chain is synthesized, contacts between the RNAP and the promoter DNA upstream of −11 remain intact, while the promoter DNA is progressively "scrunched"[41,42] and the transcription bubble is expanded[43] as additional DNA is pulled in and copied into RNA. This process builds up stress and sets up a competition between extending the RNA chain and increasing the size of the DNA bubble, or releasing both the small product and

the stress in the scrunched DNA to revert to RP_o.[41,42] When the nascent RNA reaches a critical length of about 11 nucleotides, the stress is instead relieved by disruption of the contacts between the RNAP and the promoter DNA.

The number and the length of abortive products produced prior to productive initiation are a function of promoter sequence and conditions,[44,45] but the precise "rules" governing this behavior remain unknown. For the single subunit phage T7 RNAP, single-molecule and fluorescence studies demonstrate that the probability of transition from an initiating complex to an elongation complex strongly depends on RNA length.[46] Abortive initiation was once thought to be an *in vitro* artifact or an inconsequential aspect of promoter escape. However, abortive RNAs have now been detected *in vivo*.[47] This discovery raises the intriguing possibility that the small products (e.g., 2–4 mers) may rebind the open complex and thus prime initiation *in vivo* and alter gene expression in a concentration-dependent fashion.[47]

σ^{70} is not required for elongation and is typically released from the transcription complex when the RNA reaches a length of 12–15 nt.[48] Release of σ^{70} is likely triggered by events set in motion when the nascent RNA–DNA hybrid reaches 8–9 bp. Further extension requires displacing the sigma linker connecting σ_2 and σ_4 that lies in the RNA exit channel.[49,50] While the competition with the growing RNA chain is thought to release $\sigma_{3.2}$ and σ_4, it is unclear how the remaining interaction between σ_2 and β' clamp helices[51] is disrupted. Indeed, when σ^{70} is retained, σ_2 induces promoter-proximal pausing at promoters with a −10-like sequence in the NT strand downstream of the start site.[52–55] Under some conditions, σ^{70} release is delayed beyond the transition from initiation to productive elongation. The events governing σ^{70} release *versus* retention remain to be defined. They are likely regulated and thus motivate ongoing investigations.

X-ray and electron microscopy structures of Eσ RNAP: Implications for transcription initiation

X-ray structural data for core,[56] $E\sigma^A$,[50,57] and several nucleic acid–thermophilic RNAP complexes[17,49] have had considerable impact and influence on the understanding of bacterial transcription initiation. In common with other nucleic acid polymerases (but on a larger scale), the active site lies at the bottom of a deep cleft (~70 Å deep and >100 Å long; see Fig. 1b). In addition, the $E\sigma^A$ structures detail the extensive interface formed between the highly conserved regions of sigma and core, as first deduced by biochemical and genetic studies.[58]

The highly conserved multisubunit RNAP architecture[12] itself appears to play a key role in discriminating promoter DNA from nonpromoter DNA during initiation. First, the arrangement of sigma on core and, in particular, the positioning of the promoter-recognition regions of sigma relative to the active-site cleft create a series of obstacles for the DNA to overcome to form an open complex. Interactions of σ_2 and σ_4 with the −10 and −35 elements (Fig. 2) define a promoter DNA trajectory in the initial "short footprint" closed complex RP_c (see the text below), which is at 90° with respect to the cleft. Consequently, a sharp bend at −11[59] and/or DNA opening outside the cleft[50,60,61] must be introduced for DNA to enter the cleft.

Does promoter DNA containing the start site of transcription (+1) enter the cleft as separated single strands or as a double helix? The ongoing debate regarding this question is driven, in part, by the narrow width (>25 Å) of the cleft seen in the $E\sigma^A$ crystal structures. This observation motivates the hypothesis that the cleft "screens" the state of the DNA by only allowing single-stranded, but not duplex, DNA entry.[50,60] However, this narrow width appears to be a snapshot of just one conformational state. In general, structures of the bacterial RNAP and eukaryotic RNAP II in various states of ligation and/or crystal forms exhibit a range of cleft widths, with distances (between the β' clamp and the β pincer lobes) varying from <25 Å (e.g., "open") to >15 Å ("closed") (Mukhopadhyay *et al.*[62] and references therein). Recent single-molecule fluorescence resonance energy transfer (FRET) experiments confirm that clamp opening and closing occur, demonstrating that the "hinges" in β and β' at the base of the cleft are flexible in solution (A. Chakraborty and R. Ebright, personal communication).

Comparison of RNAP structures indicates that cleft width is controlled, in part, by the conformation of "switch 2" at the base of β' (see Fig. 1b) (Mukhopadhyay *et al.*[62] and references therein). Recent work has demonstrated that the bacterial transcription initiation inhibitors myxopyronin (myxo), corallopyronin, ripostatin, and lipiarmycin (lpm) target switch 2.[62–65] Although RNAP–myxo crystal structures indicate that myxo binding stabilizes a "partly closed" conformation of the clamp,[62,63] footprinting data on myxo–RNAP–promoter DNA complexes suggest that myxo inhibits melting of the start site region (−2 to +2) but does not prevent the entry of duplex DNA into the cleft.[63] Like myxo ternary complexes, lpm–RNAP–promoter DNA complexes are protected downstream to at least +15 from DNase I cleavage.[65] However, unlike myxo–RNAP–promoter DNA complexes, no permanganate-reactive thymines are detected in the presence of lpm.[65] These data, along with results summarized below, argue that cleft width does not preclude duplex DNA entry during formation of RP_o.

An additional block to forming the open complex is created by the acidic N-terminal domain of σ^{70}. The single-stranded nucleic acid mimic $\sigma_{1.1}$ binds in the

cleft, blocking access to the active site.[66,67] For proper orientation of the start site base with respect to the active site Mg^{2+}, $\sigma_{1.1}$ must be repositioned and the T strand must descend ~70 Å from its location in the closed complex (see the text below). After NTP binding and short RNA synthesis, the transition from initiation to elongation (promoter escape) requires displacing the flexible linker ($\sigma_{3.2}$) connecting σ_2 and σ_4 from the RNA exit channel,[50,57] and breaking the contacts of σ_2 and σ_4 with the −10 and −35 elements of promoter DNA, respectively.

Although high-resolution structures are not available for *E. coli* RNAP, a recent 20-Å electron microscopy structure of a ternary complex (*E. coli* holoenzyme-CRP-DNA)[10] and a complete model of the *E. coli* core enzyme[11] reveal the locations of three large sequence insertions (SIs) in the *E. coli* β and β′ subunits that are absent in the thermophilic RNAP.[12] Two of these insertions, βSI1 and β′SI3, lie at the downstream end of the cleft (see Fig. 1b). β′SI3 occupies a particularly prominent position: it forms a tethered independent domain with the highly conserved β′ "jaw." The β′ jaw/SI3 domain is highly mobile[68] and likely provides an additional steric "block" to prevent nonpromoter duplex DNA from accessing the active site.[16]

Steps in RP_c-to-RP_o Isomerization

Mechanistic studies

How is the start site DNA opened, placed in the active site, and stabilized? During RP_o formation, how and when are the obstacles that prevent nonpromoter DNA from accessing the cleft and being opened overcome? For several decades, kinetic mechanistic and footprinting studies have been employed to determine the sequence of conformational changes and the nature of intermediate complexes on the pathway from the initial promoter-recognition complex RP_c to RP_o. At the lac UV5 and λP_R promoters, at least two steps are required to convert the initial closed complex into the final stable open complex RP_o.[69–71] However, the "isomerization" intermediates that separate the closed complex and RP_o are relatively unstable and short lived (<1 ms to 1 s; see Fig. 3). To date, they have resisted characterization by crystallography, cross-linking, FRET, and single-molecule approaches. Their size currently precludes NMR characterization.

While methods for characterizing transient isomerization intermediates were being developed, attention was focused on initial promoter recognition (forming RP_c) and its regulation by promoter sequence and by activator or repressor proteins. Information about RP_c and other potential intermediates has been obtained by equilibrium footprinting either at low temperatures (0–15 °C) (e.g., Kovacic,[73] Cowing *et al.*,[74] and Schickor *et al.*[75]) or with variant RNAPs[40] unable to effect promoter DNA melting. These complexes, all closed, exhibit different hydroxyl radical (·OH) or DNase footprint boundaries at different temperatures or promoters (see the text below) and have been given different names.[59,69,70,72,76–78] Because of the challenges in characterizing kinetically significant but unstable intermediates in real time, the mechanism of forming RP_o has often been condensed into two steps: $R + P \rightarrow RP_c \rightarrow RP_o$. This mechanism collapses all of isomerization, including DNA opening and placement of the start site base of the T strand in the active site, into a single step.

In the association direction of the three-step mechanism for the lac UV5 and λP_R promoters,[70,71] the first kinetically-significant intermediate (designated I_1 at the λP_R promoter) is found to be closed[16]. I_1 is more "advanced" than RP_c, protecting DNA to +20. While RP_c likely forms first, it never accumulates at the λP_R promoter. I_1 isomerizes in the rate-determining step to a second intermediate (designated I_2 at the λP_R promoter and found to be open[36]), which rapidly converts to RP_o. In the dissociation direction, the reverse direction of this same step ($I_2 \rightarrow I_1$) is rate determining. Thus, the $I_1 \rightarrow I_2$ and $I_2 \rightarrow I_1$ steps are the bottleneck steps in each direction. In the forward direction, use of high concentrations of RNAP creates a "burst" in the population of the closed complex immediately preceding this rate-limiting isomerization step.[77–79] To create bursts of intermediates formed after I_1, RP_o is destabilized by using salts and solutes that do not destabilize these intermediates. To trap the elusive second intermediate, temperature downshifts were attempted.[70,72,80,81] However, because I_2 is an open complex, this approach also destabilized it rather than leading to its accumulation.[14,36]

DNA footprinting studies of intermediate complexes

Methods for investigating late intermediates of isomerization have been developed recently.[14,36] These methods, combined with rapid-quench (<2 ms) mixing, allow one to perform "real-time" kinetic and chemical footprinting experiments on the timescale of the formation and disappearance of transient intermediates.[36,77,78] To date, all structural information about complexes known to be on-pathway intermediates in RP_o formation has come from chemical and enzymatic DNA footprinting methods.

Recent advances in describing the large-scale conformational changes that occur after recruitment of RNAP to the promoter and initial specific binding are summarized in Fig. 3. In addition to opening of the promoter DNA, strong evidence is obtained for DNA wrapping and for coupled folding and

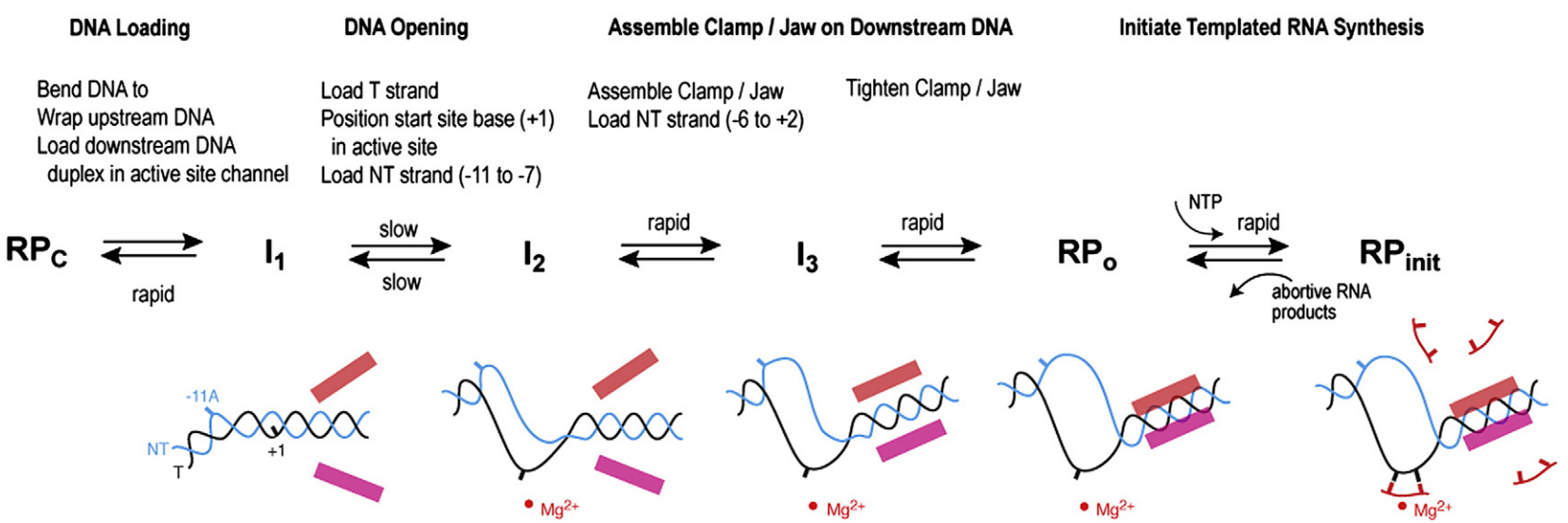

Fig. 3. Summary of the proposed isomerization steps that form the initiating complex (RP_{init}) after recruitment of RNAP to form an initial complex at the promoter (RP_c). Formation of the closed complex RP_c triggers a series of subsequent large-scale conformational changes. The RNAP molecular machine places start-site duplex DNA in the active-site cleft in I_1, opens it to form I_2, and stabilizes the open form by assembling a clamp in I_3 and RP_o (model based on Gries *et al.*,[36] Kontur *et al.*,[13,15] Davis *et al.*,[16] and Craig *et al.*[72]). Once promoter DNA is open, NTPs can bind, and transcription initiates. I_2 and I_3 are open complexes; current studies are addressing whether they can bind NTPs and initiate transcription.

domain repositioning of RNAP. Because all of these conformational changes are driven by binding free energy, the motions in the RNAP machinery are linked to the DNA sequence and structure in the interfaces that form, as well as to solution conditions. Below we detail our current understanding of the progression of conformational changes, the structures and stabilities of the intermediates, and the controversies and questions that remain.

RP_c formation: Initial recognition of duplex promoter DNA sequences

Initial specific interactions of RNAP with promoter DNA form a closed complex in which the promoter DNA is fully duplex. Based on the low-temperature footprinting data (see the text above) and the structures for holoenzyme[50,57] and for a complex of RNAP with a "fork junction" promoter fragment,[49] a model of such an initial closed complex has been proposed. In this model, duplex promoter DNA interacts with σ_2 and σ_4, and a continuous DNA duplex extends downstream of the −10 element, projecting away from the active-site cleft and therefore is cleavable by DNase I or ·OH.[60,82]

Real-time ·OH footprinting after mixing high concentrations of RNAP and T7A1 promoter DNA shows progression in the protection of the downstream boundary after mixing.[77,78] In these studies, a series of closed intermediates, which initially establish protection between −80 and −55 and then progressively extend it downstream, has been proposed. These snapshots suggest that interactions involving αCTDs and the T7A1 UP element and/or other far-upstream DNA contacts with RNAP are established first (see also Borukhov and Nudler[8]). The footprint then extends downstream as contacts between σ_2, σ_3, and σ_4 and the −10 element, and between the spacer and the −35 element of the promoter DNA, respectively, are established to form a RP_c complex with a downstream boundary of −5.[77,78]

More advanced closed complexes, including the upstream-wrapped closed intermediate I_1 with downstream duplex in the active-site cleft

Downstream boundary and its implications for closed complexes. More "advanced" closed complexes with downstream boundaries extending to +15–20 have also been characterized.[16,72,74,80,83–86] Conversion from an "RP_c"-like complex into one that protects one to two turns of the DNA helix downstream of the start site (+1) has been observed by increasing the temperature to 16–20 °C[74,80,83] or, in the case of the *rrnB* P1 promoter, by increasing the temperature to 37 °C in the absence of NTPs.[84,86] In all cases, chemical probes (dimethyl sulfate and permanganate) do not detect open/unstacked bases in these "advanced" complexes, and the periodic patterns of protection observed from ~−55 to −12 transition to full protection of both strands of the helix between −11 and +15.

At the λP_R promoter, both the transient, kinetically significant intermediate closed complex (I_1) that accumulates early in the time course of open complex formation[16] and the low-temperature (0 °C) closed complex,[87] also known to be I_1 by extrapolation of thermodynamic data from 7 °C and higher temperatures, have an extended downstream footprint (to +20–25). Based on the holoenzyme structures, we proposed that protection of both strands from −11 to +15 results from a sharp ~90° bend at the upstream end of the −10 element that inserts downstream duplex DNA into the active-site cleft.[59] Additional protection to +20 and +25 likely arises from mobile elements at the downstream end of the cleft that block access to the DNA backbone (e.g., β′ SI3; see Fig. 1b). Conversion of RP_c to a more I_1-like complex appears to be driven by increasing temperature, by favoring the bend at −11/−12 and/or the interactions that stabilize the bend. At λP_R, I_1 is the most advanced closed complex because it opens in the next kinetic step.

In closed complexes with downstream boundaries between −5 and +15,[33,84,86] downstream DNA is presumably only partially inserted into the cleft. Is this because of the difficulty in bending the DNA or the difficulty in inserting the duplex in the cleft? Possible examples of both scenarios are available. Davis *et al.* found that the I_1 intermediate formed by RNAP at an upstream-truncated (UT-47) λP_R promoter exhibits a downstream boundary of +2(NT)/+7(T).[16] Since the interactions of the −10 element are presumably identical in UT-47 and full-length λP_R, the difference in the insertion of the downstream duplex in the cleft may therefore indicate an obstacle in the cleft that is removed as a result of interactions with far upstream DNA (above −47).

Upstream boundary and its implications for closed complexes. Interactions with DNA upstream of the −35 hexamer are established in the steps of promoter recognition.[16,77,78] They influence the stability of the intermediate I_1 and the rate of converting it to the next intermediate I_2. The αCTDs (see the text above) mediate upstream interactions by binding DNA either specifically[21,88] or nonspecifically[32–35] and by interacting with activator proteins.[1,89] Intriguingly, all proteins (including the αCTDs, based on DNase I enhancements) that bind upstream of the −35 hexamer and modulate transcription bend DNA [e.g., CRP (*c*yclic AMP *r*eceptor *p*rotein), IHF (*i*ntegration *h*ost *f*actor), and FIS (*f*actor for *i*nversion *s*timulation)]. The function of these DNA bends may be simply to provide better interactions with the activator and/or the αCTDs. However, many transcription factors

bind to sites upstream of −90, presumably out of "reach" of the flexibly tethered αCTDs. In addition, the phasing of upstream binding sites for these factors (e.g., see Dethiollaz *et al.*[90] and Giladi *et al.*[91]) affects their action. Shifting the UP element in the upstream direction relative to the −35 hexamer of a given promoter abolishes UP element activation of transcription, regardless of phasing; lengthening the αCTD–αN-terminal domain linker does not restore full-length transcripts to their nondisplaced UP element levels.[92] Moving the UP element upstream prevents the formation of a complementary interface between σ_4 and the adjacent (proximal) αCTD.[93,94] Likewise, shifting sites for transcription factors that naturally abut the −35 hexamer upstream destroys the interface that they form with σ_4.[95,96] In the wild-type context, formation of these complementary interfaces is a key event in transcription initiation.

The above data have engendered several hypotheses about the role of upstream interactions. One hypothesis is that protein–protein interfaces communicate "allosterically" with the active-site channel to affect steps in initiation. Alternatively, or in addition to possible allosteric effects, we have proposed that the network of interactions between σ_4 and the αCTDs (and transcription factors, when present) and DNA bends the DNA from ~−30 to −55 and thereby sets the trajectory of far upstream DNA in the early steps of RP_o formation[16] (see Supplementary Fig. 1). The importance of the upstream DNA trajectory is based on several observations. First, the presence of DNA upstream of the −35 element at the lac UV5 and λP_R promoters profoundly accelerates the bottleneck isomerization step (see Fig. 3), now established as the DNA opening step at λP_R.[36] The isomerization rate constant k_2 (conversion of I_1 into I_2) for full-length λP_R is ~50-fold larger than that for truncated λP_R, with the upstream DNA deleted beyond −47 (UT-47), an effect as large as or larger than that exerted by activator binding. Surprisingly, deletion of this upstream DNA has little effect to no effect on the stability of I_1.[33,34] Second, deletion of upstream DNA leads to a "less advanced" closed complex that only protects downstream DNA to +2 (T)/+7 (NT)[16] relative to ~+20 observed for the full-length λP_R promoter.

How might upstream interactions facilitate the loading of downstream DNA in the cleft? In the final stable open complex at the λP_R promoter, the upstream boundary defined by DNase I or ·OH cleavage ends at ~−65. However, ·OH footprints of I_1 on full-length λP_R reveal modest protection of the DNA backbone on both strands to at least −85.[16] DNase I hypersensitive sites in I_1[16,72] indicate that a bend occurs just upstream of the −35 element. Mapping the I_1 protection pattern and the inferred bend onto available X-ray structures indicates most simply that RNAP wraps DNA around the "back" of the β′ subunit (Supplementary Fig. 1). Alternatively, the upstream interactions responsible for the large effect on the kinetics of the DNA opening step discussed above and for the far upstream ·OH footprint could involve the mobile αCTDs (see the text below).

In the upstream wrapping model,[16] the bend induced by σ_4 and the two αCTDs places far upstream DNA near the downstream end of the cleft (see Supplementary Fig. 1). Because the pattern of protection from −65 to −85 is not periodic (as typically observed for the αCTDs[97]), we proposed that this region of DNA is directed into a surface groove formed by β′ and the N-terminal domain of the associated α subunit. If so positioned, the upstream DNA lies near a conserved mobile element in β′, termed the upstream clamp. The upstream clamp is directly connected to other dynamic elements at the downstream end of the cleft: conserved jaw, trigger loop, and β′SI3. Based on our kinetic and footprinting data,[16,33] we hypothesize that interactions between upstream DNA and upstream clamp restrain the movements of the jaw domain, trigger loop, and β′SI3. Without this constraint, it appears that these elements sterically interfere with the loading of the downstream DNA in the cleft (see Fig. 1b; Supplementary Fig. 1). Predictions of this model are currently being tested.

Alternative hypotheses for the role of upstream DNA invoke a direct role for the αCTDs in mediating the acceleration of the DNA melting step beyond bending upstream DNA. In the absence of the αCTDs, the presence of upstream DNA only increases k_2 by ~2.5-fold.[34] Cross-linking of RP_o indicates that the αCTDs can occupy multiple sites on the upstream DNA.[35] Discerning whether the role of upstream DNA is to simply provide additional nonspecific αCTD binding sites or whether the αCTDs and σ_4 together set a trajectory required for wrapping interactions between upstream DNA and other elements of RNAP in I_1 or other early complexes awaits further experiments.

$I_1 \rightarrow I_2$: DNA opening is the bottleneck step in RP_o formation at the λP_R promoter

How is DNA opened by RNAP? Two conflicting hypotheses describing this critical step exist in the field. One hypothesis, based on structural data, posits that opening is nucleated by DNA breathing above the active-site cleft, after which the T strand enters the cleft and diffuses to the active site.[50,60] Evidence in support of this proposal has been obtained from molecular dynamics simulations on modeled structures formed by the bacterial RNAP[61] and from a comparison of the time evolution of downstream ·OH and MnO_4^- footprints in association experiments at the T7A1 promoter.[77] A second hypothesis proposes that the DNA duplex is first

loaded in the cleft, where it is then actively opened by elements on RNAP.[15,36] This proposal is supported by extensive kinetic and footprinting (equilibrium and real time) experiments on the λP_R promoter (cf. Fig. 3), equilibrium footprinting experiments at other promoters (see the text above), and equilibrium footprinting experiments performed in the presence of antibiotics that block DNA melting (discussed above).

Closed promoter DNA–RNAP complexes that protect the DNA backbone to at least +15 demonstrate that duplex DNA can bind in the active-site cleft. However, in most cases, evidence that complexes populated at equilibrium were on-pathway kinetic intermediates was not obtained. However, extensive kinetic studies (filter binding) of RP_o formation at λP_R, combined with real-time footprinting experiments, provide strong evidence that duplex DNA (−11 to +20) occupies the active-site cleft in the final closed on-pathway intermediate I_1. Once bound in the cleft, the next step ($I_1 \rightarrow I_2$) opens DNA (−11 to +2), as detailed below.

At λP_R, DNA in I_1 is not MnO_4^- reactive and is continuously protected from ·OH and DNase I on both strands from −11 to positions +20–25.[16,72] Thymines in the subsequent kinetically significant intermediate I_2 are MnO_4^- reactive at all positions reactive in RP_o.[36] In addition, real-time footprinting experiments reveal that the extended downstream footprint (protected from ·OH cleavage to ~+20) of I_1 develops in 100 ms. In contrast, the MnO_4^- reactivities of thymines detected in RP_o develop much more slowly (tens of seconds; Heitkamp, Drennan, *et al.*, in preparation). Therefore, we conclude that duplex DNA binds in the cleft in I_1, and that the entire bubble opens concertedly in the cleft at λP_R in one kinetic step.[15,16,36,59]

In addition to evidence cited directly above, the following data also indicate that DNA opens in the cleft. The rate constant for $I_1 \rightarrow I_2$ is strongly temperature dependent.[59] The corresponding activation energy (34 kcal) is consistent with the cooperative opening of at least 6–7 bp in the I_1–I_2 transition state. While salt and other solutes exhibit large effects on DNA opening in solution, they only exert small effects on the DNA opening (k_2) and closing (k_{-2}) steps for the λP_R promoter.[13–15,36,59] Most simply, these data indicate that DNA opening occurs in the sequestered environment of the cleft and not outside it. Alternatively, compensating for the stabilizing and destabilizing effects of these salts and other solutes may accompany opening.

$I_2 \rightarrow RP_o$: Evidence for the assembly and DNA binding of a downstream clamp/jaw to stabilize the open complex

Evidence to date indicates that the final steps of isomerization involve the interconversion of multiple different open complexes, including on-pathway intermediates I_2 and I_3, as well as the final open complex RP_o. This striking discovery has significant implications for the regulation of transcription initiation.[14,36] Major changes in DNA and RNAP in the conversion of I_2 into RP_o were revealed for the first time by burst footprinting of the dissociation intermediate I_2 and by analysis of dissociation data as a function of both stabilizing and destabilizing solutes and salts.[13–15,36] These include downstream folding and assembly of >100 residues of mobile elements of RNAP (Figs. 1 and 3; Supplementary Fig. 1) to form a clamp/jaw on downstream DNA, as well as establishment of in-cleft interactions. Evidence for the latter includes a 2-fold increase in the MnO_4^- reactivity of thymine bases in the downstream region of the NT strand (−4, −3, and +2) in the conversion of I_2 into RP_o. Thymine bases in the upstream half of the NT strand remain protected from MnO_4^- oxidation by being bound to RNAP (σ_2) in both I_2 and RP_o. The start site thymine (+1, T strand) is equally MnO_4^- reactive in I_2 and RP_o, suggesting that it is correctly positioned in I_2.

Downstream interactions in the open complexes

Numerous lines of evidence demonstrate that RNAP undergoes a large-scale conformational change in the steps following DNA melting. The extreme effects of solutes, temperature, and salts on the steps converting I_2 into RP_o suggest that the late steps of RP_o formation create a new protein–DNA interface in a process that involves coupled folding.[98] Quantitative analysis of the effects of multiple solute probes on the dissociation rate constant K_d indicates that 75–100 residues fold in the conversion of I_2 to RP_o; the effects of salt are consistent with a burial of 10 or more DNA phosphates.[15] The conformational changes that occur in these steps appear to be comparable in scale to those that occur in the conversion of the initiation complex to the elongation complex of the T7 phage RNAP (~300 amino acids refold[99]).

To interpret these results, we proposed[13–15] that several large (50–70 residue) mobile regions of the β′ and β subunits fold and assemble into a jaw/clamp superstructure that binds to duplex DNA (see Fig. 1b; Supplementary Fig. 1) after DNA opening. Given the tight binding interactions established upstream in the early steps, assembly of the downstream clamp is likely delayed to allow unimpeded rotation of the downstream DNA on its helical axis by 1.3 turns (470°).[13–15] Regions in β′ include: (i) the jaw; (ii) a highly positively charged helix hairpin helix; (iii) β′SI3; and (iv) a C-terminal region adjacent to (i). Individual deletions of (i)–(iii) all destabilize RP_o;[13,84,100] deletions/mutations in (iv) have not been studied. In addition, βSI1 also likely forms part of the clamp.[14] In the recent model

of RP_o[10] (as well as in the complete model of an *E. coli* transcription elongation complex[11]), all of the regions above are positioned near the downstream DNA, but do not necessarily directly interact with it. This may be a consequence of the low resolution of DNA in electron microscopy images and of basing the model on the transcription elongation complex, which only protects downstream DNA to ~+10–15[101] (~1 turn of DNA shorter than RP_o). Nonetheless, these models clearly show that β′SI3, the jaw, and βSI1 are positioned to clamp downstream DNA from +10 to +20.[10,11]

In both the RP_o and the transcription elongation complex models, the trigger loop, which connects to β′SI3 through flexible linkers, is unfolded. Intriguingly, using the folded form of the trigger loop[17] creates steric clash, leading to the proposal that the jaw/β′SI3 domain likely toggles between two positions as the trigger loop folds and unfolds with each cycle of NTP addition.[10,11] In addition, the RP_o model shows density for βSI1 near +15–20. The volume of this density increases relative to the partially disordered state in free *E. coli* RNAP,[68] supporting our hypothesis that βSI1 folds on binding downstream DNA.[13]

In-cleft interactions in the open complexes

Are these large conformational changes in the RNAP downstream machinery during the conversion of I_2 to RP_o connected/correlated with smaller-scale but very significant conformational changes in the active site and surrounding regions of the cleft? Differences in the MnO_4^- reactivities of bases on the downstream portion of the NT strand in I_2 and RP_o suggest that rearrangements in the NT strand are coupled to the formation of the downstream DNA clamp in the conversion of I_2 to RP_o.[13–15,59,102] Details of the communication between the cleft and the clamp/jaw remain to be established. This communication, if established in the isomerization steps, would likely persist in regulating the subsequent steps of the transcription cycle. Does the sequence/length of the discriminator region (−6 to +1) affect the "repositioning" of the NT strand and, thus, the stability of RP_o at different promoters? Are I_2 and I_3 functional in transcription? Do the different open complexes (one unstable and one stable) play distinct functional roles of open complex formation (e.g., Which open complex is the target of the stress sensor protein DksA?[85])? How do the contacts established in the late steps of RP_o formation affect promoter escape? Addressing these questions will likely bring new challenges and surprises, and will advance our understanding of the regulation of these late steps as a function of promoter sequence and solution conditions.

The role of σ^{70} in promoter interactions, open complex formation, and early RNA synthesis

Promoter recognition

Structures of RNAP holoenzyme from *Thermus aquaticus* and *Thermus thermophilus* reveal that σ^{70} consists of several independently folded domains ($\sigma_{1.2}$, σ_2, σ_4, and likely the N-terminal ~60 residues of $\sigma_{1.1}$ as well[103]) connected by flexible linkers ($\sigma_{3.2}$ and the highly negatively charged C-terminal residues of $\sigma_{1.1}$). Recent evidence reveals that the structure of the free sigma factors is compact, and that $\sigma_{1.1}$ and σ_4 interact.[103] This interaction may lead to the observed autoinhibition of promoter DNA binding by free primary sigma factors.[104] Autoinhibition has not been observed for interaction with the double-stranded promoter DNA of free sigma factors lacking region 1.1 by deletion[105] or naturally,[106,107] or for free σ^{70} interacting with the NT strand of promoter DNA.[108]

In the holoenzyme, the interactions between sigma and the core cover an extended surface area of both proteins[58] (see Fig. 1a): various regions of σ^{70}, including the linkers, interact with core RNAP, thus affording the bound σ^{70} considerable structural rigidity compared to the free σ^{70}. For example, $\sigma_{2.3/2.4}$ (−10 recognition) and $\sigma_{4.2}$ (−35 recognition) are now at a fixed orientation with respect to each other, imposing a rather strict limitation on the length of the spacer DNA separating the −10 and −35 elements (17±1 bp). Indeed, there are conditions where only σ_2 of sigma is bound to RNAP in a paused transcription complex where the RNA is 16–17 nt—long enough to have displaced σ_4 from the β flap. At this point, −10-like and −35-like elements that occur just downstream and upstream of the start site of transcription, respectively, can be jointly contacted by the now flexibly tethered σ_2 and σ_4 even if they are separated by only 1 bp.[109]

Redundancy in σ^{70} promoter elements

The −10 and −35 elements together constitute the classical prokaryotic promoter. However, in addition to the −10 and −35 elements mentioned above, there are several other regions of promoter DNA contacted by σ^{70} or the α subunit (see the text above). This raises the question of whether other pairs of promoter elements can also constitute an active promoter. Promoter DNA melting initiates within the −10 element, rendering this the most important and the least dispensable of the promoter elements. Can other regions substitute for the −35 element? This has indeed been found to be the case (reviewed by Hook-Barnard and Hinton[7]). Notable

among these are the extended −10 (TG+−10) and UP+−10, which has so far been only characterized as an artificial construct. Under *in vitro* conditions, promoter DNA strand separation has been observed with DNA containing just the −10 element.[110] Even the combination of −35+TG has been found to be active, if provided with an A+T-rich region that has the all important −11A and −7T in the NT strand.[110]

Stringent promoter requirements for holoenzymes containing "minor" sigma factors

In addition to the primary ("housekeeping") sigma factor, most bacteria have one or more minor sigma factors that can impart to RNAP the ability to transcribe genes whose products allow cells to deal with various types of stress. An example is the heat shock sigma factor σ^H (usually called σ^{32}),[111,112] which helps to mitigate the cytoplasmic consequences of transient exposure to higher-than-optimal temperatures. Promoter recognition occurs through a −35 element similar to that of σ^{70} promoters and an extended −10 element of which the sequence deviates considerably from that recognized by $E\sigma^{70}$.[113] Importantly, while $E\sigma^{70}$ is relaxed in its ability to recognize promoter DNA sequences, $E\sigma^{32}$ is found to be stringent in requiring promoters with consensus or near-consensus sequences. The reason for the difference is the DNA melting region (2.3) of σ^{70}: a broader promoter-recognition spectrum for $E\sigma^{32}$ is generated by replacing just two residues of region 2.3 with homologous aromatic σ^{70} residues.[114] These experiments indicate not only that σ^{32} is intrinsically melting deficient but also that this deficiency can be overcome by the use of consensus promoters. Indeed, similar behavior has also been observed for σ^{70} mutants rendered melting deficient by substitutions in region 2.3: if provided with a very good promoter, RNAP containing a defective σ^{70} could still form an open complex.[79] Observations similar to those described above for σ^{32} have also been made for RNAP containing σ^{28}, another minor *E. coli* sigma factor.[114]

Nucleation of DNA melting: Role of conserved residues in σ^{70}

The −11A element in the NT strand of *E. coli* σ^{70} promoters plays an important role in the formation of an open complex, as detailed in a number of studies.[76,115–119] It is the most conserved base pair in the −10 element.[120] Substitution of the −11A element by 2-AP[117] or loss of the base at this position[121,122] has much larger negative effects than at other positions. At the upstream end of the −10 region, the A-T base pair at −12 remains double-stranded in RP_o. A clear indication that strand separation is initiated at the −11 position is derived from studies demonstrating a correlation between the reduced stability of base pairing at −11 bp and the ability of promoter DNA to be melted by RNAP.[116]

Residues Y421, Y425, F427, T429, Y430, Y425, W433, and W434 (*E. coli* σ^{70} numbering; see Fig. 4) are nearly invariant among 53 sigma factors analyzed[124] and are found within a short distance of each other and of the −11A base[49] (at the upstream single-stranded–double-stranded DNA boundary over σ_2 in Fig. 1a). T429, Y430, and W433 are near the double-stranded–single-stranded junction of the model DNA cocrystallized with the RNAP, consistent with their involvement in the initiation of DNA melting. Evidence for the vital roles of Y430 and W433 includes deleterious effects of substitutions on open complex formation[125–127] and their high extent of conservation. Compelling evidence has been obtained for an interaction of Y430 with −11A.[79] However, it is likely that Y430 (and W433) additionally also recognizes other bases, and that other amino acid residues recognize −11A. Support for such a network of interactions, with multiple roles for the participating groups, is derived from two sets of observations. First, the effects of various substitutions for Y430 and W433 on the ability of RNAP to form stable promoter complexes are evident even in the absence of the base at −11 of the NT strand.[122] Second, a variant RNAP containing a multiply substituted σ^{70} (alanine

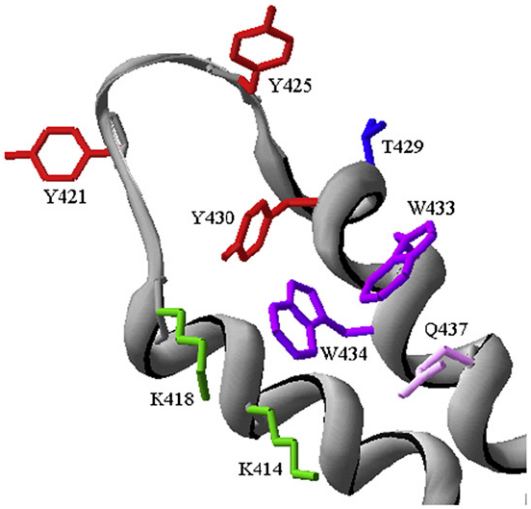

Fig. 4. Structure of region 2.3 of σ^{70}.[123] In the N→C direction are helix 13 (lower helix), a loop, and helix 14 (upper helix). The side chains of K414, K418, Y425, T429, Y430, W433, W434, and Q437 (Lys, green; Tyr, red; Thr, blue; Trp, purple; Gln, pink) stick out towards the viewer from approximately the same face of the protein, where they can interact with promoter DNA. Y430 has been shown to stack with −11A of the −10 region.[79] Y421 sticks out in another direction but may be able to interact with DNA. The structures of the *T. aquaticus* and *T. thermophilus* $\sigma_{2.3}$ are very similar.[50,57]

substitutions for F427, Y430, W433, and W434) has retained the capability for sequence recognition at −11.[40]

It is envisioned that interactions with basic amino acid residues of σ_2 (including K414 and K418) anchor the promoter DNA to the surface of the RNAP in the closed complex.[128] Promoter DNA melting is likely initiated by the rotation (or "flipping") of −11A out of the DNA helix[76,129] so that it now can stack onto Y430. Aromatic amino acid residues T429[122] and W433[125,128] of $\sigma_{2.3}$ are likely closely involved in the actual process of flipping −11A out of the DNA helix.[79] From the flipped −11A, DNA strand opening would proceed in downstream direction to +2.

Melting 12–14 bp of duplex DNA at 25–37 °C in the absence of RNAP has a large enthalpic cost (~70–84 kcal), thought to arise primarily from base unstacking and not from breaking of hydrogen bonds.[130] However, the activation enthalpy for opening the transcription bubble at λP_R is approximately half as large (~34 kcal[59]). Preservation of intrastrand stacking and favorable interactions between bases on the NT strand and aromatic residues (see the text above) likely reduce the enthalpic cost of opening the initiation bubble. Evidence for this hypothesis comes from the lack of permanganate reactivity of thymines at −7 and −10, indicating that these bases, once opened, either remain stacked with their neighbors or interact with residues in σ_2. In addition, NT strand bases (−4, −3, and +2) may be partially stacked in I_2, since they are only half as MnO_4^- reactive as in RP_o.[36] Thus, the model presented above is perhaps best described as "bind–bend/flip–melt," followed by clamping. In this model, RNAP is an active participant in achieving DNA strand separation: both RNAP-induced DNA bending and the side chains of amino acids T429, Y430, and W433[79,122,125,126,128] facilitate the DNA strand separation reaction. In addition, various elements in the cleft, such as the "fork loop 2" of β and multiple "tracks" of positively charged residues, appear positioned to capture and stabilize the open state via interactions with the DNA phosphate backbone.[15]

Sigma release *versus* retention

Since the discovery of sigma factor over 40 years ago,[18,19] it has been thought that an obligate late step in transcription initiation was the release of sigma factor from RNAP. However, two studies[131,132] clearly demonstrated the retention of σ^{70} in transcription complexes beyond the early phases of transcription. Compelling evidence for the presence of σ^{70} in elongation complexes was obtained from both FRET experiments and analysis of the proteins in the complexes. Global analysis techniques[133] provided support for σ^{70} retention *in vivo*. While these studies did not demonstrate a function for the retained sigma factor, other work demonstrated that the retained sigma factor was instrumental in generating a promoter-proximal pause of transcription during the synthesis of bacteriophage (i.e., bacterial virus) λ mRNAs. Such a pause is vital for endowing the transcription complexes with the ability to read through termination signals.[134] Subsequently, similar sigma-dependent pausing was demonstrated for the transcription of various bacterial genes as well[53–55,135,136] (see also a recent review by Artsimovitch[135]).

The role of the retained σ^{70} is to recognize −10-like sequences on the NT strand of the transcribed DNA.[53–55,134] The interaction of σ^{70} with such regions was found to be similar to its interaction with bases of the −10 element on the NT DNA. The interaction may be further strengthened by contacts to G-C base pairs[52] positioned similarly to the G-C base pairs at −5 and −6, which are contacted by $\sigma_{1.2}$ in RNAP–promoter complexes.[27,28] Interestingly, the −10-like element may not be absolutely necessary, although it greatly stabilizes the interaction of sigma with the transcription complex.[54] Recently described was an atypical example of σ^{70}-dependent pausing where the −10 element was lacking but pausing was shown to be dependent on the TG sequence of the extended −10 element of the actual promoter and contacts to a C-rich region at +2–6 of the NT strand.[136] The σ^{70}–NT DNA contacts serve to lock the elongating RNAP in position, thus impeding further movement of the elongation complex. The duration of the sigma-facilitated pause is reduced by GreA and GreB proteins *in vitro*, and evidence is consistent with this also being the case *in vivo*.[53–55,137] This behavior is indicative of backtracking of transcription complexes during the pause.

It remains to be established whether sigma retention is characteristic of most, or all, σ^{70} promoters. It may be that retained σ^{70} is not detected unless the NT strand has the proper −10-like sequence for σ^{70}-dependent pausing. Then the interaction of σ^{70} with the transcription complex would be stabilized,[53,54] further delaying sigma release. Alternatively, delayed release of sigma may be a promoter-specific event for which the signals have not yet been discerned. In support of the former, experimental evidence has been obtained for a stochastic release of sigma factor from the elongating complex, which would manifest itself as a certain half-life (estimated to be on the order of 5 s) for σ^{70} release from RNAP during transcription elongation.[138] A decreased stability of the RNAP–sigma complex is consistent with the suggestion that sigma's attachment to core during elongation differs from that in free holoenzyme.[53] Indeed, it is likely that the nascent RNA, by the time it has reached 10 nt in length, will have pried loose the contacts

between $\sigma_{3.2}$ and the β' subunit in the RNA channel of the core enzyme. At about 16 nucleotides, the contacts between σ_4 and the core enzyme β flap would be disrupted. Just the interaction of σ_2 and the clamp helices would then tie the sigma to the transcription complex. Interestingly, recent work indicates that release of σ_2 from the β' clamp helices is required to load the elongation factors NusG[139] and rfaH,[140] which thus may play a role in the release of sigma from the elongation complex. Regardless of the nature of the contacts, sigma apparently is held in a position that allows it to scan the sequence of the NT strand.

Conclusion

Determination of high-resolution structures of free and promoter-bound holoenzymes, together with advances in our understanding of how salts and solutes interact with biopolymer surfaces and perturb biopolymer processes, has led to rapid progress in our understanding of the events of RNAP recruitment and promoter recognition to form the initial closed complex RP_c, and the massive conformational changes in RNAP and promoter DNA that occur to convert it to the most stable open complex RP_o. Challenges for the future include developing a molecular understanding of how the start site region is opened and how the T strand is placed in the active site; how conformational changes in the cleft involving $\sigma_{1.1}$ and the downstream NT strand in the conversion of the initial open complex (I_2) to RP_o are sensed by the assembling downstream clamp/jaw apparatus; how upstream DNA trajectory and interactions with the αCTDs and the upstream clamp allow the entry of downstream duplex DNA into the cleft; and how all these steps of isomerization are regulated by DNA sequence, factors, ligands, and environmental variables in the response of the cell to changing growth conditions or stress.

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2011.01.018

Acknowledgements

We thank the reviewers for helpful comments, R. H. Ebright for conversations and communication of unpublished results, and our current and former colleagues for many fruitful collaborations. We thank Caroline Davis, Wayne Kontur, Theodore Gries, and Lisa Schroeder for their significant contributions to the mechanistic work reviewed here. We gratefully acknowledge Craig Bingman and Irina Artsimovitch for discussion of and assistance with Figs. 1 and 3 and Supplementary Fig. 1, and Olga Kourennaia for discussion of and assistance with Fig. 4. R.M.S. and M.T.R. acknowledge research support from National Institutes of Health grant GM23467.

References

1. Browning, D. F. & Busby, S. J. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65.
2. Haugen, S. P., Ross, W. & Gourse, R. L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6**, 507–519.
3. Wigneshweraraj, S., Bose, D., Burrows, P. C., Joly, N., Schumacher, J., Rappas, M. *et al.* (2008). Modus operandi of the bacterial RNA polymerase containing the σ^{54} promoter-specificity factor. *Mol. Microbiol.* **68**, 538–546.
4. deHaseth, P. L., Zupancic, M. L. & Record, M. T., Jr. (1998). RNA polymerase–promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* **180**, 3019–3025.
5. Gruber, T. M. & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466.
6. Paget, M. S. & Helmann, J. D. (2003). The σ^{70} family of sigma factors. *Genome Biol.* **4**, 203.
7. Hook-Barnard, I. G. & Hinton, D. M. (2007). Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Biol.* **1**, 275–293.
8. Borukhov, S. & Nudler, E. (2008). RNA polymerase: the vehicle of transcription. *Trends Microbiol.* **16**, 126–134.
9. Ghosh, T., Bose, D. & Zhang, X. (2010). Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol. Rev.* **34**, 611–627.
10. Hudson, B. P., Quispe, J., Lara-Gonzalez, S., Kim, Y., Berman, H. M., Arnold, E. *et al.* (2009). Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proc. Natl Acad. Sci. USA*, **106**, 19830–19835.
11. Opalka, N., Brown, J., Lane, W. J., Twist, K. A., Landick, R., Asturias, F. J. & Darst, S. A. (2010). Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol.* **8**.
12. Lane, W. J. & Darst, S. A. (2010). Molecular evolution of multisubunit RNA polymerases: structural analysis. *J. Mol. Biol.* **395**, 686–704.
13. Kontur, W. S., Saecker, R. M., Davis, C. A., Capp, M. W. & Record, M. T., Jr. (2006). Solute probes of conformational changes in open complex (RP_o) formation by *Escherichia coli* RNA polymerase at the λP_R promoter: evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RP_o. *Biochemistry*, **45**, 2161–2177.
14. Kontur, W. S., Saecker, R. M., Capp, M. W. & Record, M. T., Jr. (2008). Late steps in the formation of *E. coli* RNA polymerase–λP_R promoter open complexes:

characterization of conformational changes by rapid [perturbant] upshift experiments. *J. Mol. Biol.* **376**, 1034–1047.
15. Kontur, W. S., Capp, M. W., Gries, T. J., Saecker, R. M. & Record, M. T., Jr (2010). Probing DNA binding, DNA opening and assembly of downstream clamp/jaw in *Escherichia coli* RNA polymerase–λP_R promoter complexes using salt and the physiological anion glutamate. *Biochemistry*, **49**, 4361–4373.
16. Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr. & Saecker, R. M. (2007). Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **104**, 7833–7838.
17. Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. (2007). Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**, 157–162.
18. Burgess, R. R., Travers, A. A., Dunn, J. J. & Bautz, E. K. (1969). Factor stimulating transcription by RNA polymerase. *Nature*, **221**, 43–46.
19. Travers, A. A. & Burgess, R. R. (1969). Cyclic re-use of the RNA polymerase sigma factor. *Nature*, **222**, 537–540.
20. Lonetto, M., Gribskov, M. & Gross, C. A. (1992). The σ^{70} family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**, 3843–3849.
21. Gourse, R. L., Ross, W. & Gaal, T. (2000). UPs and downs in bacterial transcription initiation: the role of the α subunit of RNA polymerase in promoter recognition. *Mol. Microbiol.* **37**, 687–695.
22. Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980). *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell*, **20**, 269–281.
23. Barne, K. A., Bown, J. A., Busby, S. J. & Minchin, S. D. (1997). Region 2.5 of the *Escherichia coli* RNA polymerase σ^{70} subunit is responsible for the recognition of the 'extended -10' motif at promoters. *EMBO J.* **16**, 4034–4040.
24. Mitchell, J. E., Zheng, D., Busby, S. J. & Minchin, S. D. (2003). Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.* **31**, 4689–4695.
25. Sanderson, A., Mitchell, J. E., Minchin, S. D. & Busby, S. J. (2003). Substitutions in the *Escherichia coli* RNA polymerase σ^{70} factor that affect recognition of extended -10 elements at promoters. *FEBS Lett.* **544**, 199–205.
26. Keilty, S. & Rosenberg, M. (1987). Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**, 6389–6395.
27. Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I. *et al.* (2006). A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol. Cell*, **23**, 97–107.
28. Haugen, S. P., Berkmen, M. B., Ross, W., Gaal, T., Ward, C. & Gourse, R. L. (2006). rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell*, **125**, 1069–1082.
29. Jeong, W. & Kang, C. (1994). Start site selection at lacUV5 promoter affected by the sequence context around the initiation sites. *Nucleic Acids Res.* **22**, 4667–4672.
30. Liu, J. & 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**, 2938–2945.
31. Walker, K. A. & Osuna, R. (2002). Factors affecting start site selection at the *Escherichia coli* fis promoter. *J. Bacteriol.* **184**, 4783–4791.
32. Tang, Y., Murakami, K., Ishihama, A. & deHaseth, P. L. (1996). Upstream interactions at the λP_{RM} promoter are sequence nonspecific and activate the promoter to a lesser extent than an introduced UP element of an rRNA promoter. *J. Bacteriol.* **178**, 6945–6951.
33. Davis, C. A., Capp, M. W., Record, M. T., Jr. & Saecker, R. M. (2005). The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **102**, 285–290.
34. Ross, W. & Gourse, R. L. (2005). Sequence-independent upstream DNA-αCTD interactions strongly stimulate *Escherichia coli* RNA polymerase–lacUV5 promoter association. *Proc. Natl Acad. Sci. USA*, **102**, 291–296.
35. Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V. & Ebright, R. H. (2000). Structural organization of the RNA polymerase–promoter open complex. *Cell*, **101**, 601–611.
36. Gries, T. J., Kontur, W. S., Capp, M. W., Saecker, R. M. & Record, M. T., Jr. (2010). One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. *Proc. Natl Acad. Sci. USA*, **107**, 10418–10423.
37. Roberts, C. W. & Roberts, J. W. (1996). Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. *Cell*, **86**, 495–501.
38. Haugen, S. P., Ross, W., Manrique, M. & Gourse, R. L. (2008). Fine structure of the promoter–sigma region 1.2 interaction. *Proc. Natl Acad. Sci. USA*, **105**, 3292–3297.
39. Burr, T., Mitchell, J., Kolb, A., Minchin, S. & Busby, S. (2000). DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Res.* **28**, 1864–1870.
40. Cook, V. M. & deHaseth, P. L. (2007). Strand opening-deficient *Escherichia coli* RNA polymerase facilitates investigation of closed complexes with promoter DNA: effects of DNA sequence and temperature. *J. Biol. Chem.* **282**, 21319–21326.
41. Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S. & Ebright, R. H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, **314**, 1144–1147.
42. Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, **314**, 1139–1143.

43. Tchernaenko, V., Halvorson, H. R., Kashlev, M. & Lutter, L. C. (2008). DNA bubble formation in transcription initiation. *Biochemistry*, **47**, 1871–1884.
44. Deuschle, U., Kammerer, W., Gentz, R. & Bujard, H. (1986). Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *EMBO J.* **5**, 2987–2994.
45. Hsu, L. M., Cobb, I. M., Ozmore, J. R., Khoo, M., Nahm, G., Xia, L. *et al.* (2006). Initial transcribed sequence mutations specifically affect promoter escape properties. *Biochemistry*, **45**, 8841–8854.
46. Tang, G. Q., Roy, R., Bandwar, R. P., Ha, T. & Patel, S. S. (2009). Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase. *Proc. Natl Acad. Sci. USA*, **106**, 22175–22180.
47. Goldman, S. R., Ebright, R. H. & Nickels, B. E. (2009). Direct detection of abortive RNA transcripts *in vivo*. *Science*, **324**, 927–928.
48. Mooney, R. A., Darst, S. A. & Landick, R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell*, **20**, 335–345.
49. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme–DNA complex. *Science*, **296**, 1285–1290.
50. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S. & Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, **417**, 712–719.
51. Young, B. A., Anthony, L. C., Gruber, T. M., Arthur, T. M., Heyduk, E., Lu, C. Z. *et al.* (2001). A coiled-coil from the RNA polymerase β subunit allosterically induces selective nontemplate strand binding by σ^{70}. *Cell*, **105**, 935–944.
52. Ring, B. Z. & Roberts, J. W. (1994). Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317–324.
53. Brodolin, K., Zenkin, N., Mustaev, A., Mamaeva, D. & Heumann, H. (2004). The σ^{70} subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. *Nat. Struct. Mol. Biol.* **11**, 551–557.
54. Nickels, B. E., Mukhopadhyay, J., Garrity, S. J., Ebright, R. H. & Hochschild, A. (2004). The σ^{70} subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter. *Nat. Struct. Mol. Biol.* **11**, 544–550.
55. Hatoum, A. & Roberts, J. (2008). Prevalence of RNA polymerase stalling at *Escherichia coli* promoters after open complex formation. *Mol. Microbiol.* **68**, 17–28.
56. Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell*, **98**, 811–824.
57. Murakami, K. S., Masuda, S. & Darst, S. A. (2002). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*, **296**, 1280–1284.
58. Sharp, M. M., Chan, C. L., Lu, C. Z., Marr, M. T., Nechaev, S., Merritt, E. W. *et al.* (1999). The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized. *Genes Dev.* **13**, 3015–3026.
59. Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr, Capp, M. W. & Record, M. T., Jr. (2002). Kinetic studies and structural models of the association of *E. coli* σ^{70} RNA polymerase with the lambdaP(R) promoter: large scale conformational changes in forming the kinetically significant intermediates. *J. Mol. Biol.* **319**, 649–671.
60. Murakami, K. S., Masuda, S. & Darst, S. A. (2003). Crystallographic analysis of *Thermus aquaticus* RNA polymerase holoenzyme and a holoenzyme/promoter DNA complex. *Methods Enzymol.* **370**, 42–53.
61. Chen, J., Darst, S. A. & Thirumalai, D. (2010). Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proc. Natl Acad. Sci. USA*, **107**, 12523–12528.
62. Murakami, K. S. & Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. *Curr. Opin. Struct. Biol.* **13**, 31–39.
63. Belogurov, G. A., Vassylyeva, M. N., Sevostyanova, A., Appleman, J. R., Xiang, A. X., Lira, R. *et al.* (2009). Transcription inactivation through local refolding of the RNA polymerase structure. *Nature*, **457**, 332–335.
64. Ho, M. X., Hudson, B. P., Das, K., Arnold, E. & Ebright, R. H. (2009). Structures of RNA polymerase–antibiotic complexes. *Curr. Opin. Struct. Biol.* **19**, 715–723.
65. Tupin, A., Gualtieri, M., Leonetti, J. P. & Brodolin, K. (2010). The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site. *EMBO J.* **29**, 2527–2537.
66. Nagai, H. & Shimamoto, N. (1997). Regions of the *Escherichia coli* primary sigma factor σ^{70} that are involved in interaction with RNA polymerase core enzyme. *Genes Cells*, **2**, 725–734.
67. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N. *et al.* (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase–promoter open complex. *Cell*, **108**, 599–614.
68. Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G. & Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
69. Roe, J. H., Burgess, R. R. & Record, M. T., Jr. (1984). Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λP_R promoter. *J. Mol. Biol.* **176**, 495–522.
70. Buc, H. & 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**, 2712–2723.
71. Roe, J. H., Burgess, R. R. & Record, M. T., Jr. (1985). Temperature dependence of the rate constants of the *Escherichia coli* RNA polymerase–λP_R promoter interaction. Assignment of the kinetic steps corresponding to protein conformational change and DNA opening. *J. Mol. Biol.* **184**, 441–453.
72. Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr, Capp, M. W., Saecker, R. M. & 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**, 741–756.
73. Kovacic, R. T. (1987). The 0 °C closed complexes between *Escherichia coli* RNA polymerase and two promoters, T7-A3 and lacUV5. *J. Biol. Chem.* **262**, 13654–13661.
74. Cowing, D. W., Mecsas, J., Record, M. T., Jr. & Gross, C. A. (1989). Intermediates in the formation of the open complex by RNA polymerase holoenzyme containing the sigma factor σ^{32} at the groE promoter. *J. Mol. Biol.* **210**, 521–530.
75. Schickor, P., Metzger, W., Werel, W., Lederer, H. & Heumann, H. (1990). Topography of intermediates in transcription initiation of *E. coli*. *EMBO J.* **9**, 2215–2220.
76. Helmann, J. D. & 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**, 5959–5967.
77. Rogozina, A., Zaychikov, E., Buckle, M., Heumann, H. & Sclavi, B. (2009). DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 °C and results in the accumulation of an off-pathway intermediate. *Nucleic Acids Res.* **37**, 5390–5404.
78. Sclavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M. & 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. USA*, **102**, 4706–4711.
79. Schroeder, L. A., Gries, T. J., Saecker, R. M., Record, M. T., Jr, Harris, M. E. & deHaseth, P. L. (2009). Evidence for a tyrosine–adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of *Escherichia coli* RNA polymerase to promoter DNA. *J. Mol. Biol.* **385**, 339–349.
80. Spassky, A., Kirkegaard, K. & 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**, 2723–2731.
81. Tsodikov, O. V., Craig, M. L., Saecker, R. M. & Record, M. T., Jr. (1998). Quantitative analysis of multiple-hit footprinting studies to characterize DNA conformational changes in protein–DNA complexes: application to DNA opening by $E\sigma^{70}$ RNA polymerase. *J. Mol. Biol.* **283**, 757–769.
82. Murakami, K. S., Masuda, S. & Darst, S. A. (2003). Crystallographic analysis of Thermus aquaticus RNA polymerase holoenzymeand a holoenzyme/promoter DNA complex. *Methods Enzymol.* **370**, 42–53.
83. Mecsas, J., Cowing, D. W. & Gross, C. A. (1991). Development of RNA polymerase–promoter contacts during open complex formation. *J. Mol. Biol.* **220**, 585–597.
84. Bartlett, M. S., Gaal, T., Ross, W. & Gourse, R. L. (1998). RNA polymerase mutants that destabilize RNA polymerase–promoter complexes alter NTP-sensing by *rrn* P1 promoters. *J. Mol. Biol.* **279**, 331–345.
85. Li, X. Y. & McClure, W. R. (1998). Characterization of the closed complex intermediate formed during transcription initiation by *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **273**, 23549–23557.
86. Rutherford, S. T., Villers, C. L., Lee, J. H., Ross, W. & Gourse, R. L. (2009). Allosteric control of *Escherichia coli* rRNA promoter complexes by DksA. *Genes Dev.* **23**, 236–248.
87. Craig, M. L., Suh, W. C. & Record, M. T., Jr. (1995). HO and DNase I probing of $E\sigma^{70}$ RNA polymerase–λP_R promoter open complexes: Mg^{2+} binding and its structural consequences at the transcription start site. *Biochemistry*, **34**, 15624–15632.
88. Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A. *et al.* (1993). A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science*, **262**, 1407–1413.
89. Busby, S. & Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293**, 199–213.
90. Dethiollaz, S., Eichenberger, P. & Geiselmann, J. (1996). Influence of DNA geometry on transcriptional activation in *Escherichia coli*. *EMBO J.* **15**, 5449–5458.
91. Giladi, H., Igarashi, K., Ishihama, A. & Oppenheim, A. B. (1992). Stimulation of the phage lambda pL promoter by integration host factor requires the carboxy terminus of the α-subunit of RNA polymerase. *J. Mol. Biol.* **227**, 985–990.
92. Meng, W., Belyaeva, T., Savery, N. J., Busby, S. J., Ross, W. E., Gaal, T. *et al.* (2001). UP element-dependent transcription at the *Escherichia coli* rrnB P1 promoter: positional requirements and role of the RNA polymerase α subunit linker. *Nucleic Acids Res.* **29**, 4166–4178.
93. Chen, H., Tang, H. & Ebright, R. H. (2003). Functional interaction between RNA polymerase α subunit C-terminal domain and σ^{70} in UP-element- and activator-dependent transcription. *Mol. Cell*, **11**, 1621–1633.
94. Ross, W., Schneider, D. A., Paul, B. J., Mertens, A. & Gourse, R. L. (2003). An intersubunit contact stimulating transcription initiation by *E. coli* RNA polymerase: interaction of the α C-terminal domain and sigma region 4. *Genes Dev.* **17**, 1293–1307.
95. Lee, D. J., Busby, S. J. & Lloyd, G. S. (2003). Exploitation of a chemical nuclease to investigate the location and orientation of the *Escherichia coli* RNA polymerase α subunit C-terminal domains at simple promoters that are activated by cyclic AMP receptor protein. *J. Biol. Chem.* **278**, 52944–52952.
96. Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S. & Gross, C. (1998). Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J. Mol. Biol.* **284**, 1353–1365.
97. Kolb, A., Igarashi, K., Ishihama, A., Lavigne, M., Buckle, M. & Buc, H. (1993). *E. coli* RNA polymerase, deleted in the C-terminal part of its α-subunit, interacts differently with the cAMP–CRP complex

at the lacP1 and at the galP1 promoter. *Nucleic Acids Res.* **21**, 319–326.
98. Spolar, R. S. & Record, M. T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science*, **263**, 777–784.
99. Yin, Y. W. & Steitz, T. A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science*, **298**, 1387–1395.
100. Ederth, J., Artsimovitch, I., Isaksson, L. A. & Landick, R. (2002). The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing. *J. Biol. Chem.* **277**, 37456–37463.
101. Zaychikov, E., Denissova, L. & Heumann, H. (1995). Translocation of the *Escherichia coli* transcription complex observed in the registers 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble". *Proc. Natl Acad. Sci. USA*, **92**, 1739–1743.
102. Brodolin, K., Zenkin, N. & Severinov, K. (2005). Remodeling of the σ^{70} subunit non-template DNA strand contacts during the final step of transcription initiation. *J. Mol. Biol.* **350**, 930–937.
103. Schwartz, E. C., Shekhtman, A., Dutta, K., Pratt, M. R., Cowburn, D., Darst, S. & Muir, T. W. (2008). A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem. Biol.* **15**, 1091–1103.
104. Camarero, J. A., Shekhtman, A., Campbell, E. A., Chlenov, M., Gruber, T. M., Bryant, D. A. *et al.* (2002). Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc. Natl Acad. Sci. USA*, **99**, 8536–8541.
105. Dombroski, A. J., Walter, W. A., Record, M. T., Jr, Siegele, D. A. & Gross, C. A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor σ^{70} exhibit specificity of binding to promoter DNA. *Cell*, **70**, 501–512.
106. Chen, Y. F. & Helmann, J. D. (1995). The *Bacillus subtilis* flagellar regulatory protein σ^{D}: overproduction, domain analysis and DNA-binding properties. *J. Mol. Biol.* **249**, 743–753.
107. Sevim, E., Gaballa, A., Belduz, A. O. & Helmann, J. D. (2011). DNA-binding properties of the *Bacillus subtilis* and *Aeribacillus pallidus* AC6 σ^{D} proteins. *J. Bacteriol.* **193**, 575–579.
108. Mekler, V., Pavlova, O. & Severinov, K. (2011). The interaction of *Escherichia coli* RNA polymerase σ^{70} subunit with promoter elements in the context of free σ^{70}, RNA polymerase holoenzyme, and the β'–σ^{70} complex. *J. Biol. Chem.* **286**, 270–279.
109. Devi, P. G., Campbell, E. A., Darst, S. A. & Nickels, B. E. (2010). Utilization of variably spaced promoter-like elements by the bacterial RNA polymerase holoenzyme during early elongation. *Mol. Microbiol.* **75**, 607–622.
110. Niedziela-Majka, A. & Heyduk, T. (2005). *Escherichia coli* RNA polymerase contacts outside the -10 promoter element are not essential for promoter melting. *J. Biol. Chem.* **280**, 38219–38227.
111. Grossman, A. D., Erickson, J. W. & Gross, C. A. (1984). The htpR gene product of *E. coli* is a sigma factor for heat-shock promoters. *Cell*, **38**, 383–390.
112. Grossman, A. D., Straus, D. B., Walter, W. A. & Gross, C. A. (1987). σ^{32} synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. *Genes Dev.* **1**, 179–184.
113. Koo, B. M., Rhodius, V. A., Campbell, E. A. & Gross, C. A. (2009). Dissection of recognition determinants of *Escherichia coli* σ^{32} suggests a composite -10 region with an 'extended -10' motif and a core -10 element. *Mol. Microbiol.* **72**, 815–829.
114. Koo, B. M., Rhodius, V. A., Nonaka, G., deHaseth, P. L. & Gross, C. A. (2009). Reduced capacity of alternative sigmas to melt promoters ensures stringent promoter recognition. *Genes Dev.* **23**, 2426–2436.
115. Guo, Y. & Gralla, J. D. (1998). Promoter opening via a DNA fork junction binding activity. *Proc. Natl Acad. Sci. USA*, **95**, 11655–11660.
116. Heyduk, E., Kuznedelov, K., Severinov, K. & Heyduk, T. (2006). A consensus adenine at position -11 of the nontemplate strand of bacterial promoter is important for nucleation of promoter melting. *J. Biol. Chem.* **281**, 12362–12369.
117. Lim, H. M., Lee, H. J., Roy, S. & Adhya, S. (2001). A "master" in base unpairing during isomerization of a promoter upon RNA polymerase binding. *Proc. Natl Acad. Sci. USA*, **98**, 14849–14852.
118. Matlock, D. L. & 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**, 12274–12283.
119. Tsujikawa, L., Strainic, M. G., Watrob, H., Barkley, M. D. & DeHaseth, P. L. (2002). RNA polymerase alters the mobility of an A-residue crucial to polymerase-induced melting of promoter DNA. *Biochemistry*, **41**, 15334–15341.
120. Shultzaberger, R. K., Chen, Z., Lewis, K. A. & Schneider, T. D. (2007). Anatomy of *Escherichia coli* σ^{70} promoters. *Nucleic Acids Res.* **35**, 771–788.
121. Li, X. Y. & McClure, W. R. (1998). Stimulation of open complex formation by nicks and apurinic sites suggests a role for nucleation of DNA melting in *Escherichia coli* promoter function. *J. Biol. Chem.* **273**, 23558–23566.
122. Schroeder, L. A., Choi, A. J. & deHaseth, P. L. (2007). The -11A of promoter DNA and two conserved amino acids in the melting region of σ^{70} 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**, 4141–4153.
123. Malhotra, A., Severinova, E. & Darst, S. A. (1996). Crystal structure of a σ^{70} subunit fragment from *E. coli* RNA polymerase. *Cell*, **87**, 127–136.
124. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O. *et al.* (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell*, **9**, 527–539.
125. Fenton, M. S., Lee, S. J. & Gralla, J. D. (2000). *Escherichia coli* promoter opening and -10 recognition: mutational analysis of σ^{70}. *EMBO J.* **19**, 1130–1137.
126. Panaghie, G., Aiyar, S. E., Bobb, K. L., Hayward, R. S. & de Haseth, P. L. (2000). Aromatic amino acids in

region 2.3 of *Escherichia coli* σ^{70} participate collectively in the formation of an RNA polymerase–promoter open complex. *J. Mol. Biol.* **299**, 1217–1230.
127. Juang, Y. L. & 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**, 1470–1488.
128. Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J. & deHaseth, P. L. (2001). Different roles for basic and aromatic amino acids in conserved region 2 of *Escherichia coli* σ^{70} in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase–promoter complexes. *J. Biol. Chem.* **276**, 31891–31896.
129. Helmann, J. D. & Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**, 839–872.
130. Holbrook, J. A., Capp, M. W., Saecker, R. M. & Record, M. T., Jr. (1999). Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry*, **38**, 8409–8422.
131. Bar-Nahum, G. & Nudler, E. (2001). Isolation and characterization of σ^{70}-retaining transcription elongation complexes from *Escherichia coli*. *Cell*, **106**, 443–451.
132. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W. & Ebright, R. H. (2001). Translocation of σ^{70} with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, **106**, 453–463.
133. Raffaelle, M., Kanin, E. I., Vogt, J., Burgess, R. R. & Ansari, A. Z. (2005). Holoenzyme switching and stochastic release of sigma factors from RNA polymerase *in vivo*. *Mol. Cell*, **20**, 357–366.
134. Ring, B. Z., Yarnell, W. S. & Roberts, J. W. (1996). Function of *E. coli* RNA polymerase sigma factor σ^{70} in promoter-proximal pausing. *Cell*, **86**, 485–493.
135. Artsimovitch, I. (2008). Post-initiation control by the initiation factor sigma. *Mol. Microbiol.* **68**, 1–3.
136. Stepanova, E., Wang, M., Severinov, K. & Borukhov, S. (2009). Early transcriptional arrest at *Escherichia coli* rplN and ompX promoters. *J. Biol. Chem.* **284**, 35702–35713.
137. Marr, M. T. & Roberts, J. W. (2000). Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol. Cell*, **6**, 1275–1285.
138. Shimamoto, N., Kamigochi, T. & Utiyama, H. (1986). Release of the sigma subunit of *Escherichia coli* DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. *J. Biol. Chem.* **261**, 11859–11865.
139. Mooney, R. A., Schweimer, K., Rosch, P., Gottesman, M. & Landick, R. (2009). Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* **391**, 341–358.
140. Belogurov, G. A., Mooney, R. A., Svetlov, V., Landick, R. & Artsimovitch, I. (2009). Functional specialization of transcription elongation factors. *EMBO J.* **28**, 112–122.