Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective

Andrey Feklistov,1 Brian D. Sharon,2,3,4 Seth A. Darst,1 and Carol A. Gross3,4,5

1The Rockefeller University, 1230 York Avenue, New York, NY 10065; email: darst@rockefeller.edu
2Graduate Group in Biophysics, 3Department of Microbiology and Immunology, 4California Institute of Quantitative Biology, and 5Department of Cell and Tissue Biology, University of California, San Francisco, California 94158; email: cgrossucsf@gmail.com

Abstract

Transcription initiation is the crucial focal point of gene expression in prokaryotes. The key players in this process, sigma factors (σs), associate with the catalytic core RNA polymerase to guide it through the essential steps of initiation: promoter recognition and opening, and synthesis of the first few nucleotides of the transcript. Here we recount the key advances in σ biology, from their discovery nearly 50 years ago to the most recent progress in understanding their structure and function at the atomic level. Recent data provide important structural insights into the mechanisms whereby σs initiate promoter opening. We discuss both the housekeeping σs, which govern transcription of the majority of cellular genes, and the alternative σs, which direct RNA polymerase to specialized operons in response to environmental and physiological cues. The review concludes with a genome-scale view of the extracytoplasmic function σs, the most abundant group of alternative σs.

Keywords

sigma, RNA polymerase, transcription, gene expression
### OVERVIEW

The discovery that bacterial RNA polymerase (RNAP) required a specialized subunit, sigma (σ, for selectivity), to recognize promoters ushered in a new era in regulatory biology of seeking to understand how the cell ensures fidelity of transcription initiation. This finding presaged the discovery that all multisubunit RNAPs require separate initiation functions. In retrospect, this division of tasks is not surprising. To perform the central cellular function of transcription, RNAP must efficiently traverse a large fraction of the genome without preference for DNA sequence during the elongation phase of transcription. Nevertheless, the beginning of transcription units must be precisely specified and their use tightly regulated. Cells have solved this problem by encoding transcription start sites with specific DNA sequences (promoters; Figure 1a) and allowing the cellular multisubunit RNAP to locate these sequences through a dissociable, sequence-specific DNA-binding protein that functions primarily at initiation.

Whereas initiation functions are divided among many proteins in eukaryotes, the prokaryotic σ performs all initiation functions, making σ an ideal model for studying initiation. Bacteria have a single essential housekeeping σ that promotes transcription of thousands of genes, and many alternative σs, each promoting transcription of a specialized gene repertoire necessary for coping with stress or development. Given the critical mechanistic and regulatory roles of σs, it is not surprising that they have been intensely studied since their discovery. We begin this review with a series of historical vignettes recounting key discoveries about the σ70 family of proteins. The body of the review is dedicated to understanding the current status of important broad trends in σ biology: how σs perform their critical functions of directing RNAP to promoters and initiating DNA-strand opening; how σs further integrate regulatory events in transcription; and how alternative σs have modified the mechanism of initiation to enhance their specificity. We end by considering the breadth and reach of the most abundant, smallest, and most divergent

---

**Contents**

OVERVIEW............................................................... 358

KEY MOMENTS IN σ HISTORY........................................... 359

- Discovery of σ70.................................................. 359
- The Bacterial Promoter Takes Shape.......................... 360
- σ Binds Promoter DNA............................................ 360
- The σ70 Family of Proteins...................................... 360

HOW DOES σ PERFORM ITS FUNCTIONS?...................... 361

- Directing the Catalytic Core of the RNAP to Transcription Start Sites........ 361
- Strand Separation of the DNA: The First Step in Formation of the Transcription Bubble.................... 363
- Suppressing Nonspecific Transcription Initiation...................... 365
- Interaction with Activators...................................... 365
- Abortive Initiation/Promoter Clearance.......................... 365
- Promoter Proximal Pausing During Early Elongation................... 367

A MODIFIED INITIATION STRATEGY ALLOWS THE DIVERGENT ALTERNATIVE σS TO PERFORM THEIR FUNCTIONS......... 368

A GENOME-SCALE VIEW OF THE ECF σS.......................... 369

PERSPECTIVE......................................................... 370
alternative $\sigma$s [called extracytoplasmic function (ECF) or Group 4 $\sigma$s]. Overall, we highlight the mechanistic, genomic, and regulatory roles of the $\sigma^{70}$ family of proteins.

### KEY MOMENTS IN $\sigma$ HISTORY

#### Discovery of $\sigma^{70}$

As a graduate student in the Watson-Gilbert group, Dick Burgess purified RNAP to homogeneity. There were few examples of multisubunit enzymes at the time, making it critical to show that all copurifying polypeptides were part of the enzyme itself. This led Dick to test additional purification steps, showing that purified RNAP retained the ability to initiate RNA chains on calf thymus DNA after passage through a phosphocellulose column. However, a fellow graduate student, Jeffrey Roberts, studying initiation on a bacteriophage $\lambda$ template, found the enzyme to be inactive (R. Burgess and J. Roberts, personal communication). Reconstituting the RNAP with other fractions from the phosphocellulose chromatography identified $\sigma^{70}$, the *Escherichia coli* housekeeping $\sigma$. The $\sigma$ worked during initiation, as it increased the number of chains (as judged by incorporation of $^{32}$P-$\gamma$ phosphate, present only at the 5' end of the chain). Ekhardt Bautz and Joshua Dunn at Rutgers performed similar experiments, and the two groups published a single paper documenting these experiments (9); other groups also identified a dissociable subunit but did not identify its mode of action (45, 49). Additional experiments demonstrated that $\sigma^{70}$ was catalytic: Once initiation occurred, $\sigma^{70}$ could dissociate from one RNAP molecule and promote
initiation by a differentially marked (rifampicin-resistant) RNAP (45, 49, 107). Thus, RNAP has two distinct forms: holoenzyme ($\alpha_2\beta'\beta\omega+\sigma$), for initiation, and core ($\alpha_2\beta\omega$), for elongation.

**The Bacterial Promoter Takes Shape**

With $\sigma$ in hand, the focus shifted to determining the DNA-binding regions that marked the promoter, a difficult feat in the absence of DNA sequencing. Undeterred, several groups initiated RNAP transcription from a strong promoter, isolated the short fragment protected from DNase digestion, and inferred the promoter sequence from that of complementary RNA using nearest-neighbor sequencing of oligos produced by digestion with various RNases. Based on the sequences of several strong promoters, David Pribrnow and Heinz Schaller identified the -10 motif (Figure 1a) (74, 90). However, RNAP could not rebind this minimal fragment, suggesting that there were missing recognition determinants. Digestion with restriction enzymes provided a longer fragment to sequence and identified the -35 motif as a second common promoter motif (Figure 1a). The importance of the -10 and -35 motifs was confirmed by showing that previously isolated mutations that altered expression from the *lac* promoter had nucleotide changes in the -10 and -35 motifs of the promoter. The tour de force was the sequencing of the entire *lac* control region by Bill Reznikoff and colleagues, identifying binding sites for Lac repressor and cyclic-AMP binding protein, as well as mutations in the -35 motif of the promoter (23, 78).

**$\sigma$ Binds Promoter DNA**

Although $\sigma$ was required for initiation, $\sigma^{70}$ alone was unable to recognize DNA (8, 117). This led to the suggestion that $\sigma$ might alter RNAP conformation, thereby exposing latent DNA-binding determinants in polymerase. In fact, the opposite is true—latent DNA-binding domains of $\sigma$ are exposed when $\sigma$ binds to RNAP. The realization that alternative $\sigma$s recruit RNAP to different -10 and -35 promoter sequences was difficult to reconcile with the idea that $\sigma$ exposed latent DNA-binding domains in RNAP. It was extremely unlikely that RNAP carried the multiple DNA-binding determinants required for recognizing many alternative promoter sequences (57). Then, several groups found mutations in $\sigma$ that either compensated for poor promoters or altered promoter preference (99, 112, 123). The finding that the N-terminus of the housekeeping $\sigma$ was autoinhibitory for DNA binding began to address why $\sigma$s might not bind DNA even though they carried the DNA-recognition determinants (24). Ultimately, structural data suggested that isolated $\sigma$s likely have intramolecular interactions that occlude their DNA-binding determinants (101, 102). Binding to RNAP splays these domains apart and displays them in the correct orientation and spacing for promoter binding (67–69). Requiring $\sigma$ to bind RNAP prior to DNA recognition ensures that the binding affinity of each $\sigma$ for RNAP influences the relative abundance of different holoenzymes and thus the spectrum of promoters recognized.

**The $\sigma^{70}$ Family of Proteins**

The initial report describing $\sigma$ suggested that the cell might have additional $\sigma$-like positive regulators with different specificity, possibly explaining changes in transcription patterns, such as those observed late in the growth cycle of lytic phages (9). Indeed, new RNAP-binding proteins appeared during *Bacillus subtilis* sporulation (31, 32) (e.g., SpoIIG, now called SigE) or after infection by the lytic phages SPO1 (105) or T4 (43, 44, 114, 115), and these altered the in vitro transcription preferences of RNAP, suggesting a role for $\sigma$s in development. Then, an *E. coli* $\sigma$-like positive regulator (HtpR, now called $\sigma^{12}$) that directed RNAP to heat shock gene promoters was
identified, indicating that σ-like proteins were general across bacteria and controlled both stress and developmental responses (30). A true understanding of the σ universe awaited the onset of DNA sequencing. First came the realization that the essential housekeeping σs of E. coli and B. subtilis were highly conserved and had strong sequence similarity to SigE and σ32, hinting that they were a protein family (29, 34, 104). As additional σs were sequenced, this idea was confirmed, and alternative σs were subdivided into Group 2 σs (most similar to housekeeping Group 1 σs) and the more divergent Group 3 σs (55). The most divergent group of alternative σs, the ECF σs (Group 4), was not identified until 1994 (56). These small σs initially escaped detection via homology searching because they lacked one conserved σ domain (Figure 1b). With the onset of large-scale sequencing of microbial genomes, thousands of σs have been identified, with ECF σs the most abundant. Except for Mycoplasma, all bacteria have multiple alternative σs.

HOW DOES σ PERFORM ITS FUNCTIONS?

The σs perform two main functions: (a) directing the catalytic core of the RNAP to appropriate transcription start sites, and (b) initiating strand separation of the double-helical DNA as the first step in the formation of the transcription bubble. Indeed, to be classified as a σ factor, a protein must perform both tasks. In addition to these central functions, a number of additional σ activities have been identified. A dozen years of elegant structural studies combined with incisive biochemical and biophysical approaches have improved our understanding of the structural basis of all of the functions of σ.

Directing the Catalytic Core of the RNAP to Transcription Start Sites

Prokaryotic promoters are marked by five motifs: the -35, extended -10, -10, and discriminator motifs recognized by σ7 and the UP element recognized by the C-terminal domain (CTD) of the RNAP α-subunit. A great deal is known about how each of these motifs is recognized, and here we outline the σ recognition strategy. The σ proteins are composed of a variable number of structured domains connected by flexible linkers. The simplest σs have two domains (Group 4 or ECF σs: σ5, σ3, σ4), some have three domains (Group 3 σs: σ2, σ1, σ3), and the housekeeping σs have four domains (σ1, σ2, σ3, σ4) (Figure 1b). Except for σ2, each domain has DNA-binding determinants: σ4, -35 motif; σ1, extended -10 motif; σ2, -10 and discriminator motifs (Figure 2). The -35 motif is recognized by a canonical helix-turn-helix (HTH) DNA-binding domain within σ4 and has been visualized with high-resolution X-ray crystal structures of -35 motif/σ4 interactions for housekeeping (14) and ECF (51) σs (Figure 2). A long α-helix of σ1 recognizes the DNA major groove of the extended -10 motif (5, 88). The -35 and extended -10 motifs are recognized as double-stranded DNA and remain double stranded throughout the initiation process. The -10 and discriminator motifs are recognized upon strand separation (Figure 2) as non-template-strand, single-stranded DNA (ssDNA, see below). Normally, one of the double-stranded motifs (-35 or extended -10) is required for initiation (although see 27), and these are thought to anchor the holoenzyme in the appropriate position and orientation to then engage the -10 motif upon strand separation (26, 28).

Although all specific recognition of promoter sequence is mediated by σ, σ does not specifically bind promoters on its own. At least two mechanisms (that are not mutually exclusive) contribute to suppressing the DNA-binding capacity of free σs.

Conformational restriction. The structure of a full-length, isolated σ has not been determined, but it is thought that the σ domains interact with each other to maintain a relatively compact conformation that buries the DNA-binding determinants, or at least does not properly present
Figure 2

The lower-right image shows an *Escherichia coli* RNA polymerase (RNAP) open complex (RPo) model (78). The molecular surface of RNAP holoenzyme is shown, color-coded as follows: gray, α-subunits; light cyan, β; light pink, β'; green, σ [but with the nonconserved region (NCR) shown in gray]. A segment of the β-subunit is cut away to expose the active site channel. The DNA, shown as a phosphate-backbone worm, is color-coded: The nontemplate DNA strand is blue, the template strand is red, the -35 motif is brown, and the -10 motif is yellow. A nascent 5mer RNA transcript is shown in pink. The catalytic Mg\(^{2+}\) ion is shown as a red sphere. The boxed regions are magnified. The close-up view on the left shows the -35 motif/σ\(_4\) interaction. The -10 motif/σ\(_2\) interaction is shown in detail on the top.

them to interact with the widely spaced promoter motifs (Figure 2). This idea is supported by three key pieces of information: (a) Structural studies of a Group 3 σ bound to its inhibitory anti-σ factor reveal a compact σ conformation with buried DNA-binding determinants. Cross-linking results suggest that the isolated σ assumes a similar conformation as that bound to the anti-σ (101, 102). (b) FRET/LRET analyses show that the σ domains are relatively compact in free σ and move to a spread-out conformation upon formation of holoenzyme (10, 11). (c) Structural results show
that the core RNAP is responsible for positioning the σ domains at the appropriate orientation and spacing to interact with the promoter motifs (Figure 2) (69).

**Autoinhibitory restriction.** The N-terminal σ₁₁, found only in housekeeping σs (Figure 1b), suppresses recognition of promoter motifs by the isolated σ. Specific promoter recognition by σ alone is weak and difficult to detect even in the absence of σ₁₁; the presence of σ₁₁ suppresses that activity even further (12, 24, 91). We note that autoinhibition and conformational restriction may be linked; one way σ₁₁ might inhibit σ promoter-binding activity is by stabilizing a compact conformation of the σ that would be recalcitrant to promoter recognition.

**Strand Separation of the DNA: The First Step in Formation of the Transcription Bubble**

Once the RNAP holoenzyme locates promoters through recognition of double-stranded, upstream DNA motifs, RNAP unwinds 13 base pairs of the duplex DNA to form the transcription bubble. Remarkably, the bacterial enzyme, containing either a housekeeping or alternative σ, accomplishes this feat without the assistance of ATP hydrolysis. Current thinking suggests that σ₂-mediated capture of a flipped base (or bases) of the nontemplate-strand -10 motif underlies the strand separation process of all σs.

**Housekeeping σs.** Helmann & Chamberlin (34) noted a number of invariant basic and aromatic residues in σ₂ and proposed, by analogy with ssDNA- and RNA-binding proteins, that this region binds one of the DNA strands in the nascent bubble to stabilize the strand-separated state. Later studies showed that these invariant residues of σ are indeed critical for transcription-bubble formation (21, 40, 99, 112, 123), functioning as a sequence-specific, ssDNA-binding determinant specific for the nontemplate strand sequence of the -10 motif (T₁₋₁₂A₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋˓
Nucleation of promoter melting by housekeeping (left) and ECF (right) σ s. In both cases, promoter melting is triggered by the capture of a single base in a complementary pocket of σ. The pocket is formed by a specificity loop and additional invariant residues. (left) A−11 is stacked between R-246 and Y-253 of *Thermus aquaticus* σA2 [Protein Data Bank (PDB) ID 3UGO; 40]. (right) C−10 is stacked between W-73 and R-65 of *Escherichia coli* σE (PDB ID 4LUP; 75).

The prevailing view was that the -10 motif must be recognized sequence specifically first as duplex DNA by conserved region 2.4 in σA2, and that the non-template-strand bases of the -10 motif are recognized after strand separation by the critical melting residues mostly in conserved region 2.3 of σA2 (22, 35, 37). However, structural modeling combined with biochemical results suggests that specific recognition of the -10 motif in duplex form is unlikely (28). Thus, -10 motif sequence recognition and initiation of strand separation are coupled. The molecular details of -10 motif/σA2 interaction were fully corroborated by a subsequent structure in the context of the entire RNAP holoenzyme and a partial promoter construct (122). This study also delineated contacts between exposed bases of the discriminator (just downstream of the -10 motif) and σA2 that stabilize the RNAP open complex (RPo; Figure 1c, 27, 33).

**Alternative σs.** The biochemical, biophysical, and structural details of the promoter-opening process for alternative σs have not been studied to nearly the same extent as those for housekeeping σs, but available data suggest that the overall process is similar (17, 18, 60). Recent work structurally characterizing an ECF σ (*E. coli* σW) in complex with its cognate non-template-strand -10 motif (T−11:G−12; T−11:C−10:A−9:A−8:A−7; 13) provides important insights into promoter melting by this highly divergent family. Like housekeeping σs, a nearly absolutely conserved base (analogous to housekeeping σA2-10 motif A−11) is flipped out of the single-strand base stack and buried in the σA2 pocket corresponding to the A−11 pocket in housekeeping σA2 (Figure 3). However, in the case of σW, the flipped-out base is a C (C−10). There is no structural equivalent of the T−7 pocket of housekeeping σA2; only one base is flipped and specifically captured in a protein pocket. Remarkably, the recognition of the identity of this all-important flipped base in the -10 motif is governed by a modular specificity loop (Figure 3). For instance, the -10 motif consensus for promoters recognized by *B. subtilis* ECF σW suggests that A is the flipped base. Swapping the σW specificity-loop into σA2 gave rise to a hybrid σ that recognizes A as the flipped base rather than C (13). This modularity allows for the rapid evolutionary diversification of ECF σs and their cognate -10 motifs.
Suppressing Nonspecific Transcription Initiation

In addition to having the critical function of imparting sequence-specific DNA-binding capacity to the RNAP holoenzyme, σ has been reported to suppress the capacity of the core RNAP to initiate transcription at nonspecific (nonpromoter) DNA sites (97). The large active site channel for binding DNA, located between the pincers of the crab-claw-like structure of core RNAP, is basic (positively charged; 19) and is expected to promote nonspecific interactions between nucleic acids and core RNAP, increasing the chances for nonspecific initiation. In addition to suppressing promoter recognition by free σ, the N-terminal σ1.1 domain of housekeeping σs may also suppress nonspecific DNA binding in the RNAP active site cleft. In the RNAP holoenzyme, the highly acidic σ1.1 situates itself in the RNAP active site cleft (2, 61; Figure 4a). Remarkably, upon formation of RPo, loading of promoter DNA into the RNAP active site cleft displaces σ1.1 out of the channel (61) (Figure 4a). A long, unstructured linker connects σ1.1 to the rest of σ (this linker is 37 residues in E. coli σ70), facilitating the large movements of σ1.1 (2). The highly dynamic nature of σ1.1 in the promoter opening process explains why mutations and truncations of σ1.1 can have significant effects on the kinetics of RPo formation (110, 116). These findings suggest that σ1.1 acts as a gatekeeper for the RNAP active site, sitting in the active site cleft and blocking access to the active site for random DNA sequences. Displacement of σ1.1 requires promoter sequences that interact specifically with σ to initiate the proper pathway for RPo formation. The mechanism of displacement is unknown. Only housekeeping σs harbor σ1.1 (Figure 1b): Suppression of nonspecific DNA binding may be especially important for housekeeping σs given their increased DNA melting capacity (see below).

Interaction with Activators

Activation of basal transcription is a major regulatory strategy in bacteria. Most bacterial transcription activators bind to DNA operators upstream of the promoter -35 motif and activate transcription through direct contacts with a target subunit of the RNAP, creating new protein-protein and protein-DNA contacts to help recruit RNAP to promoters (7). When the activator binds very near the upstream part of the -35 motif, σ4 is frequently its interaction target (100, 106; reviewed in 25).

Mutations in σ4 that selectively disrupt the function of an activator without affecting basal transcription (positive control mutants) cluster on the upstream face of σ4 and are well positioned to interact with the downstream subunit of the dimeric activators (14, 39). The interaction of σ with some transcription regulators is likely to be more complex. For instance, GrgA in Chlamydia trachomatis interacts with the housekeeping σ through a nonconserved insertion in the σ between conserved regions 1.2 and 2.1 (3).

Abortive Initiation/Promoter Clearance

In RNAP holoenzyme, the σ1 and σ4 domains are linked by an approximately 50-residue segment that includes a loop-like feature comprising σ conserved region 3.2 (53; Figure 1b). This segment winds through the RNAP active site channel and occupies the exit path for the nascent RNA (69, 109). Within this linker, a hairpin-like loop (dubbed the σ finger; 122) protrudes toward the RNAP active center (Figure 2). The σ finger plays an important role in initiation by helping the RNAP bind the initiating substrate. A 7-residue deletion of the σ finger caused a two-orders-of-magnitude increase of the Km for the 3′-initiating NTP (50). Although this region of σ cross-links to initiating nucleotide analogs (50, 93), crystal structures of initiating complexes suggest that the effect of the
σ finger on initiating NTP binding is indirect through the preorganization of the DNA template strand (122). Notably, the 3.2 linker is much shorter in alternative σs (15) (Figure 4b), pointing to a simplified initiation mechanism in this group of σs with reduced melting capacity.

Once RNAP catalyzes the first phosphodiester bond, the continued translocation and catalytic activity of the RNAP extends the 3′ end of the nascent transcript. Analysis of RNAP holoenzyme
crystal structures indicates that a transcript of only four to five nucleotides in length will encounter the σ finger in its path (69; Figure 2). At each step, the elongating RNA chain must either displace σ out of its path or dissociate from the complex in a process termed abortive initiation. The structure-based proposal that σ plays a major role in abortive initiation is supported by extensive functional evidence:

1. Substitution of two conserved residues near the beginning of the σ<sub>1</sub>-σ<sub>4</sub> linker (P-504 and S-506 in E. coli σ<sup>30</sup>) led to a dramatic reduction in the ratio of abortive to full-length transcripts (92).
2. Deletion of a seven-residue segment at the tip of the σ finger (the part of the σ finger encountered first by the elongating transcript) led to a relative decrease in the ratio of short abortive transcripts (≤10 nucleotides) to longer ones, as longer transcripts are now required for an encounter with σ (30).
3. Complete deletion of the σ<sub>1</sub>-σ<sub>4</sub> linker led to the nearly complete elimination of abortive transcripts (69).

Structural analysis indicates that the nascent RNA transcript emerges from the RNA exit channel when it reaches a length of ~16 nucleotides. This corresponds to the RNA length that results in displacement of both the σ<sub>1</sub>-σ<sub>4</sub> linker and σ<sub>4</sub> from their normal positions on RNAP holoenzyme (70). Abortive transcripts beyond the length of 16 nucleotides are only rarely observed (38), supporting a model in which the transition from the initiation to the elongation phase of transcription involves a direct competition between the elongating RNA transcript and the σ<sub>1</sub>-σ<sub>4</sub> linker and σ<sub>4</sub> for the same binding site within core RNAP. The competition gives rise to abortive initiation and also sets up the staged displacement of segments of σ from the RNAP, facilitating promoter escape (67, 69).

### Promoter Proximal Pausing During Early Elongation

Even after displacement of the σ<sub>1</sub>-σ<sub>4</sub> linker and σ<sub>4</sub> and promoter escape, σ can still be retained by core RNAP during the beginning of elongation through interactions between σ<sub>2</sub> and the β<sup>′</sup> clamp helices. During elongation, σ undergoes stochastic release from RNAP, sometimes remaining dozens of base pairs downstream of the transcription start site (4, 6, 41, 66, 71, 83, 96). While bound to elongating RNAP, σ continues to scan the DNA for promoter-like elements and is able to recognize both -10- and -35-like sequences, thereby inducing transcriptional pausing (73, 83). The σ<sup>30</sup>-dependent promoter-proximal pause observed at λ phage late gene promoters is prerequisite for the recruitment of the phage-encoded protein Q that converts RNAP into a terminator-resistant form (85). A number of E. coli promoters have been shown to have σ<sup>30</sup>-dependent pauses, but their physiological significance is still unclear (6, 71, 73). Chromatin immunoprecipitation (ChIP) assays support stochastic release of σ from elongating RNAP in vivo (63, 64, 76, 77, 111).

**Figure 4**

(a) Views into the active site cleft of *Escherichia coli* RNAP holoenzyme (left; Protein Data Bank ID 4LK1; 78) and an RNAP open complex (RPo) model (right). In holoenzyme, σ<sub>1</sub> sits in the active site cleft but is displaced by the downstream DNA duplex in RPo. (b) Schematic views of RNAP holoenzymes. A version of this figure was first published in Reference 20 by Nature Publishing Group. The core RNAP is shown (gray) with active site channel (light gray). The disposition of σ domains (green) is also illustrated. The housekeeping σ (left column) triggers promoter melting by capturing two highly conserved bases of the -10 element, A<sub>C11</sub> and T<sub>C17</sub>, in RPo (bottom row). Nonspecific interactions of nucleic acids with the active site are prevented by σ<sub>1</sub> in the downstream duplex channel of holoenzyme (top row). Entry of promoter DNA in RPo displaces σ<sub>1</sub>. Alternative σs such as *E. coli* σ<sup>70</sup> (right column) lack σ<sub>1</sub> and employ a simplified strategy relying on a single pocket for capturing C<sub>-10</sub>. Abbreviation: NCR, nonconserved region.
A MODIFIED INITIATION STRATEGY ALLOWS THE DIVERGENT ALTERNATIVE σ5S TO PERFORM THEIR FUNCTIONS

The role of the essential housekeeping σ is to promote initiation at the vast bulk of bacterial promoters, regardless of their strength. Indeed, this hardwired promoter hierarchy is an essential level of control built into the cell during evolution. The ability to recognize and initiate at promoters very divergent from the consensus promoter also enables the cell to respond to activators, which boost expression from such promoters, facilitating a dynamic response to changing conditions. The housekeeping σs can perform these roles because they are able to initiate at promoters having only a subset of recognition motifs (reviewed in 37). Of the five promoter sequence motifs, two suffice for transcription, as housekeeping σs initiate either at promoters with -10 and -35 motifs (with only loose spacing requirements) or at those with -10 and extended -10 motifs (Figure 1a). The fourth motif, the discriminator, is modulatory. Furthermore, strong UP elements, recognized by the CTD of the α-subunit of RNAP (α-CTD) can substitute for the -35 promoter motif (62).

What features of housekeeping σs enable recognition of such diverse promoters?

All housekeeping σs harbor two protein pockets that specifically capture two flipped-out bases of the nontemplate strand, complemented by the contacts with exposed bases of the discriminator (Figure 4b). These structural features give housekeeping σs a powerful DNA-melting capacity, allowing them to function in the face of highly nonoptimal promoter sequences. As a consequence, as long as the -10 region is intact, housekeeping σs can use diverse upstream regions to anchor RNAP (UP element, -35, extended -10) and still accomplish promoter melting.

In stark contrast to the housekeeping σs with their generalized role, the divergent alternative σs [Group 3 and ECF (Group 4) σs] mount concerted, focused responses to specific environmental conditions. Two important characteristics of such responses are (a) their regulons are small, ranging from as few as two to three genes to as many as several hundred, thereby minimizing the metabolic cost of the response; and (b) the relative expression of regulon genes depends almost exclusively on a hardwired promoter hierarchy. How do these two σ groups achieve these critical goals? Based on the limited exploration performed thus far, it appears that a weakened melting capacity of these σs relative to the housekeeping σs plays a significant role in their increased specificity.

The Group 3 σs have the same three DNA-recognition domains (σ12-σ13) as the housekeeping σs (46, 47) but do not use the mix-and-match promoter strategy of the housekeeping σs. Instead, the two Group 3 σs studied thus far, E. coli σ32 (heat-shock σ) and σ28 (flagella σ), require all three recognition motifs (-35, extended -10, -10) as well as rigid spacing between them for promoter activity. As evidence, σ32 and σ28 promoters have a higher information content (18.3 and 21.3 bits, respectively) (72, 82) than σ70 promoters (6.5 bits) (98). These requirements sharply restrict the number of promoters per genome. Essentially all Group 3 σs lack one or more of the four critical melting residues, which form the A_11 (or equivalent) pocket in the housekeeping σs. We suggest that only promoters with a near-perfect match to the consensus can be melted. Consistent with this hypothesis, converting the missing melting residues in the Group 3 σs to the housekeeping consensus relaxed requirements for highly conserved sequence in the promoter-recognition motifs and for highly conserved spacing. Importantly, the melting-consensus alternative σs have improved ability to melt DNA, as measured by their ability to form a stable complex at 4°C in vitro on a fork-junction template, an excellent mimic of open-complex formation. As the two Group 3 σs investigated represent very divergent branches of the family, reduced melting proficiency is likely to play a role in restricting the number of sites recognized by most, if not all, Group 3 σs (48).

The ECF σs are two-domain proteins (Figure 1b) that recognize only the -10 and -35 motifs of their promoters, immediately limiting the number of potential promoters in the genome. Extensive
studies indicate that both promoter motifs separated by rigid spacing are essential for promoter recognition by $\sigma$, an E. coli ECF $\sigma$ (79). Moreover, unlike the housekeeping $\sigma$s, recognition of the UP element by the $\alpha$-subunit of RNAP cannot substitute for the -35 motif, although UP elements do enhance expression from weak promoters (80). This observation has been validated for many ECF $\sigma$s (81). Interestingly, at the time that this $\sigma$ group was identified, it was noted that the regions of $\sigma$ implicated in initiating melting and in recognizing the -10 motif ($\sigma$ regions 2.3–2.4) were highly divergent from housekeeping $\sigma$s. As we pointed out earlier, the ECF $\sigma$ group diverges from the housekeeping paradigm by harboring only a single protein pocket to capture a flipped base of the -10 motif (rather than the two pockets of the housekeeping $\sigma$s; Figures 3 and 4b; 13, 28). Use of only a single flipped-out base, recognized by a single protein pocket formed by reordering a protein loop in the $\sigma$, is also likely to contribute to a weakened melting capacity.

In summary, weakened ability to initiate strand opening underlies the requirement for consensusr promoter sequences in both Group 3 and ECF $\sigma$s. The requirement for precise $\sigma$ alignment on promoters may also contribute to the fact that these $\sigma$s have very few activators.

A GENOME-SCALE VIEW OF THE ECF $\sigma$S

ECF $\sigma$s are the most abundant and diverse group of $\sigma$s, but until recently the extent of this family was unclear. Thorsten Mascher and colleagues undertook the first comprehensive classification of ECF $\sigma$s. They manually curated 2,700 ECF $\sigma$s from hundreds of bacterial genomes and developed models to identify 43 phylogenetically distinct subgroups of ECF $\sigma$s (103). These models were additionally used to develop a web-based ECF $\sigma$ finder (http://ecf.g2l.bio.uni-goettingen.de:8080/ECFfinder/), enabling the inclusion of ECF $\sigma$s in the MiST database (108) (http://mistdb.com/), so that as new genomes are sequenced, their ECF $\sigma$s are added to the database. Genomes average 6 ECF $\sigma$s; Plesiocystis pacifica SIR-1 holds the current record, encoding 118 ECF $\sigma$s in its genome (http://mistdb.com/). This large number of ECF $\sigma$s per genome raises the question of how transcriptional space is partitioned among ECF $\sigma$s that co-occur in bacteria. Do these $\sigma$s cross-talk with each other, sharing some fraction of their promoters? Or are they orthogonal, unable to recognize the promoters of co-occurring ECF $\sigma$s? The co-occurrence of orthogonal ECF $\sigma$s would enable insulated expression of their respective regulons, whereas the co-occurrence of cross-talking ECF $\sigma$s would potentially enable $\sigma$s to cooperate in coordinated responses.

To answer this question, it is necessary to identify promoters for the phylogenetically distinct ECF $\sigma$ subgroups and then determine the cross-recognition capacity of these $\sigma$s. Promoters were initially identified in silico by exploiting the fact that most ECF $\sigma$s transcribe their own promoters. ECF $\sigma$s in the database were binned according to subgroup, and the DNA sequences upstream of the $\sigma$ genes in each subgroup were searched for overrepresented sequences to identify promoter motifs (81, 103). Currently, computational efforts have identified promoter models for 29 of 43 subgroups. Rhodius (81) used these data to make in silico predictions of orthogonality; i.e., as to whether any two promoter models were sufficiently similar that it was likely that a single ECF $\sigma$ could initiate from both promoter types. He found that most, but not all, promoters were orthogonal, suggesting that there is not much cross talk between ECF $\sigma$s in different subgroups. Orthogonality depended on both the -35 and -10 motifs, as significantly more cross talk was predicted when a single motif was used to determine orthogonality. Importantly, Rhodius also used high throughput methodologies to determine whether the predicted promoters were functional in vivo and whether their cognate $\sigma$s were able to recognize promoters from other subgroups, largely validating the in silico predictions.
This large-scale experimental examination of ECF σ specificity also uncovered underlying trends for genome organization and the partitioning of transcriptional space (Table 1). Analysis of ECF σ co-occurrence in genomes found that non-cross-talking σs tend to co-occur with more ECF σ subgroups than cross-talking σs do. For example, 6 of the 8 ECF σ subgroups found in Caulobacter crescentus and 5 out of 6 in Rhizobium leguminosarum are orthogonal subgroups. These findings suggest that insulation of transcriptional signaling pathways is typically beneficial for cells. Indeed, the most highly orthogonal ECF σs, subgroups 41 (113) and 42, not only are broadly distributed across bacteria but also often have multiple representative σs in a single genome. This could allow cells to develop complex cooperative responses within an overarching pathway that still remains insulated from the rest of the cell’s transcriptional programs. Streptomyces coelicolor exhibits an impressive example of this strategy, with nearly half of its 30 ECF σs belonging to groups 41 and 42. In striking opposition to the broad distribution of highly orthogonal σs, the most promiscuous ECFs appear to have more restricted distributions. ECF σ subgroup 25 is a striking example, as it activates promoters from a wide range of other subgroups but co-occurs with very few of these and is only found in a small number of genomes overall.

Despite the apparent preference for insulated transcriptional signaling, cells can also evolve complex regulatory networks by utilizing cross-talking σs. The ECF σs in B. subtilis, which all respond to envelope stress, are one example. Although each is a member of a different ECF subgroup, all exhibit cross talk and thus have partially overlapping regulons. B. subtilis can therefore orchestrate a complex response in which a subset of core genes are induced by all envelope stresses, whereas additional genes are only induced by specific types of envelope stress (reviewed in 59). Experimental determinations of promoter specificity also revealed another role for cross talk. There are several examples of ECF σs that do not transcribe their own promoters. However, these ECF σs and their promoters co-occur with an ECF σ from a different subgroup that transcribes the promoter of the first ECF σ, raising the possibility of σ-σ cascades as an additional regulatory scheme (81). Such cascades can be found in a variety of divergent organisms, including Flavobacterium spp., Mycobacterium spp., and Streptomyces spp., and could control complex cellular decisions similar to the cascade of alternative σs controlling sporulation in B. subtilis.

Although these studies have begun to address general properties of the ECF σs, interesting questions remain unanswered. For highly orthogonal σs, it will be important to determine whether specific regulon functions are retained across organisms. Likewise, for cross-talking σs it will be interesting to address specific roles of σ-σ cascades, as well as whether cross-talking σs generally co-occur as part of cooperative responses or are in fact insulated by additional genome-level mechanisms.

**PERSPECTIVE**

In the 45 years since the discovery of σs, tremendous progress has been made in understanding their role in the cell, and consequently in understanding the process of transcription initiation.

**Table 1** Number of orthogonal and cross-talking ECF groups and σs in representative organisms

<table>
<thead>
<tr>
<th>Organism</th>
<th>Orthogonal</th>
<th></th>
<th></th>
<th></th>
<th>Cross-talking</th>
<th></th>
<th></th>
<th></th>
<th>Unclassified</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of</td>
<td>No. of σs</td>
<td>No. of groups</td>
<td>No. of σs</td>
<td>No. of groups</td>
<td>No. of σs</td>
<td>No. of groups</td>
<td>No. of σs</td>
<td>No. of groups</td>
<td>No. of σs</td>
<td>No. of groups</td>
<td>No. of σs</td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caulobacter crescentus</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhizobium leguminosarum</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streptomyces coelicolor</td>
<td>3</td>
<td>15</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This large-scale experimental examination of ECF σ specificity also uncovered underlying trends for genome organization and the partitioning of transcriptional space (Table 1). Analysis of ECF σ co-occurrence in genomes found that non-cross-talking σs tend to co-occur with more ECF σ subgroups than cross-talking σs do. For example, 6 of the 8 ECF σ subgroups found in Caulobacter crescentus and 5 out of 6 in Rhizobium leguminosarum are orthogonal subgroups. These findings suggest that insulation of transcriptional signaling pathways is typically beneficial for cells. Indeed, the most highly orthogonal ECF σs, subgroups 41 (113) and 42, not only are broadly distributed across bacteria but also often have multiple representative σs in a single genome. This could allow cells to develop complex cooperative responses within an overarching pathway that still remains insulated from the rest of the cell’s transcriptional programs. Streptomyces coelicolor exhibits an impressive example of this strategy, with nearly half of its 30 ECF σs belonging to groups 41 and 42. In striking opposition to the broad distribution of highly orthogonal σs, the most promiscuous ECFs appear to have more restricted distributions. ECF σ subgroup 25 is a striking example, as it activates promoters from a wide range of other subgroups but co-occurs with very few of these and is only found in a small number of genomes overall.

Despite the apparent preference for insulated transcriptional signaling, cells can also evolve complex regulatory networks by utilizing cross-talking σs. The ECF σs in B. subtilis, which all respond to envelope stress, are one example. Although each is a member of a different ECF subgroup, all exhibit cross talk and thus have partially overlapping regulons. B. subtilis can therefore orchestrate a complex response in which a subset of core genes are induced by all envelope stresses, whereas additional genes are only induced by specific types of envelope stress (reviewed in 59). Experimental determinations of promoter specificity also revealed another role for cross talk. There are several examples of ECF σs that do not transcribe their own promoters. However, these ECF σs and their promoters co-occur with an ECF σ from a different subgroup that transcribes the promoter of the first ECF σ, raising the possibility of σ-σ cascades as an additional regulatory scheme (81). Such cascades can be found in a variety of divergent organisms, including Flavobacterium spp., Mycobacterium spp., and Streptomyces spp., and could control complex cellular decisions similar to the cascade of alternative σs controlling sporulation in B. subtilis.

Although these studies have begun to address general properties of the ECF σs, interesting questions remain unanswered. For highly orthogonal σs, it will be important to determine whether specific regulon functions are retained across organisms. Likewise, for cross-talking σs it will be interesting to address specific roles of σ-σ cascades, as well as whether cross-talking σs generally co-occur as part of cooperative responses or are in fact insulated by additional genome-level mechanisms.

**PERSPECTIVE**

In the 45 years since the discovery of σs, tremendous progress has been made in understanding their role in the cell, and consequently in understanding the process of transcription initiation.
The two fundamental properties of $\sigma$s, identifying promoters and initiating strand separation of the DNA, are central to prokaryotic physiology. On the one hand, these properties enable housekeeping $\sigma$s to carefully fine-tune transcription initiation at the majority of promoters, where they not only identify promoter sequence but also orchestrate several crucial checkpoints along the pathway of open-complex formation and monitor the abortive initiation process, promoter escape, and in some cases early elongation. On the other hand, the proliferation of alternative $\sigma$s, with slightly divergent properties, has enabled the creation of critical stress and developmental responses; many of these responses have been identified and extensively studied.

What are the next big questions? We see two major areas that will be fleshed out in the coming years. First, increasingly sophisticated technologies for structural studies will enable us to follow the process of initiation through all of its complex steps at molecular resolution. Currently, we have single snapshots of parts of the process that are used to infer a causal set of events. However, in the next several years, we will have multiple structural snapshots complemented by results of state-of-the-art solution methods (42), which will allow us to follow the path of events, including understanding the complex dynamic rearrangements of $\sigma$ during initiation. The mechanistic details of $\sigma$ association with core RNAP and dissociation from transcription complexes entering the elongation phase, as well as the nature of transient intermediates of the promoter-opening pathway (87) are among the most important questions for future research.

Second, there will be extensive genome-scale and evolutionary investigations of the alternative $\sigma$s. Genome-scale studies will address how transcriptional space is partitioned among the multitude of alternative $\sigma$s encoded in most bacterial genomes, a study that is especially important for genomes with multiple ECF $\sigma$s. The evolutionary studies will address the mechanisms for maintaining orthogonality/cross-talking relationships among $\sigma$s and explore whether the trajectory of $\sigma$ regulon evolution is sensitive either to the orthogonality of the parental $\sigma$ or to the $\sigma$s that commonly co-occur with the parental $\sigma$.

DISCLOSURE STATEMENT
The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
We thank Brian Bae for sharing unpublished information. S.A.D. is supported by the National Institutes of Health. C.A.G is supported by the National Institutes of Health. B.D.S is supported by National Institutes of Health training grant T32GM008284.

LITERATURE CITED
5. Barne KA, Bown JA, Busby SJW, Minchin SD. 1997. Region 2.5 of the *Escherichia coli* RNA polymerase σ70 subunit is responsible for the recognition of the “extended-10” motif at promoters. *EMBO J.* 16(13):4034–40


Changes may still occur before final publication online and in print


