Escherichia coli RNA Polymerase ($E\sigma^{70}$), Promoters, and the Kinetics of the Steps of Transcription Initiation

M. THOMAS RECORD, JR., WILLIAM S. REZNIKOFF, MARIA L. CRAIG, KRISTI L. McQUADE, AND PAULA J. SCHLAX

54

OVERVIEW

Initiation Rates and Gene Expression

Gene expression in *Escherichia coli* begins with the promoter-specific binding of the enzyme RNA polymerase to initiate transcription of the $>10^3$ DNA operons into mRNA. The amount of gene product synthesized is primarily determined by the rate of productive initiation. In vivo, productive initiation rates span a range of approximately 4 orders of magnitude. The maximum observed initiation rate of ~1 transcript per s for individual rRNA promoters at high growth rates (144) corresponds to $\sim 1.5 \times 10^3$ transcripts per generation. At the other extreme, an initiation rate of 10^{-4} transcripts per s at a lower growth rate corresponds to production of 1 transcript per generation.

What are the principal determinants of the rate of transcription initiation? What molecular interactions and what kinetic steps on the pathway (mechanism) between uncomplexed and initiating states of promoter and RNA polymerase determine the hierarchy of initiation rates for different operons? What are the roles of intrinsic variables (e.g., promoter sequence) and

extrinsic variables (e.g., the concentrations of RNA polymerase and repressors, the extent of DNA supercoiling, as well as cytoplasmic or in vitro solution conditions) in the regulation of the rates of these steps and the overall rate of initiation? Even after two decades of intensive study these questions remain at the frontiers of research.

Steps in the Process of Initiation

Transcription initiation is a multistep process involving several classes of steps, as illustrated by the summary mechanism (*A*) in Fig. 1.

Within each class of steps (I through IV) in mechanism (*A*), the detailed mechanism (i.e., the full set of intermediates and the order of their appearance) is currently being investigated by kinetic and structural methods, reviewed in subsequent sections. For example, several key intermediates in the conversion of the initial (closed) binary complex to the final (open) binary complex (class II) have recently been identified. Research on this exciting frontier is continuing. However, even when all these important details are known, summary mechanism (*A*) will still remain a useful general characterization of the overall molecular events of transcription initiation, just as the classical Michaelis-Menten kinetic mechanism ($E + S \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P$) remains a useful summary of the mechanism of interaction of any enzyme (E) with a substrate (S) to yield product (P).

Class I: Reversible Initial Specific Binding. RNA polymerase (R) holoenzyme, defined as the 1:1 complex of the core polymerase (subunit composition, $\alpha_2\beta\beta'$) with the appropriate σ (specificity) subunit, binds to the double-helical DNA of the promoter site (P), forming an initial "closed" complex designated RP_{c1}. ("Closed" indicates that the promoter DNA remains entirely double-helical in this complex.)

Class II: Reversible Conformational Changes Driven by Binding Free Energy To Form the Final Binary Open Complex. Initially, RP_{c1} is converted to another closed state (designated RP_{c2}), in which RNA polymerase R interacts with the promoter DNA downstream as well as upstream of the transcription start site. The intermediate complex RP_{c2} is poised for the key event of this series of steps, namely, the reversible "opening" of a specific 10- to 15-bp region of DNA at the start site of transcription. For some if not all promoters, formation of a stable binary open complex is driven entirely by binding free energy and requires no coupled hydrolysis of nucleotides or other free energy inputs. To date, two open complexes have been characterized (designated RP_{o1} and RP_{o2}). The region of the transcription start site (+1) opens in the conversion of RP_{o1} to RP_{o2} and the complex is then poised to bind the initiating ribonucleotide.

Class III: Reversible Binding of Initiating Ribonucleotides (NTP). The nucleoside triphosphate (NTP) complementary to the template strand at the open start site (+1) binds to form the first of a series of ternary initiation complexes at the promoter, here collectively designated RP_{init}. Each of these may advance by binding to the next ribonucleotide (specified by the DNA template sequence) and by catalysis of covalent bond formation to the previous ribonucleotide, or may reverse by cleavage to remove the terminal ribonucleotide. Alternatively, RP_{init} may revert to RP_{o2} by release of the entire nascent oligo RNA chain, a cycle called abortive initiation.

Class IV: Transition to Elongation (Promoter Clearance/Escape). Once a 7- to 12-nucleotide RNA is synthesized, the σ subunit is released. At this point, no specific interactions with the promoter DNA remain, and the switchover from initiation to processive elongation by core RNA polymerase occurs.

Assumption: a Common Mechanism for all Promoters

The central premise of this article is that the steps and types of intermediate complexes on the path (i.e., the mechanism) by which the initial closed complex converts to the final binary open complex and then to a transcribing complex are common to all promoters transcribed by $E\sigma^{70}$ holoenzyme (and probably to other *E. coli* RNA polymerases as well). We assume that promoter sequence and other intrinsic and extrinsic variables affect the rate constants of these steps and the relative stabilities and populations of the intermediates, but do not cause fundamental changes in the mechanism. For example, the global and local extents of DNA supercoiling are important physiological variables which in general affect the thermodynamics (i.e., equilibria) of DNA opening. Effects of supercoiling on the kinetics of open complex formation and transcription initiation are generally attributed to the stabilization of RP_{c1} or the facilitation of kinetically significant isomerization steps involved in converting RP_{c1} to RP_{o2}, under the assumption that RP_{c1} remains an integral part of the mechanism on both linear and supercoiled DNA (see below). An alternative possibility (for which there is no evidence, and which is inconsistent with this assumption) would be that negative supercoiling facilitates opening of A-T-rich regions of the promoter DNA in the absence of RNA polymerase and thereby changes the mechanism, eliminating the need for closed complexes and allowing RNA polymerase to bind directly to open DNA, as occurs in "bubble duplexes" (2, 45).

The challenge to the experimentalist is twofold: (i) to determine the important intermediates and steps in the classes defined by the proposed mechanism (*A*), and (ii) to correlate the effects of DNA sequence and extrinsic variables on initiation rates with the rates and rate constants of these steps. This chapter focuses primarily on the two groups of mechanistic steps (initial binding, isomerization) involved in forming the final binary open complex at the promoter. Relevant aspects of the nucleotide-binding and catalytic steps which made up the subsequent two groups of processes in initiation are also discussed.

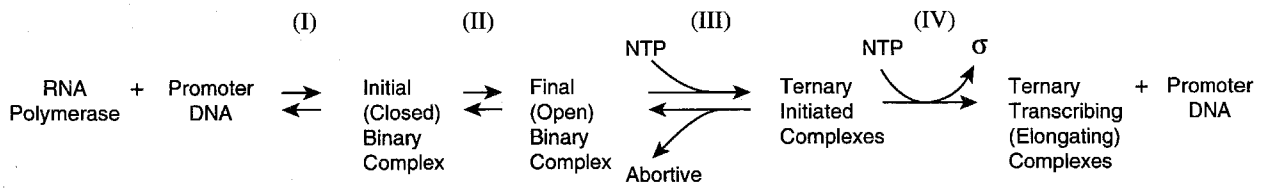

FIGURE 1 Mechanism (*A*).

Kinetic Concepts

How is the rate of formation of transcribing complexes related to reactant concentrations and to the rates and rate constants of the individual steps in the mechanism? For mechanism (*A*) or any more detailed version thereof, analyzed in the steady state, the rate of production of transcribing complexes per promoter (designated by *V*, the initiation velocity in transcripts per second) is given by an algebraic expression of the form:

$$\frac{V}{V_{max}} = \theta = \frac{[R]_F}{K_m + [R]_F} \quad (1)$$

In equation 1, θ is the overall fractional occupancy of the promoter by RNA polymerase in all complexes in the steady state. Fractional occupancies range from 0 to 1; from equation 1, θ is determined by the concentration of unbound (free) active RNA polymerase holoenzyme ($[R]_F$) and by the quantity designated K_m, a mechanism-specific collection of rate constants for the individual steps of classes I through IV in mechanism (*A*) above. The quantity designated V_{max} is the "maximum" initiation velocity obtainable for the promoter and conditions being investigated; V_{max} is the velocity at full occupancy ($\theta = 1$), achieved when $[R]_F >> K_m$. For a specified $[R]_F$, a promoter with a higher K_m has a lower steady-state occupancy by *R;* this steady-state definition of occupancy includes both closed and open binary and ternary complexes in mechanism (*A*).

Both the form and terminology of equation 1 are of course identical to those used in a completely different context to discuss the steady-state velocity of a noncooperative enzyme-catalyzed reaction. Equation 1, used to analyze in vitro initiation kinetic data for coliphage T7 and T3 RNA polymerases (references, e.g., 96, 139, 140, 187), should provide a general framework to discuss the rate of initiation of transcription both in vivo and also in multiround transcription assays in vitro, as long as the process begins with a bimolecular binding event and ends with regeneration of the binding site (promoter DNA). Equation 1 also is applicable to initial rate data in both reversible and irreversible promoter-binding kinetic assays (see below).

From equation 1, the rate of initiation at an individual promoter site (in transcripts per second) is determined by the free polymerase concentration $[R]_F$ and the mechanism-specific collections of sequence- and environment-dependent rate constants which constitute V_{max} and K_m. Multiplication of this intrinsic rate by the number of copies of that promoter per cell (or by the molar concentration of that promoter in solution) yields initiation rates in transcripts per cell per second (or moles of transcript per liter per second).

From mechanism (*A*) and the corresponding velocity equation 1, several generalizations with regard to initiation rates and their regulation can be made. (These are quantified using the steady-state approximation in a subsequent section.)

(i) A high rate of initiation is achieved when, for each intermediate complex, the rate constant of the reverse (dissociation direction) step is small in comparison to that of the next forward step, so that the forward direction is essentially irreversible. In this case, V_{max} is determined primarily by the slowest forward step, according to the "bottleneck" principle of chemical kinetics (and of traffic flow).

(ii) The highest initiation rate observed in vivo (~1 transcript s^{-1}) may correspond to the situation where the kinetics of the final two classes of steps in mechanism (*A*), namely, the regeneration of the promoter-binding site by downstream movement of the transcription complex in conjunction with very early steps in transcription elongation, are rate determining. In fact, since elongation rates are in the range of 50 to 100 nucleotides s^{-1} (138) and since the transcribing complex must move 30 to 60 bp downstream from the start site to permit another RNA polymerase to bind to the promoter, the overall rate of this promoter regeneration step is probably no greater than ~1 to 3 s^{-1}, very close to the highest observed initiation rate (144).

(iii) Alternatively, the rate of the initial bimolecular binding event to form RP_{cl} (the forward direction of step I in mechanism A) may not exceed 1 s^{-1} and therefore may be rate determining for the most rapidly initiating promoters. The rate of this step is proportional to the free RNA polymerase concentration $[R]_F$. In vitro measurements (as well as theoretical estimates) of the rate constant for this step yield values which are sufficiently large (2×10^8 to 3×10^8 M^{-1} s^{-1} in vitro [22, 27, 169]) to make the rate of this step in excess of 1 s^{-1} in vivo (and therefore faster than the maximum observed initiation velocity) *if* the in vivo $[R]_F$ is 10 nM. (Estimates of $[R]_F$ in vivo typically range from 10 to $\gtrsim$ 100 nM.) Since nonspecific binding of polymerase is thought to be significant in vivo, then facilitating mechanisms (e.g., diffusional sliding on adjacent nonspecific DNA) may extend the DNA target (229), thereby increasing the rate constant of the initial binding step. This facilitating mechanism is only effective for promoters at which the initial closed complex is stable to dissociation on the time scale of its isomerization. For those promoters, the existence of such a facilitating mechanism would permit the in vivo $[R]_F$ to be less than 10 nM, without being the rate-limiting quantity in transcription initiation. In any case, the rate of association to form RP_{cl} is determined by diffusional parameters and $[R]_F$ and is independent of promoter sequence. (All other forward and reverse steps in initiation mechanism (*A*) may depend on promoter sequence.)

(iv) The initiation rate is reduced and the possibilities for regulation of initiation rate by DNA sequence and/or extrinsic variables are greatly increased for promoters or conditions where one or more of the steps involved in forming the transcribing complex are reversible. Evidence exists for reversibility of all three classes of intermediate complexes in mechanism (*A*), at least at some promoters under some conditions. How does reversibility affect K_m and V_{max}? If the initial binding step is reversible, then the overall K_m increases if the rate constant of any reverse step increases. The magnitude of V_{max} does not depend on the initial binding step, but is reduced by an increase in the rate constant of any reverse isomerization step, as quantified in a section below.

(v) In general, how do changes in promoter sequence affect *V*, V_{max}, and K_m? Conceptually useful but mathematically approximate algebraic expressions for *V*, V_{max}, and K_m, obtained from application of the steady-state approximation to each intermediate in detailed initiation mechanisms derived from mechanism (*A*), are presented in this article. These demonstrate how the experimentally observable quantities *V*, V_{max}, and K_m may depend on the rate constants and equilibrium constants (ratios of forward and reverse rate constants) for

individual mechanistic steps. These rate and equilibrium constants in turn depend on promoter sequence and experimental conditions.

Equilibrium constants are related to free energy differences (ΔG^0) between the corresponding final and initial states; rate constants are interpreted in terms of quasithermodynamic activation free energy differences ($\Delta G^{0\ddagger}$) between the relevant "transition" state and the initial ("ground") state. Changes in the sequence of a promoter affect the free energies of various transition states and/or initial states and thereby affect the corresponding rate constants and equilibrium constants. The information now available regarding effects of sequence on free energies of initial states and transition states for related systems in vitro (e.g., references 129 and 154) indicates that (i) single base pair substitutions can have very large effects on these free energy differences, (ii) the first substitution in a site or subsite may have a much larger effect than additional substitutions, and (iii) the transition from specific to nonspecific binding can be quite abrupt. This information is presently being sought for RNA polymerase-promoter variants.

Organization of this Chapter

This chapter covers structural, thermodynamic, and kinetic studies focused on understanding the mechanism of transcription initiation. Our coverage is of necessity selective, not comprehensive. The article is intended to be readable by the nonspecialist, while at the same time discussing examples of current (quantitative) work in this exciting research area. The article is organized to introduce first the structural features of promoters and of RNA polymerase, followed by a review of selected structural and thermodynamic information regarding polymerase-promoter complexes on, or relevant to, the pathway of productive initiation. The kinetics of the steps in the overall process of initiation are then examined, with the goals of (i) relating studies of the kinetics of binding and isomerization to the overall kinetics of initiation, and (ii) relating kinetic parameters obtained from the original irreversible two-step mechanism used to interpret binding kinetic data to more realistic multistep mechanisms and to reversible binding conditions. In this regard, the kinetic and structural studies used to develop the present four-step mechanism of open complex formation at the λP_R promoter are reviewed. We then discuss the role of promoter sequence (an intrinsic variable) in determining the rate of transcription initiation and conclude with a brief summary of selected systems in which quantitative kinetic studies of the effects of supercoiling, repressors, and activators, among other extrinsic variables, have been performed. Other perspectives and discussions of these topics are provided by the classic reference books on *E. coli* RNA polymerase (133, 177, 179), comprehensive reviews (23, 25, 34, 35, 46, 75, 77, 86, 93, 108, 127, 143, 170, 178), and references therein.

THE PLAYERS

Promoter DNA Sequences Recognized by $E\sigma^{70}$

A promoter is the sequence of DNA from which RNA polymerase initiates transcription. Important elements of the promoter sequence are the positions where a change in sequence affects the rate of initiation of transcription from the specified start site. Sequences of ~300 naturally occurring *E. coli* and coliphage promoters (i.e., 15 to 30% of the total) recognized by $E\sigma^{70}$, the most abundant form of RNA polymerase holoenzyme, have been statistically analyzed (81, 131). This analysis demonstrates the key importance of three features within the "core" region of the promoter DNA: two conserved 6-bp DNA sequences centered approximately 10 and 33 bp upstream from the start site of transcription (+1), called the "–10" and "–35" hexamers; and the length of DNA that separates them, called the "spacer" region, which is most commonly ~17 bp in length but varies from 15 to 21 bp. Figure 2 summarizes the primary structure of this statistically defined (cloned and investigated in vivo and in vitro, but not yet naturally observed) $E\sigma^{70}$ consensus core promoter, specified by convention in the 5′ to 3′ direction as the sequence of the nontranscribed strand, which corresponds to the transcript sequence. When optimally aligned, the sequences of most promoters match at least 7 of the 12 bp in the consensus –35 (TTGACA) and –10 (TATAAT) hexamers (81, 131). At many promoters, the primary start site (+1; specified by the same convention) is an A and the bases flanking it are C(–1) and T(+2) (83).

The sequence conservation observed in the core promoter indicates functional importance. Indeed, *E. coli* promoters at which $E\sigma^{70}$ initiates at high rates typically exhibit no more than three deviations (in total) from consensus in the –35 and –10 regions of the promoter and from the consensus (17 bp) spacer length (157). Base substitutions that reduce a promoter's homology to consensus generally reduce the rate of transcription initiation from the promoter. In addition, other regions upstream (183) and downstream (22, 27) of the statistically defined core promoter are observed to be functionally important in at least some promoters. These regions, designated as the UP element and the downstream region (DSR), respectively, are shown in Fig. 1. More quantitative correlations between promoter structural features and the rate of transcription initiation are discussed in subsequent sections.

E. coli RNA Polymerase Holoenzyme

Subunits; Assembly. Promoter-specific initiation of transcription requires both core RNA polymerase (E) and a specificity subunit (σ) (28, 29, 91). Core polymerase is a stable noncovalent assembly of four polypeptide chains: two α (each 36,511 g mol^{-1}), one β (150,615 g mol^{-1}), and one β' (155,159 g mol^{-1}), which are the products of the *rpoA*, *rpoB*, and *rpoC* genes, respectively (see reference 28). Under normal growth conditions, transcription of most *E. coli* genes is initiated by RNA polymerase holoenzyme $E\sigma^{70}$ (459,163 g mol^{-1}), in which the specificity subunit σ^{70} (70,262 g mol^{-1}) is the product of the *rpoD* gene.

The pathway of assembly of core polymerase in vitro begins with dimerization of α to form α_2, followed by addition of β and subsequently of β' (97, 238). The core enzyme is thought to be thermodynamically stable to dissociation both in vivo and at the concentrations typically investigated ($\gtrsim$ 1 nM) in vitro. The interaction of σ^{70} with core involves region 2 of σ^{70} (128; see Fig. 4 below). The sites on the subunits of core contacted by σ remain to be determined. At one set of solution conditions, the equilibrium binding constant for association of σ^{70} with core RNA polymerase is ~2×10^9 M^{-1} (70). Under these conditions, a solution which is nominally 1 nM in holoenzyme (i.e., a 1:1 ratio of active core polymerase and σ^{70}) is actually an equilibrium mixture of holoenzyme (~0.5 nM), core polymerase (~0.5 nM), and unbound σ^{70} (~0.5 nM). Since the fractional extent of dissociation of holoenzyme of course increases with increasing dilu-

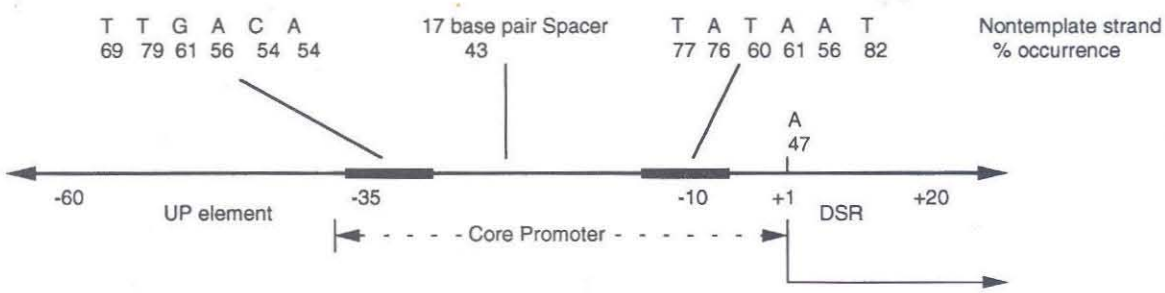

FIGURE 2 Structure of the "consensus" $E\sigma^{70}$ promoter showing functionally important regions. Consensus sequences of the –35 and –10 hexamers and the consensus spacer length shown are illustrated, along with their percentage of occurrence in the database of ~300 promoter sequences (81, 131). The upstream (UP) element (183) and downstream (DSR) region (22), which are of functional importance for some promoters (see text), are also indicated.

tion, dissociation must be of major significance in experiments performed at subnanomolar concentrations of polymerase.

Possible Structural Features of Holoenzyme Relevant for DNA Binding: Channel, Grooves, and Surrounding Flexible Regions. Neither $E\sigma^{70}$ RNA polymerase nor its subunits have been crystallized, and the holoenzyme is far too large to obtain a high-resolution solution structure by nuclear magnetic resonance. Low-resolution structural information was obtained originally by small-angle X-ray scattering (147) and more recently by electron microscopy (EM) image analyses of negatively stained two-dimensional crystals of $E\sigma^{70}$ (44). In the latter structure, with a resolution of ~27Å, the holoenzyme appears as an irregularly shaped object with dimensions of ~90 by 95 by 160Å (1 Å = 0.1 nm). The key structural feature observable at this resolution is a cylindrical channel ~25 Å in diameter and ~55 Å in length, which is surrounded by a "thumblike" projection (see Fig. 3A). Approximately 16 bp of double-helical DNA would fill this channel; this DNA would be removed from water if the surrounding thumb folded over it like the lid of a box. However, no such structural data on a $E\sigma^{70}$-promoter DNA complex are available.

Insight into other structural features of $E\sigma^{70}$ related to DNA binding comes primarily from sequence comparisons and comparisons of the very-low-resolution EM data on $E\sigma^{70}$ with high-resolution (~3 Å) X-ray crystallographic data for single-subunit polymerases (see Fig. 3A and below) and with the low-resolution (~16 Å) EM two-dimensional crystal structure of a nine-subunit variant of yeast RNA polymerase II (pol II) (43; see Fig. 3B), all

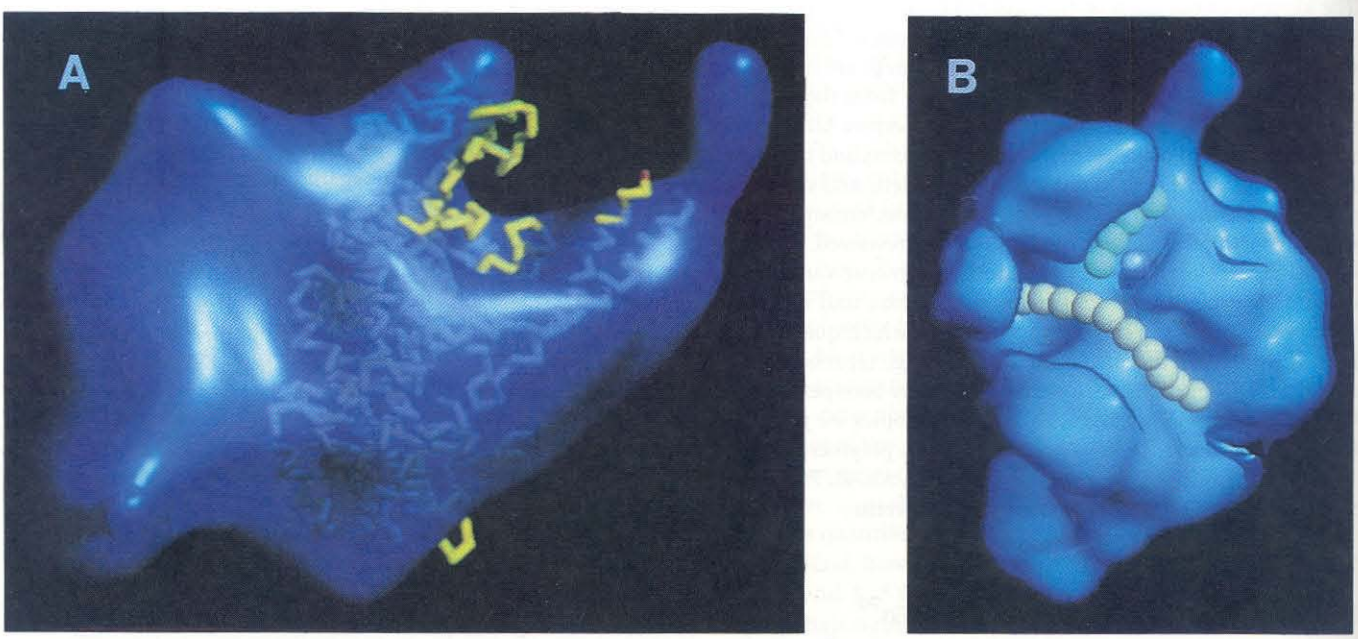

FIGURE 3 (A) Very-low-resolution (27 Å) EM crystal structure of *E. coli* $E\sigma^{70}$ RNA polymerase, with the high-resolution crystal structure of the Klenow fragment of *E. coli* DNA polymerase I superimposed. From reference 44 with permission. (B) Low-resolution (16 Å) EM crystal structure of yeast RNA polymerase II. Beads placed on the structure have a scaled diameter of 6.8 Å, corresponding to the length of 2 bp of double-helical DNA or of 2 bases of single-stranded DNA (see text). From reference 43 with permission.

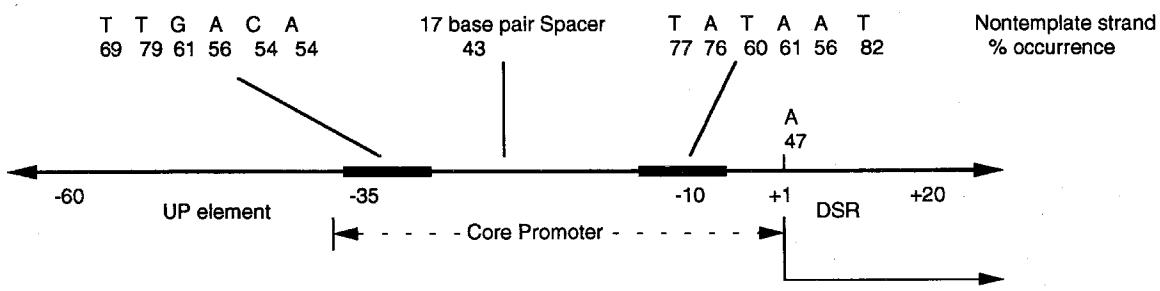

FIGURE 2 Structure of the "consensus" $E\sigma^{70}$ promoter showing functionally important regions. Consensus sequences of the –35 and –10 hexamers and the consensus spacer length shown are illustrated, along with their percentage of occurrence in the database of ~300 promoter sequences (81, 131). The upstream (UP) element (183) and downstream (DSR) region (22), which are of functional importance for some promoters (see text), are also indicated.

tion, dissociation must be of major significance in experiments performed at subnanomolar concentrations of polymerase.

Possible Structural Features of Holoenzyme Relevant for DNA Binding: Channel, Grooves, and Surrounding Flexible Regions. Neither $E\sigma^{70}$ RNA polymerase nor its subunits have been crystallized, and the holoenzyme is far too large to obtain a high-resolution solution structure by nuclear magnetic resonance. Low-resolution structural information was obtained originally by small-angle X-ray scattering (147) and more recently by electron microscopy (EM) image analyses of negatively stained two-dimensional crystals of $E\sigma^{70}$ (44). In the latter structure, with a resolution of ~27Å, the holoenzyme appears as an irregularly shaped object with dimensions of ~90 by 95 by 160Å (1 Å = 0.1 nm). The key structural feature observable at this resolution is a cylindrical channel ~25 Å in diameter and ~55 Å in length, which is surrounded by a "thumblike" projection (see Fig. 3A). Approximately 16 bp of double-helical DNA would fill this channel; this DNA would be removed from water if the surrounding thumb folded over it like the lid of a box. However, no such structural data on a $E\sigma^{70}$-promoter DNA complex are available.

Insight into other structural features of $E\sigma^{70}$ related to DNA binding comes primarily from sequence comparisons and comparisons of the very-low-resolution EM data on $E\sigma^{70}$ with high-resolution (~3 Å) X-ray crystallographic data for single-subunit polymerases (see Fig. 3A and below) and with the low-resolution (~16 Å) EM two-dimensional crystal structure of a nine-subunit variant of yeast RNA polymerase II (pol II) (43; see Fig. 3B), all

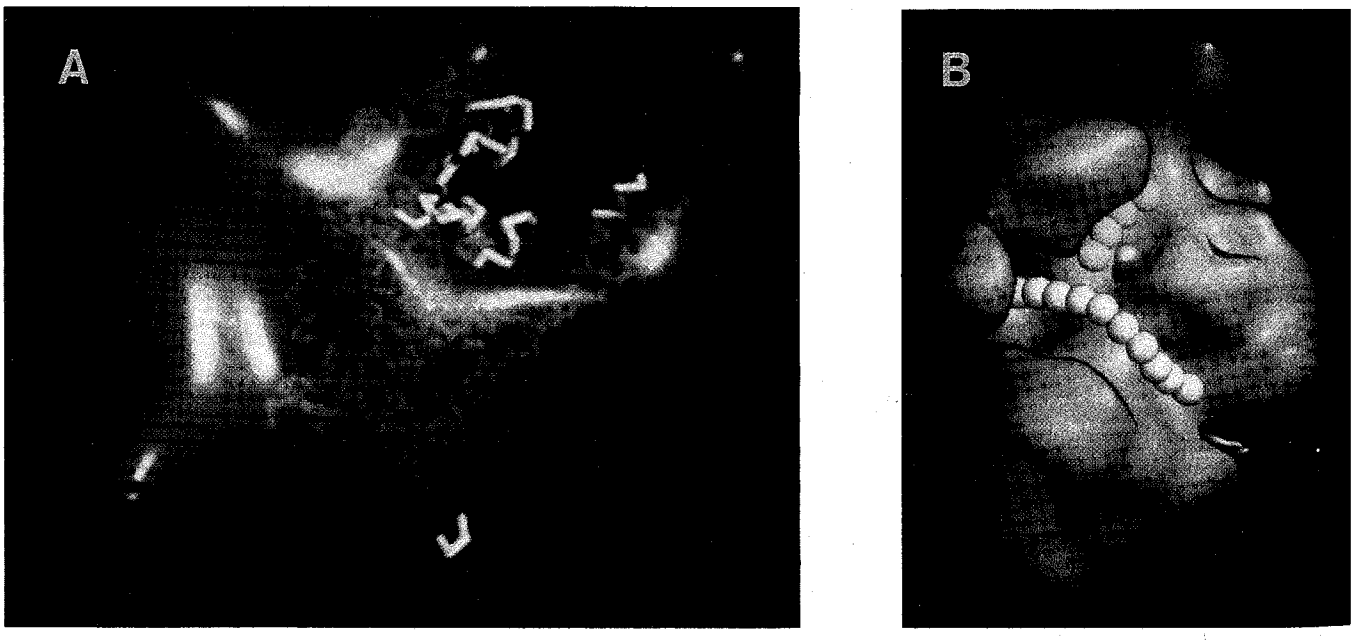

FIGURE 3 (A) Very-low-resolution (27 Å) EM crystal structure of *E. coli* $E\sigma^{70}$ RNA polymerase, with the high-resolution crystal structure of the Klenow fragment of *E. coli* DNA polymerase I superimposed. From reference 44 with permission. (B) Low-resolution (16 Å) EM crystal structure of yeast RNA polymerase II. Beads placed on the structure have a scaled diameter of 6.8 Å, corresponding to the length of 2 bp of double-helical DNA or of 2 bases of single-stranded DNA (see text). From reference 43 with permission.

obtained in the absence of DNA. Although this yeast pol II variant lacks two subunits important for specific binding to promoters, it contains the two largest subunits, which exhibit extensive sequence similarity to the large (β, β′) subunits of *E. coli* RNA polymerase (4, 219).

Key structural features of the yeast polymerase (see Fig. 3B) include a cylindrical channel ~25 Å in diameter and ~30 to 35 Å in length, which is surrounded by an "armlike" projection analogous to the thumblike projection in $E\sigma^{70}$. It is a reasonable working hypothesis to assume that these structural features are formed by conserved regions of these large subunits. Leading into the channel is a surface groove, ~25 Å in diameter and ~5 to 10 Å deep, extending ~40 Å. The groove and channel together span ~70 to 80 Å and could accommodate 20 to 25 bp of double-stranded DNA, of which ~10 bp would be in the channel, buried beneath the arm. If so bound, the DNA would be bent near the intersection of the groove and the channel. In this same region, a narrower, deep surface groove (~12 to 15 Å wide, ~20 to 25 Å deep, and ~30 Å long) is envisioned to branch from the channel. Near the branch point, a tunnel ~8 Å in diameter is proposed to extend for ~35 Å from the floor of the narrow groove through to the other side of the protein (see Fig. 3B). Given a monomer repeat of about 3.5Å (18), the narrower groove could accommodate up to ~9 bases of the nascent RNA chain (or possibly the nontemplate DNA strand); the tunnel could be an access route for NTP (or possibly could accommodate up to ~10 bases of the RNA chain) (43). Structural (footprinting) and binding (thermodynamic) data for the *E. coli* enzyme, reviewed in sections below, support the existence of a channel closed by an arm or thumb and of extensive surface binding grooves on $E\sigma^{70}$, some of which differ in their accessibility to promoter versus nonpromoter and single-stranded versus double-stranded DNA. Possible relationships between low-resolution structural features of yeast pol II and nucleic acid-binding properties of the $E\sigma^{70}$ polymerase are discussed in subsequent sections.

Structural similarities were also noted between the channel and projecting thumb of $E\sigma^{70}$ and the corresponding features of the high-resolution X-ray crystal structure of the large ("Klenow") fragment of *E. coli* DNA polymerase I, superimposed on the $E\sigma^{70}$ structure in Fig. 3A. Indeed, several single-subunit DNA and RNA polymerases, including the Klenow fragment (9, 163), human immunodeficiency virus reverse transcriptase (110), T7 and T7/T3 hybrid coliphage RNA polymerases (201, 202), and rat DNA polymerase β (166), appear to share a set of important structural features (208), including a "handlike" structure that constitutes the nucleic acid binding cleft. The "palms" (i.e., the base of the clefts) of these structures contain the key carboxyl residues which bind the Mg^{2+} ions required for catalysis of the polymerization reaction. The thumbs are flexible or disordered structures in the absence of DNA, which are thought to undergo conformational changes upon DNA binding, perhaps functioning to clamp the DNA in place. Possible relationships between these structural features and structural, thermodynamic, and kinetic-mechanistic properties of $E\sigma^{70}$-promoter complexes are discussed in subsequent sections.

Roles of Subunits in the Steps of Initiation.

The "specificity subunit" σ. While core polymerase can initiate transcription from ends, nicks, and open regions of DNA, a bound σ subunit is required for promoter-specific initiation (29). The *E. coli* genome encodes at least six sigma factors, each of which directs RNA polymerase to a different set of promoter sequences (77). The best-studied sigma factor, σ^{70}, is required for transcription of genes involved in most fundamental cell functions (metabolism, biosynthesis, etc.) during exponential growth. Other sigma factors coordinate transcription of functionally related sets of coregulated genes, including those involved in heat shock response (σ^{32} and σ^{24} [σ^{E}]), nitrogen assimilation (σ^{54}), flagellum gene expression (σ^{28}), and stationary-phase expression (σ^{38}) (see reference 77 and the Appendix to this chapter).

With the exception of σ^{54}, the above-mentioned sigma subunits are homologous (see references 77 and 86). Sequence comparisons that include sigma factors from other prokaryotes indicate that all known sigma factors are either σ^{70} type or σ^{54} type. σ^{70}-type sigma factors can be further classified into primary (e.g., σ^{70}) and alternative (σ^{32}, σ^{28}, etc.) sigma factors. σ^{70} exhibits greater similarity to primary σ factors from other organisms than to alternative (cf. σ^{32}) sigma factors in *E. coli* (see Fig. 4). We have focused this review on transcription initiation by $E\sigma^{70}$ holoenzyme.

Important contacts between the σ^{70} subunit and promoter DNA were identified by isolating mutations in the sigma-encoding *rpoD* gene that suppressed promoter mutations in the –10 and –35 sequences. Conserved region 2.4 of σ^{70} was found to be involved in recognition of the –10 sequence (198, 231). (An analogous conclusion was obtained for a minor σ factor in *Bacillus subtilis* [42]). Conserved region 4 is involved in recognition of the –35 sequence (65, 198). Key residues in conserved region 2.4 may be part of an α helix; key residues in conserved region 4 may be part of a helix-turn-helix motif, which is characteristic of many proteins that bind with specificity to sites on double-stranded DNA (see reference 77).

The σ subunit (specifically conserved region 2.3 of σ^{70}) has also been postulated to play a role in the local DNA opening that occurs during open complex formation (86). This model is consistent with identification of mutants in the analogous sigma subunit of *B. subtilis* (σ^{A}) that are defective in DNA opening (3, 103). However,

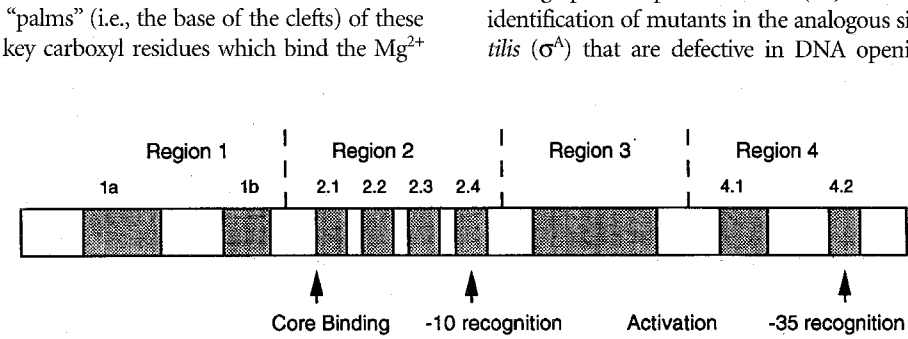

FIGURE 4 Regions of sequence conservation in primary (σ^{70}-type) prokaryotic sigma factors (adapted from reference 77). Shaded regions are conserved in primary (σ^{70}-type) prokaryotic sigma factors; functions of these regions are specified where known. Regions 2, 3, and 4 are also conserved in alternative (cf. σ^{32}) sigma factors.

mutants in comparable positions in *E. coli* σ^{70} retain transcriptional activity on a promoter in vivo (232).

Although intact σ^{70} does not bind detectably to DNA in the absence of core RNA polymerase, peptides containing regions 2 and 4 of σ^{70} bind with modest specificity to promoter DNA (51). In free σ^{70}, conserved region 1 apparently inhibits formation of specific σ^{70}-promoter complexes (50). Truncated peptides not containing region 1 bind to promoter DNA, but with much lower stability and specificity than that characteristic of $E\sigma^{70}$ holoenzyme. Core enzyme subunits must contribute directly to stability and indirectly to specificity of promoter binding.

Conserved region 3 of σ^{70} appears to be positioned in the vicinity of the binding site for the initiating nucleotide, and thus near the +1 site on the template. An initiating nucleotide analog was found to cross-link to this region of σ^{70} (as well as to the β subunit) in an open complex at the T7 A1 promoter (194).

In addition to its direct role in promoter recognition, σ^{70} has recently been implicated in some forms of positive gene regulation. In general, it appears that positive activators that bind to sites centered at or near –40 make contacts with σ^{70}. Examples include λcI activation of λP_{RM} (117, 130), PhoB activation of P*pstS* (118), and CAP activation of *gal* P1 (118). Mutations in sigma region 4 suppress cI mutants defective in activation of λP_{RM} in an allele-specific fashion, but do not suppress cI mutants defective in DNA binding. (The activation defective mutants are called positive control [pc] σ mutants [130].) Likewise, deletion analysis of sigma indicated that activation of *pstSp* by PhoB and of *gal* P1 by CAP requires region 4 and part of region 3 (118).

Although this chapter is dedicated to the analysis of transcription initiation involving the $E\sigma^{70}$ holoenzyme, it is important to point out the potential ambiguities resulting from the overlapping specificities of $E\sigma^{38}$ (also designated $E\sigma^{S}$) and $E\sigma^{70}$. $E\sigma^{38}$ (which is present during stationary phase) programs transcripts from some $E\sigma^{70}$ promoters because of the similar –10 region recognition pattern (222). This observation may complicate some in vivo studies of promoter specificity and some in vitro studies with uncertain σ^{38} contamination.

Role of α in interactions with upstream DNA regions and bound factors. The α subunit plays an important and direct role in promoter recognition at some promoters by specifically interacting with sequences upstream of the –35 region. $E\sigma^{70}$ variants in which the carboxy-terminal third of α is deleted, or changed by a point mutation, fail to utilize promoter upstream elements (UP elements; see below) such as that found in the *rrnB* P1 promoter (183). The direct role of the C-terminal domain (CTD) of α in recognition of the UP element was demonstrated by DNase I footprinting; both purified α and α-CTD specifically protect UP element DNA (16, 183).

The α subunit is also indirectly involved in the activation of several promoters, as discussed below. In these cases, the carboxy-terminal sequence appears to provide the contact point for the activating protein. No activation occurs in the absence of the carboxy-terminal sequence or for some carboxy-terminal point mutants (95, 98, 223). It is of interest that the α-UP element contact domain and the α-positive activator contact domain are in the same region of α. It is not yet clear whether these two functions are mechanistically related.

The α subunit is not believed to play a direct role in transcription initiation at many other promoters (presumably lacking UP elements). For example, the holoenzyme assembled from α peptides missing the carboxy-terminal third of the primary sequence initiated transcription in vitro with the expected specificity and efficiency at a number of promoters (95). The amino-terminal two-thirds of α (from residue 8 to 241) appears to be sufficient for holoenzyme assembly and, in these cases, for function.

Roles of β and β′ subunits in initiation. Since the large β and β′ subunits of $E\sigma^{70}$ play important roles in binding DNA and NTP and in catalysis of RNA synthesis, it is unfortunate that they are not better characterized at a structural level. Extensive sequence homology exists between these large subunits of $E\sigma^{70}$ and the largest subunits of other prokaryotic and eukaryotic DNA-dependent RNA polymerases (4, 219). The β and β′ subunits probably define the major structural features of $E\sigma^{70}$ (e.g., the channel, grooves, and thumb [or arm]) responsible for stable DNA binding in at least the processive (elongation) phase of transcription (44; see Fig. 3A) and probably in specific promoter-bound open complexes as well.

In *E. coli* polymerase, the β subunit contains the site for binding the initial nucleotide substrate and is involved in the initial polymerization steps which lead to promoter clearance. Cross-linking studies using initiating nucleotide analogs indicate that residues 1065 and 1237 of the β subunit are in the immediate vicinity of the initiating nucleotide (71, 72, 159). A number of mutants mapping in the β subunit exhibit altered substrate binding properties and are defective in promoter clearance (101). Stepwise substitution of alanine for the amino acids in the region surrounding residue 1065, thought to be involved in binding the initiating nucleotide, alters the abortive initiation and promoter clearance properties of the enzyme (106, 185). The β subunit is probably involved in DNA binding, since β can be cross-linked to DNA in a holoenzyme-DNA complex (38). The roles of the β′ subunit in the initiation process also remain largely undefined. β′ binds nonspecifically to DNA (239) and interacts with at least the nontemplate strand of the promoter DNA (38, 162).

RNA POLYMERASE-PROMOTER COMPLEXES: STRUCTURAL AND THERMODYNAMIC PROPERTIES

In the development of this field, kinetic and/or thermodynamic evidence for the intermediate complexes on the pathway to initiation in general preceded the structural characterization obtained by trapping these intermediates. We invert this order for ease of presentation of this material. Complexes believed to be mechanistically significant on the pathway of transcription initiation are described below.

Structural Characterization of Binary Complexes

The structures of polymerase-promoter complexes have been studied by probing the availability of DNA functional groups to cleavage by specific chemical or enzymatic probes. Sites available for cleavage in free DNA but less accessible in an RNA polymerase-promoter complex are considered sites of protection, while sites cleaved to a greater extent in the RNA polymerase-promoter complex than in free DNA are termed enhancements. This type of chemical and enzymatic probing, known as footprinting (see references 175 and 226 for reviews), provides specific information regarding accessibility to the probe of (i) the DNA phosphodiester backbone (hydroxyl radical [HO·] or DNase I cleav-

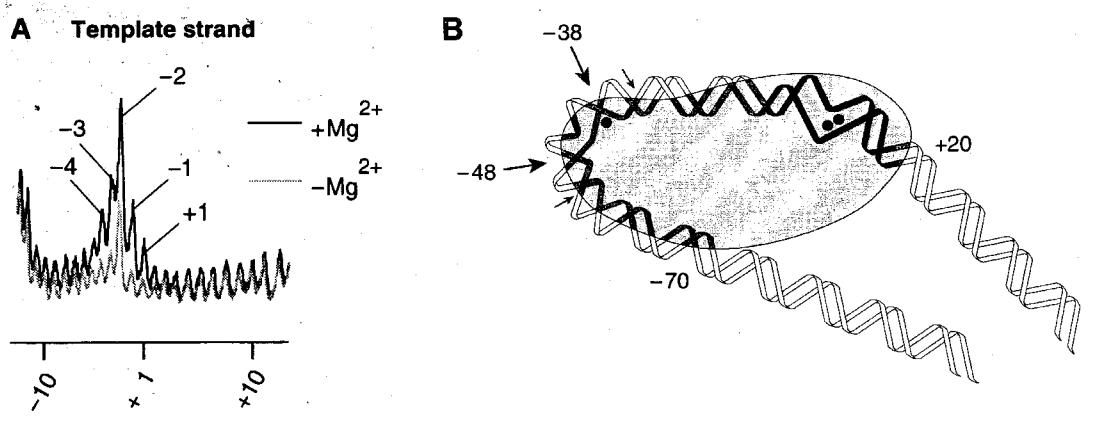

FIGURE 5 (A) Increased accessibility of the λP_R transcription start site to HO· in RP_{o2} versus RP_{o1}. Phosphorimager comparison of template strand footprints obtained in the presence and absence of Mg^{2+} (from reference 40a with permission). (B) Two-dimensional model for the λP_R-Eσ^{70} open complex (RP_{o2}). Sketch of λP_R promoter DNA on RNA polymerase based on the approximate dimensions of the protein (160 by 100 Å) determined by Darst et al. (45). In the upstream region, four DNA bends result in the wrapping of DNA around the protein. The small arrows show intrinsic DNA bends at A_n-T_n tracts; the large arrows show protein-induced bends. Spheres (not drawn to scale) indicate possible binding sites for the three Mg^{2+} ions (221) taken up in the conversion from RP_{o1} to RP_{o2} (from reference 40a with permission).

age), (ii) the major groove (methylation of guanines by dimethylsulfate [DMS]), (iii) the minor groove (methylation of primarily adenines by DMS), and (iv) normally buried surfaces of the bases, especially T (by $KMnO_4$) and C (by DMS), which become accessible due to local opening or severe bending. Interference experiments, the reverse of protection experiments, detect sites where chemical modification interferes with formation of a complex. Our review focuses on general characteristics of polymerase-promoter interactions deduced from these studies.

Closed Complexes (RP_c). The process of specific recognition of the promoter DNA sequence begins in the initial closed complex (RP_{c1}), which involves only regions of the promoter upstream of the transcription start site. Subsequent conformational changes in polymerase and DNA extend the contacted region through the transcription start site and into the downstream region, to form the closed complex designated RP_{c2}.

Initial closed complex: RP_{c1}. Initial closed complexes are characterized by trapping them at low temperatures (typically ~0 to 10°C) where the subsequent isomerization steps are thermodynamically unfavorable and/or kinetically blocked. DNase I and HO· have been used as probes of the accessibility of the DNA phosphodiester backbone in RP_{c1} to large (DNase I) and small (HO·) reagents. Studies of initial closed complexes at a variety of promoters, including those recognized by both Eσ^{70} and Eσ^{32}, indicate that RNA polymerase contacts one face of the promoter DNA from –55 to –5 in this complex (17, 40, 92, 113, 146, 188). The protected region does not include the start site (+1) or the initial transcribed region, indicating that neither Eσ^{70} nor Eσ^{32} interacts with this key region of the promoter in RP_{c1}. Within the –55 to –5 region, an alternating pattern of protection and deprotection is apparent from the HO· footprints (146, 188). This pattern exhibits a 10- to 13-bp periodicity which is approximately that of the DNA helix, indicating that binding occurs to one face (side) of the double helix (227). On the basis of the EM crystallographic data for yeast pol II reviewed above, it is reasonable to propose that RP_{c1} involves binding the promoter DNA in the wider surface groove (but not a closed channel) of the Eσ^{70} polymerase. However, the 50-bp region of DNA spanned in RP_{c1} is far longer than the length of the wide surface groove identified in yeast pol II.

Isomerized closed complex(es): RP_{c2}. Formation of the isomerized closed complex (RP_{c2}) involves major conformational changes in the polymerase and/or the DNA. RP_{c2} is trapped for study at a promoter-specific range of temperatures (typically ~10 to 15°C), above the temperature where only RP_{c1} is observed but below that where significant opening of the DNA occurs. Probing of RP_{c2} complexes at the *lac*UV5 (204), *tetR* (54), GroE (40, 146), and T7 A1 (188) promoters with DNase I and HO· yielded "footprints" spanning the range from position –55 to approximately +20 on the promoter DNA. Clearly, more of the promoter DNA interacts with RNA polymerase in RP_{c2} than in RP_{c1}. Almost all of the additional protection occurs downstream of the start site; the upstream boundary of the protected region is unchanged. The contacts between polymerase and the downstream region which form in the transition from RP_{c1} to RP_{c2} presumably play a role in the initiation (i.e., nucleation) of DNA opening.

The periodicity of protection in the upstream region (–10 to –55) is virtually identical to that observed by HO· footprinting for RP_{c1} at the *groE* (146) and T7 A1 (188) promoters. Both Eσ^{32} and Eσ^{70} polymerases therefore interact with the upstream region of the promoter DNA in RP_{c2} in the same manner as in RP_{c1}, indicating that this DNA must remain in a surface groove on the polymerase. Downstream, however, protection from HO· is complete from –10 to +20 on both template and nontemplate strands, indicating that RNA polymerase contacts the backbone of both strands of the DNA helix in this region of the promoter. This result suggests that the region from –10 to +20 in RP_{c2} may be bound in the channel, with the backbones of both strands

buried by the arm, thumb, or other flexible features of this structure. If this model is correct, then both conformational changes in the arm of the polymerase and bending of the DNA may be involved in the conversion of RP_{c1} to RP_{c2}, as judged from the features of the yeast pol II groove-channel-arm structure.

RP_{c2} is designated a closed complex on the basis of the lack of reactivity of cytosines of the promoter DNA in this complex with DMS (40, 55). (Recall that DMS reacts with accessible cytosines in open or distorted regions of DNA.) The lack of DMS reactivity of cytosine does not by itself prove that RP_{c2} is an entirely closed complex, because the opening might be confined to a few A-T base pairs whose adenines are unreactive in free DNA, and/or because any open bases might be protected by the polymerase. The possibility of protection by polymerase on both strands can be eliminated because the complementary guanines at the positions of the unreactive cytosines are reactive to (and hence accessible to) DMS. A scenario where only the nontemplate strand (which contains the cytosines) is protected cannot be ruled out. If opening were present but confined to a few A-T base pairs, these probably would be in the region –9 to –12, as judged by the data reviewed above on RP_{c2} complexes at the *groE* ($E\sigma^{32}$) and *tetR* ($E\sigma^{70}$) promoters. Evidence for some untwisting (~2 to 4 bp) of promoter DNA early in the process of open complex formation is provided by kinetic studies with supercoiled templates (216), as reviewed below.

Binary Open Complexes (RP_o). The process of DNA strand opening appears to begin upstream of the transcription start site and to extend subsequently to the start site in the "final" binary open complex (RP_{o2}), in which the template strand is accessible to initiating NTP. We begin by reviewing the available structural information for RP_{o2} and then discuss the more limited structural information regarding intermediate binary open complexes.

Final binary open complex: RP_{o2}. Chemical probing of the single-stranded regions of the final binary open complex (RP_{o2}, formed in the presence of Mg^{2+}) with DMS and $KMnO_4$ indicates that DNA strand opening extends from the –10 region (as far upstream as position –12) to the region of the start site (as far downstream as +2) (55, 107, 186, 196, 218). Most pyrimidines on both strands in this region are reactive (though the extent of reactivity varies with position). Hence the bases are accessible to small solutes and cannot be completely buried in the binding channel. On the other hand, HO· footprinting indicates that the backbone of the nontemplate strand is completely protected in the region –10 to +20 by $E\sigma^{70}$ in RP_{o2} and that only a small part of the backbone of the template strand (from –4 to +1 at λP_R [40a]) is accessible to HO· in this key region (Fig. 5A). (Recall that the backbones of both strands are protected from HO· cleavage in the region –10 to +20 in RP_{c2}.) Subsequent binding of the initiating nucleotide to form the first RP_{init} ternary complex appears to occur without a major conformational change (203).

In RP_{o2}, the promoter DNA is protected from DNase I digestion from approximately –55 to +20 (40, 40a, 54, 92, 113, 197, 204). At all promoters investigated (including *lac*UV5 [204], *tetR* [54]), fdVII P2 [92], and GroE [40]), the patterns of methylation and DNase I cleavage protection in RP_{o2} remain essentially unchanged from those in RP_{c2}, indicating that the opening of the DNA in the vicinity of the transcription start site does not require significant shifts in other areas of RNA polymerase-DNA contacts. In particular, the positions of protection and deprotection in the upstream region as well as the lack of periodicity in the protection downstream of –10 are preserved in RP_{o2} (146). Interpreted in terms of the yeast pol II structure, this result indicates that the upstream region of the promoter DNA (–55 to –10 in RP_{o2}) remains bound in a wide surface groove, as in RP_{c1} and RP_{c2} (139).

Protection of DNA from DMS methylation in open complexes at six promoters occurs primarily in the region from –32 to +8, with the most complete protection (in both the major and minor grooves) occurring in the –10 region (–8 → –13) (54). In the region from –14 to –20 (i.e., the downstream part of the spacer region), both the major and minor grooves are protected from DMS methylation. In particular, the major groove at –14 and –15 is always protected, even though the sequence at these positions is not highly conserved (exhibiting only a slight bias toward guanine [81]). Comparison of methylation interference data for seven promoters indicated that methylation of residues between positions –20 and –24 of the spacer region never interfered with open complex formation (54), suggesting that these residues have no specific role in open complex formation.

The DNA in the open complex is thought to be bent (see, e.g., references 167 and 225), based in part on the results of gel mobility shift assays (116) and hyperreactivity of bases in the spacer region to DNase I digestion (Fig. 5B) (146). Even if the DNA-binding site on the polymerase were aligned with the longest dimension (~160 Å) of the enzyme, no more than ~50 bp of linear B-DNA could be protected. Since ~70 bp of DNA is protected in both RP_{c2} and the open complexes (40, 204), the promoter DNA must be bent or wrapped around the enzyme (44). Scanning atomic force microscopy of open complexes provides additional evidence for this proposal (173). However, analysis of neutron scattering studies of an open complex of $E\sigma^{70}$ at the T7 A1 promoter suggested that if bending occurs in this complex, the bend angle must be less than 45° (89).

Investigation of positions of cross-links between RNA polymerase and promoter DNA in open complexes indicates that β, β′, and σ all contact both the –10 and –35 regions of the core promoter (21, 26, 38, 39, 65, 82, 165, 198, 199, 231). Presumably, therefore, σ lies along the groove and channel created by the β and β′ subunits. Genetic and footprinting evidence suggests that at the *rrnB* P1 promoter, α contacts the region upstream of the –35 sequence (the UP element [183]).

Intermediate open complex(es). Potential intermediate open complexes have been observed at several promoters (188, 211, 217, 218; W. C. Suh, Ph.D. thesis, University of Wisconsin, Madison, 1993). The best characterized of these is the complex formed at the λP_R promoter in the absence of Mg^{2+} (217, 218; Suh, Ph.D. thesis). This putative intermediate complex, designated RP_{o1}, differs from RP_{o2} in having a smaller number of open bases and a smaller extent of $KMnO_4$ reactivity of bases which are reactive (i.e., open). In RP_{o1}, the open region is confined principally to positions –11 to –1. Notably, the start site (+1) is relatively inaccessible to $KMnO_4$. Positions –12, +1, and +2 exhibit the largest relative increases in $KMnO_4$ reactivity in the transition from RP_{o1} to RP_{o2}. Hence, at least the final stage of opening appears to propagate bidirectionally.

The conversion $RP_{o1} \rightarrow RP_{o2}$ makes the start site accessible to probes [HO· from $Fe(EDTA)^{2-}$, $KMnO_4$] which are of comparable size to the initiating nucleotide and also negatively charged (Fig. 5A). The transition from RP_{o1} to RP_{o2} requires Mg^{2+} at millimolar

concentration and cannot be efficiently induced by elevated temperature or negative supercoiling in the absence of Mg^{2+} (218; Suh, Ph.D. thesis). Since the function of RP_{o2} is to bind the initiating nucleotide, it is plausible to propose that the requirement for Mg^{2+} in the conversion of RP_{o1} to RP_{o2}, which greatly increases the accessibility of the start site on the template strand to both MnO_4^- and HO· [i.e., to $Fe(EDTA)^{2-}$], involves binding of Mg^{2+} ions in the vicinity of the start site (Fig. 5B). By analogy with structural data on single-subunit polymerases (208), these Mg^{2+} ions may bind to the triad of carboxylate side chains involved in catalysis of phosphodiester bond formation. The increase in reactivity to negatively charged reagents at this position may in part reflect local neutralization of negative charge by binding Mg^{2+}, in addition to (or in place of) conformational changes causing additional opening (40a).

Methylation of cytosines by DMS as a function of temperature in the presence of Mg^{2+} is uniformly temperature dependent at all positions (54, 107). Therefore, opening as a function of temperature appears to behave thermodynamically as an all-or-none cooperative process, with no evidence for stable partially opened complexes such as RP_{o1} in the presence of Mg^{2+}. Since RP_{o1} (at λP_R) is trapped as a function of Mg^{2+} concentration but not temperature, other intermediates may be discovered by judicious use of ion concentrations as well as temperature to probe the mechanism of open complex formation. The proposal that RP_{o1} (trapped in the absence of Mg^{2+} but not at intermediate temperatures) is an intermediate on the kinetic pathway of opening in the presence of Mg^{2+} is supported by the rationale that RP_{o1} rapidly converts to RP_{init} upon addition of Mg^{2+} and a subset of NTP (180, 181).

Complexes which may correspond to RP_{o1} at λP_R are also observed at other promoters. In particular, synthetic *lac* variant and consensus promoters probed without Mg^{2+} exhibit $KMnO_4$ reactivities characteristic of RP_{o1}. At the *lac*UV5 promoter, a presumably open complex (probed in gel slices) exhibited no reactivity to 1,10-phenanthroline-copper in the absence of Mg^{2+}, but in the presence of Mg^{2+} it was reactive to the probe in the region between positions –3 and –6 (120). Two open complexes at the *lac*UV5 promoter are observed by gel mobility shift assays at intermediate temperatures in the presence of Mg^{2+} (209, 211). Relative amounts of these two complexes are highly temperature dependent. It is not yet clear whether either of these open complexes corresponds to the complex (RP_{o1}) formed in the absence of Mg^{2+}.

Characteristics of Ternary Initiated Complexes (RP_{init}) and the Events of Promoter Clearance

The binary open complex RP_{o2} is poised to initiate transcription, awaiting the addition of the four NTP substrates. The next event in the transcription initiation process is the choice of the start site and the associated binding of the initiating NTP to RNA polymerase. In the presence of all four nucleotides (the in vivo situation), most promoters appear to have a cluster of start sites with one usually predominating both in vivo and in vitro (83). Three factors appear to dictate the choice of primary start site. First, the start site is typically restricted to positions 6 to 8 bp downstream of the –10 hexamer. Also, although any of the four nucleotides can serve as an initiating nucleotide, given a choice, RNA polymerase appears to prefer initiating with a purine. Finally, for promoters in which a C is the initiating nucleotide, an unusually high intracellular concentration of the nucleotide is probably necessary, presumably because RNA polymerase exhibits a higher steady-state enzymatic (Michaelis) constant (K_m^{NTP}) for dissociation of C as an initiating nucleotide than for A and G (200). Key elements of the core promoter sequence may also influence the decision regarding start site (32).

Once the start site is chosen (i.e., once the initiating nucleotide is bound), RNA polymerase begins the complex series of events that eventually lead to promoter clearance. Promoter clearance (which typically occurs after the synthesis of a 7- to 12-nucleotide-long message) involves release of contacts with the promoter, loss of the σ subunit, and synthesis of full-length transcripts (31, 80, 115, 210). We designate all of the complexes on the pathway to promoter clearance collectively as RP_{init}.

RP_{init} complexes have been characterized by examining the DNA structure associated with RNA polymerase as a function of the number of nucleotides that have been incorporated by RNA polymerase (e.g., the potential length of the RNA transcript in the ternary complex), using subsets of the four nucleotides. At the *lac*UV5 promoter (which initiates with a sequence AAUU), no difference was observed between the structure of the complex formed in the absence of nucleotides (as probed by DMS) and that formed in the presence of the dinucleotide ApA (203). However, upon addition of both U and ApA to RP_{o2}, the open region shifts downstream. Similar effects are seen at the λP_R promoter (221, 224; Suh, Ph.D. thesis). It is plausible (but not proven) that these trapped complexes are also significant intermediates in the pathway of initiation in the presence of all four nucleotides.

These structurally unique properties are consistent with the observation that RP_{init} complexes are more stable to dissociation than RP_{o2} complexes (34, 137, 180). At many promoters, RP_{init} is resistant to brief competition with high salt concentrations (>0.5 M) that would cause RP_{o2} to dissociate rapidly. In general, stabilization is by no means absolute, because RP_{init} complexes typically undergo a process of recycling (abortive initiation) in which short transcripts are continuously synthesized and released, reforming RP_{o2} (discussed below). Since the pathway of dissociation of polymerase from RP_{init} is probably via RP_{o2}, formation of RP_{init} reduces the rate of dissociation from that characteristic of RP_{o2}.

The catalytic behavior of RP_{init} in vitro is complex. Even in the presence of a sufficient excess of all four NTPs, aborted as well as full-length products are synthesized. At some promoters the molar yield of aborted products is significantly greater than that of full-length products. After each nucleotide condensation step, there is a finite probability of an abortive cycle in which polymerase releases the short transcript and reforms RP_{o2}.

Alternatively, the polymerase continues with the next incorporation step. The probability of cycling versus elongation at each step is unique to a particular type of promoter, as is the yield of abortive product versus full-length product. After synthesis of a 7- to 15-nucleotide-long transcript, RNA polymerase escapes abortive cycling and continues on to synthesize a full-length transcript (30, 73, 114, 158, 210; L. M. Hsu and M. J. Chamberlin, in preparation).

Another surprising property of the RP_{init} complex is that initial short transcripts can undergo a "slippage" event to yield transcripts with additional Us (or As) at the 5′ end of the transcripts when the start site is embedded within a run of T/A (or A/T) base pairs. Slippage appears to be involved in forming stable ternary (RNA polymerase, promoter, nucleotide) complexes at the *rrnB* P1 promoter (19). Slippage of RP_{init} is known

to occur both in vitro and in vivo (78, 236). It is not clear what effect, if any, oligo(A) or oligo(U) 5′ tails have on the transcripts.

In one system (*pyrBI*), transcriptional slippage is proposed to have an interesting role leading to pyrimidine (UTP)-mediated gene regulation. In this case, slippage appears to block productive messenger elongation (this is not the case for other promoters demonstrating slippage [e.g., see references 78 and 236]), and slippage (and hence no elongation) occurs only in the presence of high levels of UTP (132).

Are properties of RP_{init} relevant in vivo? Evidence exists that at some strong promoters, where V_{TC} (equation 1a) is relatively high, steps in class III of mechanism (*A*) (i.e., after open complex formation and before promoter clearance) are primary determinants of the rate of initiation (22, 57, 58, 158, 206). It is possible that initiated complexes at some promoters pause during recycling or slippage, delaying promoter clearance events. In vivo hyperreactivity to $KMnO_4$ was observed in the region between positions +6 and +12 in such a promoter (57). This $KMnO_4$ reactivity presumably results from the open region of a ternary complex paused at a step before promoter clearance, although it has yet to be demonstrated that this complex is undergoing nonproductive initiation. Additional evidence exists for a role of recycling of RP_{init} in determining promoter activity. The DSR is important in determining the strength of some promoters (27; see below). This observation has been largely attributed to an effect on the promoter-specific recycling properties of RP_{init} (Hsu and Chamberlin, in preparation). Finally, CAP (catabolite gene activator protein)-mediated positive regulation of *malT* occurs by facilitating promoter clearance from the RP_{init} complex (151).

RNA POLYMERASE-PROMOTER COMPLEXES: STABILITY AND EQUILIBRIUM SPECIFICITY

What noncovalent interactions are involved in forming a specific, stable complex between RNA polymerase and a promoter DNA site? To introduce this topic, we briefly summarize the quantitative definitions of stability and specificity and the general consequences for stability of bringing together oppositely charged regions of protein and DNA surface and of removing nonpolar regions of protein and DNA surface from exposure to water. Then we review the available information regarding the origins of stability and specificity of holoenzyme-promoter complexes by comparing them with other specific protein-DNA complexes and with other complexes of core polymerase and holoenzyme with nonpromoter DNA.

General Thermodynamic Properties of Protein-DNA Complexes

For a protein-DNA interaction of the form $A + B \leftrightarrow AB$, stability is characterized by the equilibrium concentration quotient K_{obs}, defined as $K_{obs} \equiv ([AB]/[A][B])_{eq}$. Equilibrium specificity is characterized by the ratio of values of K_{obs} for specific and nonspecific binding ($K_{obs}^{specific} / K_{obs}^{nonspecific}$), where each DNA nucleotide (phosphate) constitutes a potential nonspecific binding site.

In general, stabilities of both specific and nonspecific protein-DNA complexes increase profoundly with a reduction in salt concentration. In many cases, virtually any desired stability can be achieved by a suitable choice of the concentration and also the chemical identity of the salt (see references 170 and 172). (Of particular importance are the valence of the cation and the chemical identity of the anion.) The strong effects of salt concentration are interpreted in terms of local reductions in the density of negatively charged phosphate groups on DNA, either by interaction with positively charged functional groups or regions on the protein or by conformational changes (e.g., local denaturation). The polyelectrolyte character of sufficiently large DNA molecules (>100 bp) guarantees that both local denaturation and binding of any (locally) positively charged ligand will be highly salt-concentration dependent as a consequence of the release of salt cations from the vicinity of the DNA when its phosphate charge density is locally reduced (18, 164, 171).

Stabilities (K_{obs}) of specific protein-DNA complexes typically show an unusual but characteristic temperature dependence, in which K_{obs} exhibits a maximum at a temperature near the physiological range (~10 to 50°C). The existence of a temperature of maximum stability is interpreted in terms of the importance of the hydrophobic effect (the removal of nonpolar surface from exposure to water) as a driving force for specific binding (79, 205). At temperatures below that of maximum stability, binding is enthalpically unfavorable but entropically favorable; above the temperature of maximum stability, binding is favorable both enthalpically and entropically, though the latter effect dominates.

From the characteristic effects of salt concentration and temperature on the stability of $E\sigma^{70}$-λP_R promoter open complexes (RP_{o2}), we conclude that extensive reductions of DNA phosphate charge density, extensive burial of nonpolar surface, and extensive coupled folding local regions of polymerase all play important roles in this interaction, as summarized below.

RNA Polymerase-DNA Complexes

E. coli RNA polymerase exhibits multiple modes of specific and nonspecific binding. In all cases investigated, binding is driven by the favorable entropy change; enthalpic effects are either small (for nonspecific binding) or highly unfavorable (for open complex formation). How can a binding interaction be so favorable entropically? In the gas phase, it could not; in aqueous salt solutions, the redistribution of water and of salt ions that accompanies binding provides a large entropic driving force, especially for processes which reduce the amount of nonpolar surface exposed to water (the "hydrophobic effect") and/or which reduce the charge density of polyelectrolyte (e.g., DNA) surfaces (the "polyelectrolyte effect"). A maximum estimate of the primarily entropic contribution of cation release (the polyelectrolyte effect) to the binding free energy may be obtained at a specified salt concentration from knowledge of SK_{obs} (see next section) (171, 205).

In addition to its specificity for promoter sequences, $E\sigma^{70}$ holoenzyme also exhibits specificity of binding to ends of double-stranded DNA and to as-yet-uncharacterized "tight-binding" but noninitiating sites (148). Nonspecific (random) binding of both $E\sigma^{70}$ holoenzyme and core polymerase appears to occur on all regions of double-stranded and single-stranded DNA. Both core and holoenzyme bind much more strongly to single-stranded than to double-stranded DNA, and holoenzyme binds more strongly than does core to single-stranded DNA (47). Hence both core and holoenzyme are "melting proteins," which by their presence destabilize double-stranded DNA thermodynamically (99).

Early studies of binding of $E\sigma^{70}$ RNA polymerase to promoters, generally performed at low salt concentration, indicated that open complex formation was essentially irreversible on the time

scale of subsequent steps in transcription initiation. Hence the physiologically important determinant of specificity of promoter recognition was thought to be the overall rate constant of open complex formation. More recently, it has been recognized that formation of RP_{o2} and/or RP_{init} (in excess NTP) may be a reversible process at some promoters on the time scale of productive initiation, and that an equilibrium definition of binding specificity ($K_{obs}^{specific}/K_{obs}^{nonspecific}$) may therefore be applicable. This is yet another reason for the need to understand the origins of stability of specific promoter complexes and of those nonspecific complexes formed at all DNA nucleotide phosphates.

Large Effects of Salt Concentration on Stability: Thermodynamic Signature of Reduction of DNA Phosphate Charge Density upon Binding. A dominant characteristic of all specific and nonspecific modes of binding of RNA polymerase to DNA is their extraordinarily large dependence on salt concentration. In the absence of Mg^{2+}, binding constants K_{obs} vary as large negative powers of the univalent salt concentration, with exponents (called SK_{obs}, the log-log derivative of K_{obs} with respect to [salt]) specific to the mode of binding, ranging from –7 to –21. These large salt-concentration dependences indicate that the phosphate charge density of the promoter DNA site is extensively reduced as a consequence of the interactions with $E\sigma^{70}$ (171). Two ways in which local DNA phosphate charge density may be reduced are (i) by interaction with positively charged regions of RNA polymerase and (ii) by local strand opening, if the phosphates in the open region were not already interacting with positively charged regions in $E\sigma^{70}$. Interpretation of the strong effect of salt concentration on the extent of binding can provide quantitative information about the contribution of charge-charge (coulombic) interactions to the binding "energetics" (see references 151, 164, and 171). In addition, salt concentration provides a powerful variable for investigating the mechanism of binding, as described in a subsequent section.

Magnitudes of SK_{obs}, which in general are different for different binding modes and even for different promoters (J.-H. Roe, Ph.D. thesis, University of Wisconsin, Madison, 1984), are among the largest observed for any protein-DNA interaction. Binding modes in which holoenzyme forms (presumably) closed complexes with double-stranded DNA (e.g., at the T7 A1 promoter at 0°C), at fragment ends, and at nonspecific sites (47, 148, 195, 214, 215) all are less salt-concentration dependent ($-SK_{obs}$, ~7 to 10) than are modes in which the DNA is locally open at the promoter or globally single stranded ($-SK_{obs}$, ~12 to 20) (47, 182, 195, 214, 215). Hence, the equilibrium distribution of holoenzyme between different binding modes depends on salt concentration, since both the stability (K_{obs}) and the specificity ratio ($K_{obs}^{specific}/K_{obs}^{nonspecific}$) change with changes in salt concentration. Nonspecific complexes of core polymerase with both double-stranded and single-stranded DNA are very salt-concentration dependent ($-SK_{obs}$ = 16 to 21 [47]).

Can some insight into the location of σ^{70} on core polymerase, and into the sites on polymerase contacted by DNA in the various binding modes, be obtained from comparisons of K_{obs} and SK_{obs} values? In a previous section we proposed that the initial closed complex (RP_{c1}) involves binding to a wide surface groove of polymerase. Based on values of SK_{obs}, it is reasonable to propose that this same surface groove is used to form complexes of $E\sigma^{70}$ holoenzyme with fragment ends and double-stranded nonpromoter DNA sites, and perhaps is used as part of the binding site in all the other complexes discussed below. Nonspecific complexes of core RNA polymerase with double-stranded DNA involve more ion release and hence more extensive reduction in DNA phosphate charge density than observed for the nonpromoter complexes of holoenzyme with double-stranded DNA. As a structural basis for this behavior, we propose that σ^{70} may block access of nonpromoter double-stranded DNA to the channel adjacent to the surface groove, but that this channel is used by core polymerase to bind double-stranded DNA, resulting in a much larger magnitude of SK_{obs} (which is comparable to that observed for open promoter complexes or nonspecific complexes of both core and holoenzyme with single-stranded DNA). Since complexes formed by holoenzyme with single-stranded DNA are more stable than those formed by core polymerase, but approximately as strongly dependent on salt concentration, the presence of σ^{70} must introduce some favorable but relatively nonspecific interactions with single-stranded DNA. Whether these occur in the channel or in a putative narrower groove like that of yeast pol II is of course unknown.

Existence of a Temperature of Maximum Stability: Thermodynamic Signature of the Hydrophobic Effect (Reduction in Exposure of Nonpolar Surface upon Binding). At low temperatures (below 20°C), the enthalpy of binding of $E\sigma^{70}$ to the λP_R promoter is large (~ 70 kcal [ca. 292.9 kJ]) and positive. Binding under these conditions is highly endothermic (i.e., requires heat) and is driven by a large positive (favorable) entropy change. Binding of $E\sigma^{70}$ to the λP_R promoter becomes increasingly favorable at higher temperatures, as would be true of any endothermic interaction. In this case, however, the interaction becomes much less endothermic at higher temperatures, exhibiting a smaller enthalpy change of ~30 kcal (ca. 125.5. kJ) near 30°C. These data indicate a large negative heat capacity change upon binding, such that the enthalpy change became more exothermic (i.e., less endothermic) by ~2.4 kcal (ca. 10.0 kJ) per degree increase in temperature (182). This analysis predicts a maximum in K_{obs} (near 42°C for $E\sigma^{70}$-λP_R), at which temperature the enthalpy change is reduced to zero. Maxima in K_{obs} in the physiological temperature range and large negative heat capacity changes are in fact common features of specific protein-DNA interactions (79, 205).

This characteristic thermodynamic behavior has been interpreted in terms of the removal of nonpolar macromolecular surface from exposure to water (i.e., the hydrophobic effect). Estimates based on model systems indicate that the amount of nonpolar surface buried in complexation is more than twice that available in rigid contact surfaces between $E\sigma^{70}$ and promoter DNA, and hence lead to the proposal that the large-scale conformational changes in polymerase that accompany binding to form RP_{o2} must bury additional nonpolar surface (79).

A comprehensive analysis of all available structural and thermodynamic data for protein processes (folding; association; ligand binding, including site-specific protein-DNA binding) led to the proposal that in many binding interactions "additional" nonpolar surface is buried as part of local protein-folding transitions that create parts of the complementary recognition surfaces (205). Analysis of the entropy change upon complexation provides a test of this hypothesis. The entropy change should contain a "thermodynamic signature" of coupled folding, just as the heat capacity change provides a signature of burial of nonpolar surface (205). Indeed many protein-ligand, protein-protein,

TABLE 1 Stabilities of $E\sigma^{70}$-DNA complexes at ~0.1 M NaCl–0.1 M $MgCl_2$ (pH ~8)

DNA sequence	Binding mode	Temp (°C)	K_{obs} (M^{-1})	Reference(s)
Nonspecific	Nonspecific	0–37	$\sim 10^5$	47
T7A1	RP_{c1}	0	$\sim 10^9$	195, 214, 215
T7A1	RP_{o2}	37	$\sim 10^{10}$	195, 214, 215
λP_R	RP_{c1}	0	$\sim 10^7$	181
λP_R	RP_{o2}	37	$\sim 10^{11}$	181

and protein-DNA interactions exhibit a much less favorable entropy change than that expected from the hydrophobic effect, based on a quantitative analysis of the burial of nonpolar surface. If this "entropy deficit" is interpreted in terms of coupled folding, an estimate of the number of amino acid residues folded is obtained which is consistent with both structural and heat capacity data in systems where these are available (205). For $E\sigma^{70}$ RNA polymerase, no high-resolution structural data are yet available. The thermodynamic analysis predicts folding of $\sim 10^2$ residues, burying additional nonpolar surface but simultaneously ordering these residues. Both at a qualitative and semiquantitative level, this proposal appears consistent with emerging structural data on polymerases. Approximately 50 disordered or flexible residues exist in the Klenow polymerase at the tip of the arm or thumb (163). These cannot be detected in the high-resolution crystal structure, but appear at the low resolution of the $E\sigma^{70}$ EM structure (Fig. 3A; 142). In addition, approximately 40 residues at the end of the thumb of T7/T3 hybrid polymerase are thought to be disordered in the crystal structure of the uncomplexed protein (202). Folding of 40 to 50 residues of $E\sigma^{70}$ in the process of forming RP_{o2} at the λP_R promoter would account for almost half of both the excess heat capacity change and the deficit in the observed entropy change.

Stability and Specificity of RNA Polymerase-Promoter Complexes

Table 1 summarizes experimentally determined stabilities (K_{obs}) of the various binding modes of RNA polymerase at the λP_R and T7 A1 promoters at one set of solution conditions (~0.1 M NaCl–0.01 M $MgCl_2$, pH ~8). Nonspecific binding of holoenzyme exhibits $K_{obs}^{NS} \cong 10^5 M^{-1}$, relatively independent of temperature (47, 195). Specific binding to the T7 A1 promoter is characterized by $K_{obs}^{T7} \cong 10^8 M^{-1}$ at 0°C (presumably forming RP_{c1}) and $K_{obs}^{T7A1} \cong 10^{10} M^{-1}$ at 37°C (presumably forming RP_{o2}) (195, 214, 215). For the λP_R promoter at a similar ionic condition, $K_{obs}^{\lambda P_R} \cong 10^{7M-1}$ at 0°C (RP_{c1}) and $K_{obs}^{\lambda P_R} \cong 10^{11} M^{-1}$ at 37°C (RP_{o2}) (180–182).

From the appropriate ratios of K_{obs}, the equilibrium specificity of holoenzyme for promoters relative to nonpromoter DNA phosphate sites at 37°C (RP_{o2}) is $\sim 10^5$ (T7 A1) and $\sim 10^6$ (λP_R). These large specificity ratios are comparable to that of *Eco*RI endonuclease and somewhat less than that of *lac* repressor ($\sim 10^7$) (see reference 172 and references therein). Similar specificities apply to RP_{o1} at 0.2 M Na^+, estimated from data in the absence of Mg^{2+}. At 0°C (RP_{c1}), equilibrium specificity is much less: $\sim 10^2$ (λP_R) and $\sim 10^3$ (T7 A1). These small specificity ratios indicate that competition between the $\sim 10^3$ promoter sites and $\sim 10^7$ nonpromoter DNA sites for $E\sigma^{70}$ should be very significant at the level of the initial closed complex.

Binding of $E\sigma^{70}$ holoenzyme to blunt-ended (*Hae*III) restriction fragments exhibits $K_{obs}^{E} \cong 10^8 M^{-1}$ at 0°C (149), indicating the importance of another type of competition in in vitro promoter-binding studies using small DNA fragments. Since end binding of holoenzyme appears to be of comparable strength and specificity to the initial closed complex (RP_{c1}), the kinetics and the extent of specific binding to a promoter may be significantly reduced if end-bound polymerase overlaps the promoter site. Therefore, large fragments (with centrally located promoter sites) and excess polymerase should be used in kinetic and equilibrium studies of promoter binding.

KINETICS OF TRANSCRIPTION INITIATION AND OF FORMATION OF INTERMEDIATE BINARY AND TERNARY COMPLEXES

This section reviews the background for discussing the current status of kinetic and mechanistic studies of promoter binding and indicates the relevance of these studies to the kinetics of transcription initiation and promoter strength. We present approximate algebraic relationships between experimentally measured rates of initiation or binding and rate constants for individual mechanistic steps, or classes of steps. These relationships, obtained by application of the "steady-state approximation" of chemical kinetics to all intermediates, provide a basis for the qualitative statements regarding the kinetics of binding and initiation in the first section of this review. Many association kinetic data have previously been interpreted in terms of a two-step mechanism of formation of an open complex (RP_{o2} or RP_{init}), in which the first (binding) step is assumed to be rapidly reversible and the second (isomerization) step is assumed to be irreversible. Since the mechanism of formation of RP_{o2} is now thought to include at least four steps, and since all of these steps may be reversible, previous analyses of kinetic data which assumed an irreversible two-step mechanism should be reconsidered. The conclusions of the steady-state analysis presented below are (to our

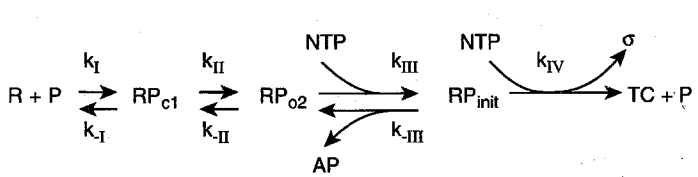

FIGURE 6 Mechanism (*B*). RP_{c1} represents the initial "closed" complex, RP_{o2} is the final binary open complex, RP_{init} is a ternary (initiated) complex, AP is abortive product, and TC is the transcribing complex.

knowledge) not available elsewhere for RNA polymerase-promoter interactions and transcription initiation. We recognize that for many readers this will represent a new and seemingly complex way of discussing these topics. However, it provides essential fundamental insights into the factors responsible for promoter strength and provides a basis for understanding the relationship between kinetic studies and mechanism. We also recognize that our use of the steady-state approximation, while the only route to algebraic rather than numerical solutions of the differential equations of kinetics, is approximate and therefore less suitable for analysis of kinetic data than a global nonlinear fit. However, in our opinion the above-mentioned fundamental insights are best captured by this approximate analysis. Readers are invited to work though the algebraic details.

Velocity of Initiation of Transcription (V_{TC})

In vivo, relative velocities of forming transcribing complexes from different promoters are deduced from a comparison of levels of expression of a common reporter protein (e.g., β-galactosidase or galactokinase). In vivo comparisons of this type may be complicated by uncontrolled variables, including changes in the promoter copy number (which can be estimated but not accurately determined) and variations in mRNA stability and translation initiation frequencies, resulting from differences in the 5′ ends of the mRNA sequences of the different promoter constructs.

Velocities of transcript production in vitro are routinely determined from multiround runoff transcription assays on promoter-containing DNA fragments. These should be done under conditions of excess active RNA polymerase (total concentration $[R]_T >> [P]_T$, so $[R]_F \cong [R]_T$). Incorporation of radiolabeled NTP into polymeric RNA product is monitored as a function of time. If the velocity of elongation is rapid in comparison to that of forming a transcribing complex (V_{TC}), a condition which can be obtained by use of a sufficiently short transcribed sequence (the elongation rate is typically ~50 to 100 nucleotides s^{-1}) (138), then this assay yields V_{TC}.

Steady-State Analysis of a Prototype Mechanism. The general form of equation 1, reproduced here, is applicable to describe the steady-state rate of production of transcribing complexes at a promoter:

$$V_{TC} \equiv \frac{1}{[P]}\frac{d[TC]}{dt} = \frac{V_{TC}^{max}[R]_F}{K_m + [R]_F} \qquad (1a)$$

In this section, we indicate how the parameters (V_{TC}^{max}, K_m) of equation 1a are related to the individual rate constants of a mechanism. The simplest inclusive mechanism representative of the classes of steps involved in forming a transcribing complex [mechanism (*B*): identical to mechanism (*A*) but with symbols in place of words] is shown in Fig. 6.

We use mechanism (*B*) to describe general features of the classes of steps of transcription initiation and of the analysis of kinetic data. We start with mechanism (*B*) instead of more detailed versions thereof for three reasons: (i) the general principles of the analysis are independent of the number of steps in the mechanism, but the algebra is simple for a smaller number of steps; (ii) even a more complete mechanism [e.g., mechanism (*E*) (Fig. 9), see below] will undoubtedly prove incomplete as more intermediates are discovered; and (iii) the first two steps (I, II) of mechanism (*B*) correspond to the classical two-step mechanism previously widely used to analyze binding kinetic data. Limitations of mechanism (*B*) in analyzing kinetic data are summarized later in this article.

Our notation for describing the rate constants has been carefully chosen. Rate constants that refer to the classes of steps specified in mechanisms (*A*) and (*B*) are designed by the roman numerals corresponding to those classes of steps. Rate constants with roman numeral subscripts therefore describe collections of reaction steps, rather than a single (elementary) reaction step. Subsequently we will use arabic number subscripts on rate constants that describe the more or less "elementary" steps in a mechanism including all known intermediate complexes.

Contributions to the initiation velocity V_{TC} from the individual steps of mechanism (*B*), or more detailed versions thereof, are best discussed using the reciprocal of equation 1a:

$$V_{TC}^{-1} = (V_{TC}^{max})^{-1} + K_m(V_{TC}^{max})^{-1}[R]_F^{-1} \qquad (2)$$

For mechanism (*B*), analyzed in the steady-state approximation (assuming the concentrations of all intermediates are independent of time) in excess NTP, one obtains the approximate (but algebraic) results:

$$V_{TC}^{-1} \cong (k_I[R]_F^{-1}) + k_{II}^{-1}\{1+(K_I[R]_F)^{-1}\} + k_{III}^{-1}\{1+K_{II}^{-1} + (K_IK_{II}[R]_F)^{-1}\} + k_{IV}^{-1}\{1+K_{III}^{-1}(1+K_{II}^{-1}) + (K_IK_{II}K_{III}[R]_F)^{-1}\} \qquad (3)$$

so that

$$(V_{TC}^{max})^{-1} \cong k_{II}^{-1} + k_{III}^{-1}(1+K_{II}^{-1}) + k_{IV}^{-1}\{1+K_{III}^{-1}(1+K_{II}^{-1})\} \qquad (4)$$

and

$$K_m(V_{TC}^{max})^{-1} \cong k_I^{-1} + (K_Ik_{II})^{-1} + (K_IK_{II}k_{III})^{-1} + (K_IK_{II}K_{III}k_{IV})^{-1} \qquad (5)$$

Though this analysis is approximate, it allows the following general conclusions. The first term in equation 3 represents the reciprocal of the initial velocity of forming RP_{c1}. The sum of the first two terms is the reciprocal of the initial velocity of appearance of the final binary open complex (RP_{o2}). The first three terms represent the reciprocal of the initial velocity of formation of the ternary complex RP_{init} (cf. equation 11, below). Each of these components of V_{TC}^{-1} may be determined independently, as discussed below. In equation 3, if the equilibria favor binding/isomerization (i.e., $K_I[R] >> 1$, $K_{II} >> 1$, etc.), the smallest forward rate constant of the set ($k_I[R]$, k_{II}, k_{III}, k_{IV}) is the principal determinant of the velocity V_{TC}, because the *reciprocals* of these rate constants are added to obtain V_{TC}^{-1}.

Kinetic Inferences: Central Role of $[R]_F$. What determines the rate of transcript initiation at a given promoter in *E. coli*? What is the relationship between promoter sequence and rate constants of the various isomerization (k_{II}, k_{III}) and dissociation (k_{-II}, k_{-III}) steps which specify the kinetic behavior of an individual promoter?

Equation 3 demonstrates that the velocity of productive initiation per promoter is determined by the concentration of free RNA polymerase and by the intrinsic rates (rate constants) of

$$R + P \underset{k_{-I}}{\overset{k_I}{\rightleftarrows}} RP_{c1} \underset{k_{-II}}{\overset{k_{II}}{\rightleftarrows}} RP_{o2}$$

FIGURE 7 Mechanism (*C*).

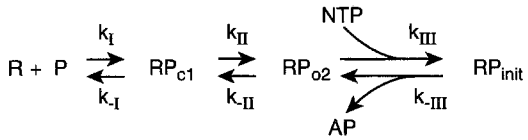

FIGURE 8 Mechanism (*D*).

the various binding, isomerization, and clearance events at the promoter. Key elements of system design in vivo are therefore (i) the concentration of free polymerase $[R]_F$; (ii) the cytoplasmic variables (e.g., *T*, viscosity, ion concentrations) which determine the rate constant for initial association (k_I); and (iii) the sequence-specific polymerase-promoter interactions and cytoplasmic variables (*T*, ion concentrations) which determine K_m and the rate constants k_{-I}, k_{II}, and k_{-II} (and probably also contribute to k_{III}, k_{-III}, and k_{IV}). We discuss $[R]_F$ here. Discussion of the other design elements is deferred to later sections.

In vivo, the total concentration of RNA polymerase is determined by a balance between the net rates of synthesis and degradation and by the cytoplasmic volume under the growth conditions examined. However, macromolecular crowding at the high protein and nucleic acid concentration of the cytoplasm may make the effective concentration of RNA polymerase substantially higher than its physical concentration (33). In vivo, RNA polymerase is thought to be distributed approximately as follows: actively transcribing core, ~50%; specifically bound holoenzyme, ~25%; nonspecifically bound core and holoenzyme, ~25%; and free holoenzyme, $\lesssim$1% (144). The free RNA polymerase concentration $[R]_F$ may be buffered against change during cell growth at constant conditions by the large reservoir of nonspecifically bound polymerase, which should equilibrate rapidly with free polymerase on the time scale of the events in mechanism (*B*) (228); $[R]_F$ in principle may vary with changes in growth conditions which change ion or water concentrations in the cytoplasm (33). In vitro, transcription and binding experiments should be performed in excess RNA polymerase (i.e., total concentration $[R]_T >> [P]_T$), so $[R]_F \cong [R]_T$ throughout the time course. This reduces experimental uncertainties regarding $[R]_F$ as a consequence of nonpromoter binding and allows a pseudo-first-order analysis of the kinetics.

In addition to its roles in drastically reducing and buffering the free RNA polymerase concentration, nonspecific binding may also serve to enhance the rate of initiation (V_{TC}) formation by reducing the dimensionality of the diffusional search for the promoter from three to one (or two) (229). However, while evidence exists for sliding of RNA polymerase on nonspecific DNA (104), no significant contribution of sliding to the velocity V_{TC} of productive initiation has yet been demonstrated.

Relationship of Kinetic Studies of Binding to Initiation Velocity

Association and dissociation kinetic studies performed as a function of temperature, salt concentrations, and other solution variables in the absence of NTP provide a means of investigating the details of the mechanism of forming RP_{o2} at a particular promoter sequence and/or determining the rates of the steps (and substeps) of the simplest inclusive binding mechanism (*C*), corresponding to the first two classes of steps of mechanism (*B*) (Fig. 7). Separation assays, in which polymerase-DNA complexes are separated from free radiolabeled promoter DNA on nitrocellulose membrane filters (e.g., references 91 and 180) or by electrophoretic gel mobility shifts (e.g., references 63, 66, and 211), have been used to measure the kinetics of formation and dissociation of RP_{o2}. Most spectroscopic methods are of insufficient sensitivity to detect complexation at the concentrations required for equilibrium and kinetic studies.

The abortive initiation assay (142, 145), subsequently modified for fluorescence detection of pyrophosphate release in nucleotide incorporation (13, 14), in which a subset of initiating NTP is used to obtain repetitive synthesis of a short (abortive) RNA product (AP), provides a means of investigating the steps and substeps of the simplest inclusive mechanism (*D*), of formation of a particular RP_{init} (Fig. 8). [Mechanism (*D*) corresponds to the first three classes of steps of mechanism (*B*).] The details of prototype mechanism (*D*) may also be investigated by performing binding assays in the presence of a subset of NTP yielding uniquely stable ternary complexes (180).

In general, formation of both RP_{o2} and RP_{init} (in excess NTP) is a reversible process, which proceeds to equilibrium and not to completion. In excess RNA polymerase ($[R]_T >> [P]_T$), the experimentally determined reversible kinetics of forming both RP_{o2} and RP_{init} are observed to fit rate equations of the general form (189; P. J. Schlax, Ph.D. thesis, University of Wisconsin, Madison, 1994).

$$\frac{d[RP_x]}{dt} = \alpha_{RP_x}[P]_T - \beta_{RP_x}[RP_x] \tag{6}$$

where RP_x is either RP_{o2} or RP_{init}. Because the equilibrium concentration $[RP_x]^{eq}$ is independent of time ($d[RP_x]/dt = 0$ at equilibrium), the promoter-concentration-independent, experimentally determined quantities α_{RP_x} and β_{RP_x} must be related by the equilibrium condition:

$$\frac{\alpha_{RP_x}}{\beta_{RP_x}} = \frac{[RP_x]^{eq}}{[P]_T} \tag{7}$$

Substitution of equation 7 into equation 6 yields the result that β_{RP_x} is the observed first-order rate constant for the approach of $[RP_x]$ to its equilibrium value. Mechanisms (*C*) and (*D*), as well as more detailed versions thereof [e.g., mechanism (*E*), below], when analyzed under pseudo-first-order conditions using the steady-state approximation, invariably yield rate equations of the form of equation 6 (189; Schlax, Ph.D. thesis).

In excess RNA polymerase, the quantity $\alpha_{RP_x}^{-1}$ is found to be a linear function of $[R]_T^{-1}$, both experimentally and by steady-state analyses of C, D, and extended versions thereof:

$$\alpha_{RP_x}^{-1} = (k_{a,RP_x}[R]_T)^{-1} + k_{i,RP_x}^{-1} \tag{8}$$

In equation 8, $\alpha_{RP_x}^{-1}$ is equivalent to the previously defined time constant $\tau_{obs}^{init} = k_{obs}^{-1}$ for irreversible association of RNA polymerase with a promoter (142). Double-reciprocal plots analogous to that indicated by equation 8 are commonly used in chemical kinetics to separate concentration-dependent and concentration-independent contributions to an observable kinetic quantity. Equation 8 defines two observable composite rate constants, a second-order ("association") rate constant (k_a) and a first-order ("isomerization") rate constant (k_i); these can be interpreted in terms of collections of elementary rate constants of the corresponding mechanism. Different mechanisms result in different interpretations of k_a and k_i (see below).

Dissociation of RP_{o2}, made irreversible by addition of a polyanionic competitor to bind free R and monitored by the reduction in concentration of RP_{o2} with time, is invariably found to be a first-order process characterized by the rate constant k_d.

For the two-step summary mechanism (*C*) of formation of the open complex RP_{o2}, analyzed using the steady-state approximation:

$$k_{a,RP_{o2}}^{-1} \cong (K_I k_{II})^{-1} + k_I^{-1}$$

$$k_{i,RP_{o2}}^{-1} \cong k_{II}^{-1} \tag{9}$$

and

$$k_{d,RP_{o2}}^{-1} \cong k_{II}^{-1} + K_{II} k_{-I}^{-1}$$

For the analogous three-step summary mechanism (*D*) of formation of an initiating complex RP_{init} to the same approximation:

$$k_{i,RP_{init}}^{-1} \cong k_{II}^{-1} + k_{III}^{-1} + (K_{II} k_{III})^{-1}$$

$$k_{a,RP_{init}}^{-1} \cong (K_I K_{II} K_{III})^{-1} + (K_I K_{II})^{-1} + k_I^{-1}$$

$$k_{i,RP_{init}}^{-1} \cong k_{II}^{-1} + k_{III}^{-1} + (K_{II} k_{III})^{-1} \tag{10}$$

and

$$k_{d,RP_{init}}^{-1} \cong k_{III}^{-1} + K_{III} k_{-II}^{-1} + K_{III} K_{II} k_{-I}^{-1}$$

Since formation of both RP_{o2} and RP_{init} is in general reversible, two different experimental strategies exist for determining α_{RP_x} as a function of $[R]_T$, and hence for determining the corresponding rate constants k_{a,RP_x} and k_{i,RP_x}. (i) From the kinetics of the approach of $[RP_x]$ to its time-independent equilibrium value $[RP_x]^{eq}$, one obtains the so-called "relaxation" rate constant β_{RP_x} and calculates α_{RP_x} from equation 7. (The dissociation rate constant k_{d,RP_x} is also obtained from this relaxation experiment.) (ii) From the initial rate per promoter, defined as $V_{o,RP_x} \equiv [P_T]^{-1} (d[RP_x]/dt)_{t\to 0}$, one obtains α_{RP_x} directly from $V_{o,RP_x} = \alpha_{RP_x}$. Hence:

$$V_{o,RP_x}^{-1} = (k_{a,RP_x} [R]_T)^{-1} + k_{i,RP_x}^{-1} \tag{11}$$

which may also be expressed in the form of equation 1 with $V_{max} = k_{i,RP_x}$ and $K_m = k_{i,RP_x}$. (This method is impractical with the abortive initiation assay.)

At a specified free RNA polymerase concentration $[R]_F$, the initial velocities of formation of RP_{o2} and RP_{init} may be related, using the steady-state approximation, to the initiation velocity V_{TC} (equation 3) by:

$$V_{TC}^{-1} \cong V_{o,RP_{o2}}^{-1} + k_{III}^{-1}\{1 + K_{II}^{-1} + K_I (K_{II}[R]_F)^{-1}\}$$
$$+ k_{IV}^{-1}\{1 + K_{III}^{-1}(1 + K_{II}^{-1}) + (K_I K_{II} K_{III}[R]_F)^{-1}\} \tag{12}$$
$$\cong V_{o,RP_{init}}^{-1} + k_{IV}^{-1}\{1 + K_{III}^{-1}(1 + K_{II}^{-1}) + (K_I K_{II} K_{III}[R]_F)^{-1}\}$$

A plausible correlation has been observed between in vivo values of V_{TC}, inferred from levels of expression of a β-galactosidase reporter protein from the transcript, and in vitro values of $V_{o,RP_{init}} = \alpha_{RP_{init}} = \tau_{obs}^{-1}$ measured at excess $[R]_T \cong [R]_F = 30$ nM under ionic conditions which, while nonphysiological, may be equivalent in the sense of yielding kinetic and thermodynamic constants similar to those appropriate in vivo (143). (Ignoring crowding effects, an in vivo concentration $[R]_F = 30$ nM corresponds to ~1% of the total holoenzyme concentration $[R]_T$ in the cell.) The existence of this correlation indicates that the in vivo clearance/elongation term $k_{IV}^{-1}\{1 + K_{III}^{-1}(1 + K_{II}^{-1}) + (K_I K_{II} K_{III}[R]_F)^{-1}\}$ is kinetically insignificant in comparison to $V_{o,RP_{init}}^{-1}$ at the promoters investigated. Clearly these terms are significant for other promoters (22, 57, 58).

Relationship of Experimentally Observed Rate Constants (k_a, k_i, k_d) and Rate Constants of the Classical Two-Step Mechanism (*C*) to Rate Constants of Individual Mechanistic Steps

Mechanism (*C*) for formation of RP_{o2} corresponds to the classical two-step mechanism of open complex formation, an irreversible version of which has been extensively used to interpret association kinetic data (k_a, k_i, k_d) in terms of the quantities K_I (previously designated as K_B [142]) and k_{II} (previously designated k_f [142]) (compare equations 9 and 10). Although this two-step mechanism, like the analogous two-step treatment of enzyme-catalyzed reactions, summarizes the division between bimolecular and unimolecular steps, mechanism (*C*) is insufficient in detail for any promoter. Neither mechanism (*C*) nor the rapid equilibrium approximation conventionally assumed to analyze it should be used to interpret k_a and k_i without direct experimental evidence for their applicability. In particular, the ratio of k_a to k_i should not be interpreted as the equilibrium constant for formation of RP_{c1} without verifying these approximations.

The structural information reviewed above and the kinetic/mechanistic studies reviewed in a subsequent section indicate that the minimal mechanism of formation of RP_{o2} is shown in Fig. 9.

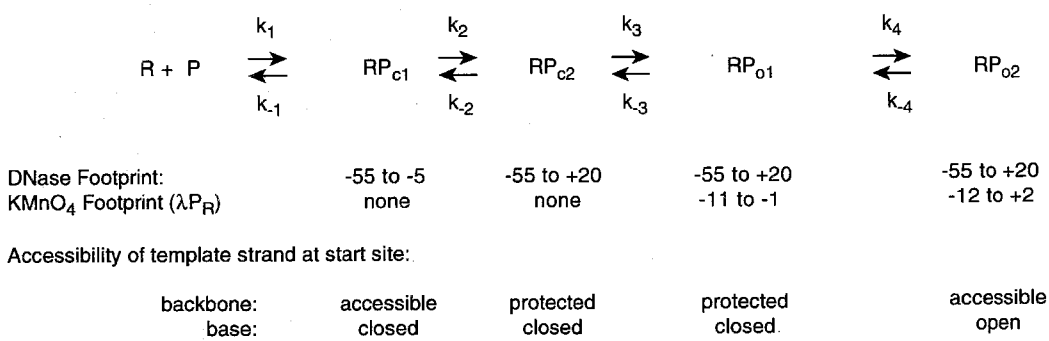

FIGURE 9 Mechanism (*E*).

For both λP_R and *lac*UV5 promoters, the isomerizations between the two closed complexes are thought to be the rate-determining steps in both association and dissociation of RP_{o2} under most experimental conditions.

To analyze effects of sequence changes and other variables on the kinetics of open complex formation, one must answer the following questions: (i) how are the experimentally determined rate constants k_a, k_i, and k_d and the kinetic and thermodynamic parameters of the two-step mechanism K_I, k_{II}, and k_{-II} (previously interpreted as k_a, k_i, and k_d, respectively) related to the rate constants and equilibrium constants of mechanism (*E*)? and (ii) how (if at all) would these interpretations change if (or when) more intermediates between RP_{c1} and RP_{o2} are discovered and mechanism (*E*) expands to include them?

Steady-state (approximate) solutions for k_a and k_i for linear mechanisms of the form of E with any number of isomerization steps are available (e.g., Schlax, Ph.D. thesis). Though approximate (see reference 60) and therefore less satisfactory than global numerical analysis of kinetic data (e.g., references 15, 152, 153, 240, and 241) in any particular context, a steady-state analysis is the only reasonably general approach to obtain approximate algebraic relationships between experimentally determined and elementary rate constants, and provides a semiquantitative description of the effect of reversibility on the observed kinetics. In the appropriate limit, the steady-state approximation becomes equivalent to the less general approximation of rapid equilibrium, as demonstrated below.

For the particular case of the four-step mechanism (*E*), analyzed with the steady-state approximation, the relationships between experimental rate constants (k_a, k_i, k_d) and the two-step (composite) and four-step ("elementary") rate constants are:

$$k_a^{-1} \cong k_i^{-1} + (K_I k_{II})^{-1} \cong (K_1 k_2)^{-1} (1 + k_2 k_{-1}^{-1} + Q_1) \quad (13)$$

$$k_i^{-1} \cong k_{II}^{-1} \cong k_2^{-1} (1 + Q_1)(1 + Q_2) \quad (14)$$

$$k_d^{-1} \cong k_{-II}^{-1} + K_{II} k_I^{-1} \cong K_4 K_3 k_{-2}^{-1} (1 + k_2 k_{-1}^{-1} + Q_1) \quad (15)$$

where the initial factor on the right-hand side of each equation, highlighted in boldface type, is the term thought to be the dominant contributor to the experimentally determined rate constant for the λP_R promoter under most conditions investigated (180–182, 217) and $k_2 k_{-1}^{-1}$, Q_1, and Q_2 are "correction" terms which must be small in comparison to unity for the initial factor to indeed be the dominant term in each equation. In equations 13 through 15 the quotients of rate constants specified as Q_1 and Q_2 in the parenthetical correction factors are:

$$Q_1 = k_{-2} k_3^{-1} (1 + k_{-3} k_4^{-1}) \quad (16)$$

$$Q_2 = \frac{k_3^{-1} + k_4^{-1} + (K_3 k_4)^{-1}}{k_2^{-1} + (K_2 k_3)^{-1} + (K_2 K_3 k_4)^{-1}} \quad (17)$$

and it is assumed that $k_I = k_1$ (see below).

The following is a brief paraphrase of equations 13 through 15. In equation 13, k_a is typically interpreted as $K_1 k_2$; for this to be a reasonable interpretation, the corresponding correction terms must be small ($k_2 k_{-1}^{-1} << 1$, $Q_1 << 1$). From equation 14, for $k_i \cong k_2$, it is necessary that $Q_1 << 1$ and $Q_2 << 1$. If $k_a = K_1 k_2$, then (cf. equation 15) $k_d = K_4 K_3 k_{-2}^{-1}$, consistent with $k_d / k_d = K_1 K_2 K_3 K_4$.

The Initial Bimolecular Association Step. For promoters in the interior of relatively small DNA fragments (<1 kbp), it is likely that the initial association step is at or near the value predicted from an orientation-corrected calculation of the rate of productive encounters resulting from three-dimensional diffusion of the DNA and the polymerase in solution:

$$R + P \xrightarrow{k_1} R\,P_{c1}$$

Hence $k_I = k_1$, the elementary rate constant for the diffusion/collision process. If all collisions between the promoter site and polymerase lead to reaction, we estimate $k_1 \cong 3 \times 10^9\ M^{-1}\ s^{-1}$ (229). Measurements of k_1 for association of polymerase with strong promoters on DNA fragments yield estimates of k_I in the range of 2×10^8 to $3 \times 10^8\ M^{-1}\ s^{-1}$ (22, 27, 169), suggesting that ~10% of collisions have the correct orientation (neglecting long-range coulombic effects, which are of unknown importance). On larger DNA molecules (>>1 kbp), sliding or other mechanisms of facilitated (one- or two-dimensional) diffusion on nonspecific DNA may extend the DNA target from the promoter site to much or all of the DNA molecule, thereby increasing k_I above the predicted three-dimensional diffusion limit (229). Since such facilitating processes will also increase k_{-1}, they will only increase the velocity of initiation V_{TC} (or of open complex formation) if the initial binding step is relatively irreversible (i.e., $k_I[R]_T >> k_{-I}$).

Dissociation of RP_{c1}: Relationship of k_{-I} to k_{-1}. If $k_I = k_1$, as discussed above, then the relationship between k_{-I} and elementary rate constants is:

$$k_{-I} = k_{-1} (1 + Q_2)^{-1} \quad (18)$$

Only if the isomerization of RP_{c1} to the next intermediate complex (RP_{c2}) is the rate-determining step in the forward direction (mathematically equivalent to the statement that k_2^{-1} is much larger than k_3^{-1}, k_4^{-1}, or the other terms in Q_2, so that $0 \cong Q_2 << 1$) will k_{-I} in the two-step mechanism C be identical with the elementary rate constant k_{-1} for dissociation of RP_{c1}.

Isomerization of RP_{c1}: Relationship of k_{II} to k_2. The composite isomerization rate constant k_{II} corresponds graphically to the intercept term in a "τ-plot" (142) of $\alpha_{RP_{o2}}^{-1} = \tau_{obs,RP_{o2}}$ versus $[R]_T^{-1}$, generated from data from a binding kinetic assay that detects RP_{o2} directly. Alternatively, k_{II} may be estimated from the abortive initiation assay for RP_{init}, provided that $k_{III} >> k_{II}$ and $K_{II} >> 1$. If formation of the species detected (RP_{o2}, RP_{init}) is not essentially irreversible, then it is significantly more difficult to obtain $\alpha_{RP_x}^{-1}$ and k_{II} from kinetic data. Lack of data regarding reversibility complicates the interpretation of some of the kinetic literature in this field; such studies indicate where large effects can be observed but in our opinion should not be considered quantitative or definitive without a demonstration of irreversibility or a correction for reversibility.

Subject to the above considerations regarding its measurement, the quantity k_i is interpreted as:

$$k_i^{-1} \cong k_{II}^{-1} \cong k_2^{-1} (1 + Q_1)(1 + Q_2) \quad (14)$$

If $Q_1 << 1$ and $Q_2 << 1$, corresponding to the case where the isomerization $RP_{c1} \rightarrow RP_{c2}$ is the slow step in the association direction (so that k_2 is the smallest term in Q_2), and where the forward direction of conversion of RP_{c2} and subsequent intermediates is favored kinetically ($k_3 >> k_{-2}$; $k_4 >> k_{-3}$), then $k_i = k_{II} \cong k_2$. For both λP_R and *lac*UV5 promoters, it appears that these conditions are met for experiments performed at temperatures above ~20°C. Where the DNA opening steps become kinetically

significant (presumably at lower temperature or higher salt concentration), and/or where reversal of the intermediate isomerizations becomes kinetically significant, then $k_i = k_{II} < k_2$.

Equilibrium Binding To Form RP_{c1}: Relationship of k_a/k_i to K_1. From equations 13 and 14 the ratio of experimentally determined composite rate constants k_a/k_i is interpreted as:

$$\frac{k_a}{k_i} = K_1 \frac{(1+Q_1)(1+Q_2)}{(1+k_2k_{-1}^{-1}+Q_1)} \qquad (19)$$

Only if $Q_2 << 1$ and $k_2k_{-1}^{-1}(1 + Q_1)^{-1} << 1$ does $k_a/k_i = K_1$, which is the interpretation of this quantity in the commonly used (but seldom critically tested) rapid-equilibrium limit of the two-step mechanism (C).

Dissociation of RP_{o2}: Dissociation Rate Constant k_d. For both λP_R and *lac*UV5 promoters, under the conditions studied, the isomerization $RP_{c2} \rightarrow RP_{c1}$ is thought to be the slow step in the dissociation direction. Hence, k_d^{-1} is most conveniently expressed by equation 15:

$$k_d^{-1} = K_4K_3k_{-2}^{-1}(1 + k_2k_{-1}^{-1} + Q_2) \qquad (15)$$

If $Q_2 << 1$ and $k_2 << k_{-1}$, then $k_d^{-1} \cong K_4K_3k_2^{-1}$. This behavior of k_d has been assumed in previous studies (e.g., reference 217).

Case Study: Evidence for Intermediate Closed (RP_{c2}) and Open (RP_{o1}) Complexes at the λP_R Promoter

RP_{c2}. The existence of an intermediate complex, subsequently designated RP_{c2}, in the λP_R mechanism was deduced from studies of the effects of temperature and univalent salt concentration on the phenomenological rate constants k_a, k_i, and k_d for the $E\sigma^{70}$-λP_R promoter interaction (180–182). Kinetic studies as functions of these variables provide information about whether an observed rate constant (e.g., k_a, k_i, or k_d) is a composite of equilibrium constants and rate constants for multiple steps, or whether it is a rate constant of an elementary step. At the time, the question was whether a two-step mechanism [as in mechanism (*C*)] was sufficient to describe the kinetics of formation of RP_{o2}, or whether an additional intermediate complex was required. The key evidence that an additional intermediate and therefore a three-step mechanism was required for open complex formation came from an inability to reconcile both the salt-concentration dependence of the association rate constant k_a and the temperature dependence of the dissociation rate constant k_d with the same two-step mechanism. (i) The association rate constant k_a was substantially less than the diffusion limit, and decreased strongly with increasing salt (NaCl) concentration: $k_a \propto [Na^+]^{-12}$ in the absence of Mg^{2+} or other cations. This findings indicated that an intermediate (RP_{c1}), in rapid equilibrium with reactants, contributed to the association kinetics. If $k_a \cong K_Ik_{II} \cong K_1k_2$ (cf. equation 13; rapid equilibrium implies that $k_{-1} >> k_2$), then the strong Na^+ dependence of k_a is that of K_1, the equilibrium constant for forming RP_{c1} (i.e., $K_1 \propto [Na^+]^{-12}$). (ii) The dissociation rate constant k_d decreased with increasing temperature. The resulting large negative Arrhenius activation energy could not result from an elementary rate constant (i.e., that of a single step), and therefore one or more intermediates also contributed to the dissociation kinetics. Consistent with this, k_d was also very Na^+ dependent: $k_d \propto [Na^+]^8$. These intermediates in dissociation, now known to be RP_{c2} (and RP_{o1}), which apparently are in rapid equilibrium (see equation 15) with RP_{o2} on the time scale of conversion of RP_{c2} to RP_{c1}, could not be the same as RP_{c1}. Hence, it was necessary to conclude that the mechanism contained at least three steps. The activation energies obtained from k_a and k_d are themselves strikingly temperature dependent, an observation which is consistent with the overall large heat capacity change in the process of open complex formation and indicates that large changes in the amount of water-accessible nonpolar surface occur in the kinetically significant steps of both association and dissociation. The interconversions of the two closed complexes (RP_{c1}, RP_{c2}) were deduced to be the slow steps in both directions of the process under the conditions investigated.

A concurrent set of kinetic studies at the *lac*UV5 promoter (24) indicated that a mechanism involving at least three steps was required to describe formation of RP_{init}. Structural studies (204) indicated that the proposed intermediate complex following RP_{c1} on the association pathway was not open (because the cytosines were inaccessible to methylation). Hence an intermediate closed complex (RP_{c2}) is a common intermediate in the mechanisms of open complex formation at λP_R and *lac*UV5 parameters. The details of these studies on *lac*UV5 and λP_R have been thoroughly reviewed (23, 127).

RP_{o1}. The existence of an intermediate open complex (RP_{o1}) was deduced from dissociation kinetic and $KMnO_4$ footprinting data in the presence and absence of Mg^{2+} (217, 218). Dissociation of open complexes in the presence of Mg^{2+} was less dependent upon the concentration of Mg^{2+} and under many conditions was slower than predicted from studies in the absence of Mg^{2+} (217). Parallel $KMnO_4$ reactivity studies showed that addition of Mg^{2+} extended the "open" (i.e., $KMnO_4$-reactive) region into the transcription start site (+1, +2) (218). This result was anticipated by equilibrium binding studies (195, 214, 215) on the T7 A1 promoter. In this case Mg^{2+} exerted a large competitive effect on binding of $E\sigma^{70}$ holoenzyme at 0°C (presumably forming RP_{c1}), but exerted only a small competitive effect at 37°C (presumably because its competitive effect on stability of open complexes was counterbalanced by its role in conversion of RP_{o1} to RP_{o2}).

With the inclusion of univalent (e.g., Na^+) and Mg^{2+} cations in the appropriate steps, mechanism (*E*) becomes mechanism (*E'*) (Fig. 10). Steps involving major reductions in DNA axial charge density (initial binding of RNA polymerase as RP_{c1}; opening to form RP_{o1}) are affected by cation concentration; cation uptake (in the thermodynamic sense) is required in the direction in which the DNA axial charge density increases (dissociation or RP_{c1}; closing of RP_{o1}) (182). The Mg^{2+} requirement for formation of RP_{o2} is a specific one, not satisfied by univalent cations and only inefficiently satisfied by Ca^{2+} or Ba^{2+} (218). Although the site of binding of Mg^{2+} in RP_{o2} is unknown, it is possible that some or all of these ~3 Mg^{2+} ions bind to a "trio of catalytic carboxylates," as observed at the base of the DNA-binding channel of several single-subunit polymerases and thought to be the catalytic site for RNA phosphodiester bond formation (208). This would be consistent with the proposal that RP_{o2}, the final binary open complex, is poised to carry out polymerization upon binding the initiating NTP.

INTRINSIC REGULATION: EFFECTS OF PROMOTER SEQUENCE AND STRUCTURE

Introduction: Homology to the Core Consensus Promoter Sequence

Most studies of effects of promoter primary structure on promoter function have focused on the major structural elements of

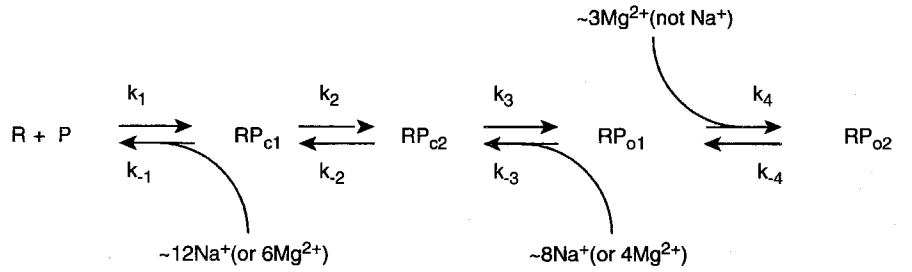

FIGURE 10 Mechanism (E').

the core promoter, namely the sequence of the –35 and –10 hexamers and the length of the intervening spacer region. The fundamental premises of these studies are that the statistically defined consensus core promoter structure yields the highest initiation velocity (called the promoter strength) of any core promoter in a given context (i.e., choice of flanking sequences) and that the reductions in initiation velocity observed for other core sequences in the same context correlate with their reduced homology to the consensus core promoter. Although important exceptions exist, results to date in general support both of the above hypotheses.

A qualitative definition of a given promoter's homology to the consensus promoter sequence is obtained from the number of positions at which the sequence matches the consensus sequence. In general the in vivo strength of a promoter correlates with its qualitative homology to the consensus sequence. Mutations increasing homology to consensus increased promoter strength; with one exception, mutations reducing homology decreased promoter strength (83). A direct relationship between the qualitative homology of a promoter sequence to consensus and in vivo and in vitro promoter strength was shown for a series of variants in the –10 and –35 regions of the $\lambda P_{RM\text{-}UP1}$ promoter (220). Changes in the length of the spacer region separating the –10 and –35 hexamers affected promoter strength in the direction expected from the homology-to-consensus model (8, 156, 220). In addition, promoter strength depended on which nonconsensus base was present at a given position (83, 220).

A more quantitative approach has therefore been used to define homology and interpret or predict relative promoter strengths. In this approach, at each position in the –10 and –35 hexamers, the homology score of each base is assigned based upon its statistical frequency of occurrence (156). For example, for three promoters differing only in the base at position –8, the promoter with an A (the "consensus" base, found in 56% of scored promoters) is given a higher score than the promoter with a C (observed 20% of the time), which, in turn, is given a higher score than the promoter with a G (observed 8% of the time) (131). The contribution of the consensus base at a particular position to the overall homology score also depends on its frequency of occurrence. For example, the consensus A at position –11 (occurring with a frequency of 76%) contributes more to the homology score than an A at position –8 (with a frequency of 56%) (131). A contribution to homology score from spacer length is also included, based on frequency of occurrence.

Use of a quantitative homology score assumes that the bases at different positions are recognized independently. Quantitative studies of other site-specific protein-DNA interactions, however, show many examples of context-dependent effects of base substitution, suggesting that this assumption may not be generally valid for polymerase-promoter interactions (e.g., references 129 and 154).

In spite of the above caveat, a systematic comparison of in vivo activity as a function of –10 and –35 sequence in the context of the P22 P_{ant} promoter indicated that the quantitative homology score can be a good predictor of promoter strength (76, 155). The wild-type P22 P_{ant} promoter deviates from consensus only at position –8. One at a time, each base in the –10 and –35 hexamers was systematically changed to the three alternatives. In all but one case, the relative promoter strengths of these variants correlate with the statistical probability of occurrence of that base in the collection of core promoter sequences. At position –8, the wild-type (nonconsensus) base pair was slightly preferred to the consensus base pair.

Is a homology score based solely on the core promoter structure sufficient to predict relative promoter strength? The answer in general is "No." For example, the consensus core promoter sequence, in the context of a particular flanking sequence, was found to be ~17-fold less strong than the T7 A1 promoter, which has three nonconsensus bases in the core promoter but a different flanking sequence (49). Hence, flanking sequences may play a large role in determining promoter strength, as discussed in more detail subsequently. Changes in flanking sequences may either stabilize or destabilize intermediates or the RP_{init} complex and therefore influence promoter strength (57, 58). Thus, the context (i.e., flanking sequences) of the core promoter must be considered when predicting or comparing promoter strength. In addition, the design of promoters with very high homology to consensus may simultaneously introduce constraints which reduce promoter strength (e.g., RNA polymerase may bind so tightly that it does not readily advance through RP_{init} to form an elongating complex) (57, 58).

Changes in the –35 and –10 Hexamers

Do changes in the individual regions of a promoter (e.g., the sequence of the –35 hexamer, the –10 hexamer, or the flanking regions of the promoter) affect single steps or many steps in transcription initiation? The classical "bipartite model" of promoter structure-function (69), which was proposed contemporaneously with the two-step kinetic mechanism, hypothesized that the –35 sequence affected primarily the initial binding of polymerase to the promoter and the –10 sequence affected primarily opening of the DNA. This model has been tested in vitro by comparing various mutants with their wild-type counterparts.

The general conclusion from these studies is that the simple bipartite model for promoter structure-function is not correct. The bipartite model was a useful first step in defining a model for promoter structure and function. However, since the mechanism of open complex formation has been shown to contain at least four steps, it is clear that the bipartite model is no longer adequate.

A comprehensive study of the *lac*UV5 promoter illustrates the apparent limitations of the bipartite model. The four possible variants at each position of the –35 hexamer of the *lac*UV5 promoter were examined. Single-round in vitro transcription and abortive initiation assays indicated that both the time required to obtain the appropriate product and the final amount of product depend on the sequence of the –35 hexamer (109). Though quantitative interpretation of these data in terms of equation 3 is not possible, the qualitative conclusion of this study, namely, that changes in –35 sequence affect multiple steps of transcription initiation, appears justified. Which isomerization steps are affected is not yet known.

Changes in the –10 sequence of a promoter also appear to affect the rates of multiple steps of formation of RP_{init}. The abortive initiation assay was used to investigate three variants of *lac*P_1 (*lac*P_1, *lac*P_s, and *lac*UV5): *lac*UV5 has a consensus –10 region, *lac*P_s has one nonconsensus base pair in the –10 region, and the wild-type *lac*P_1 has two nonconsensus base pairs in the –10 region (135). Subject to the assumption of irreversibility, k_a and k_i were both largest for the *lac*UV5 promoter and smallest for the wild-type promoter, suggesting that more than one step is affected by the changes in sequence of the –10 hexamer (cf. equations 13 and 14). Do the kinetic effects of these changes in sequence correlate with structural differences between complexes formed at these promoters? To examine this, $KMnO_4$ was used to probe the opening regions of *lac*P_s, *lac*P_1, and *lac*UV5 promoters (186). Significant differences were observed in the pattern of reactive bases and in the extent of reactivity of those bases in the open regions of these promoters, consistent with the proposal that the –10 sequence is of importance in the steps of opening the promoter DNA to form RP_{o2}.

Changes in Spacer Length and Sequence

Any role for the –35 and –10 hexamers in determining promoter strength is likely to also involve a role of the spacer sequence in aligning the –35 and –10 hexamers (6, 7). Large effects of changing spacer length on the kinetics of open complex formation have been reported (e.g., reference 233 and references therein). Systematic changes (±1 bp) in spacer length have indicated that the consensus 17-bp length yields the largest values of k_a and k_i for the *tac* and *lac*P_s promoters (156, 207). However, the effects of spacer length on the mechanistic steps of transcription initiation appear to depend upon the sequences of the –35 and –10 hexamers. This is evident from a comparison of the effects of spacer length on the dissociation rate constant k_d at the *lac*P_s and *tac* promoters. For the *tac* promoter variants, k_d is smallest for the 17-bp spacer length and largest for the 16-bp spacer length (156). For *lac*P_s variants, k_d is largest for the 17-bp spacer and smallest for the 16-bp spacer (207). Differences in the trends exhibited by k_d for the two sets of constructs indicate that factors in addition to the length of the spacer must be important in the helix-closing (and opening) and other isomerization steps which determine k_d (see equation 15). Significant increases in k_a are observed when part of the sequence of the spacer region of $\lambda P_{RM\text{-}UP1}$ is replaced by tracts of C, G, or CG (6). Generally, these results are consistent with the hypothesis that the role of spacer length (and to a lesser extent sequence) in determining promoter strength is to orient the positions of the –35 and –10 hexamers (see reference 46).

Changes in Start Site

Although RNA polymerase exhibits a preference for ATP or GTP starts, this observation does not directly indicate whether a change in the sequence of the start site (+1) affects the velocity of transcription initiation. Few systematic studies have addressed this question. Within the context of the *lac*UV5 promoter system, a number of mutations have been constructed with altered start site sequences (100; X. F. Xiong, Ph.D. thesis, University of Wisconsin, Madison, 1992). Analysis of the transcripts programmed by these mutant promoters confirms that A start sites are chosen preferentially, but that the other three base pairs can also function as start sites. Interestingly, C start sites yield significantly lower levels of expression. The deleterious effect and bias against C start sites are also the apparent molecular basis for regulation of *Salmonella pyrC* and *pyrD* promoters in which the choice of a C start only occurs when the cell has unusually high internal pools of CTP (200). The effects of C starts may be an indirect effect of promoter sequence; e.g., they may result from a higher K_m^{NTP} for this nucleotide in the initiating-nucleotide binding site of the polymerase.

Non-Core Sequences Affect Promoter Strength In Vivo

As discussed above, homology to the consensus core promoter elements is not always sufficient to determine promoter strength levels in vivo. Sequences upstream, downstream, and between the core promoter elements can also have significant effects. We consider these separately because they probably are important only for a subset of promoters. Understanding how these sequences affect promoter strength will provide important information regarding the overall process of transcription initiation.

Extended –10 Promoters. Some σ^{70} promoters such as *gal* P1 contain no obvious –35 region sequence. In these cases the sequence immediately upstream of the –10 region has been found through mutation studies to affect promoter activity (see references 119 and 236 and references therein). The optimal sequence in these cases appears to be TnTGn followed by the –10 region. It is not known whether an extended –10 region can influence promoter strength in the presence of a consensus –35 region sequence, nor what step(s) in the transcription process is influenced by this sequence. Extended –10 promoters are efficiently recognized in vitro by $E\sigma^{70}$ containing a deletion of the σ subunit region 4 (the recognition element for the –35 region recognition element) (119).

UP Element. Replacement of the upstream region (–40 to –60) (UP element) of the *rrnB* P1 promoter with vector or random sequences reduced the strength of the promoter in vivo to 1/30 its wild-type amount, indicating that the UP element contributes greatly to the strength of the promoter (126, 160, 169). Replacement of the naturally occurring sequence of the *lac* or *lac*UV5 promoter (–40 to –60) with the UP elements of the *rrnB* P1, *rrnB* P2, *mer*, or RNA II promoters resulted in increases in promoter activity, suggesting that the UP element of *rrnB* P1 has a unique function which is separable from the function of the

core promoter sequence (169, 183). UP elements are A-T rich, but the critical sequences or structures have not yet been determined.

How do UP elements increase the velocity of transcription initiation? Replacement of the natural UP element at the *rrnB* P1 promoter with vector sequences results in a reduction in k_a and possibly in k_i (126, 169). The UP element may therefore function mechanistically by stabilizing the initial closed complex such that, upon removal of the UP element, the rate constant for the dissociation of the initial closed complex (k_{-1} or k_{-1}) becomes significant in the kinetics of open complex formation (169). Since the α subunit and α-CTD bind this region of DNA (16, 183) and since the amount of transcription from the *rrnB* P1 promoters is reduced in carboxy-terminal α mutants, it is likely that the binding of the α subunit is responsible for increasing K_1, the equilibrium constant for formation of RP_{c1}.

Other Upstream Sequence Effects. The effects of curved or bent (e.g., A-tract) sequences located upstream of the –35 region on the kinetics of transcription initiation have been studied at many promoters (12, 20, 64, 67, 122, 126, 169). Replacement of curved upstream sequences by random or vector sequences reduces k_a and/or k_i, suggesting that initial binding and/or isomerization steps may be influenced by the upstream sequence (reviewed in reference 167). Replacement of the upstream region of a promoter with sequences of other promoters or A tracts influenced both the in vivo promoter strength and the kinetics of forming a stable (presumably open) complex (27, 58, 105). Generally, insertion of A tracts or sequences from strong phage promoters upstream of the –35 sequence increases the rate at which RNA polymerase forms the first competitor-resistant complex. However, an imperfect correlation was observed between this rate constant and in vivo promoter strength, suggesting that steps after open complex formation may be rate limiting in vivo and that the presence of A-tracts or other specific upstream sequences may reduce the rates of these subsequent steps (58). It is not clear which of the observed effects of changes in upstream sequence are related to UP element function.

At some promoters, the base pair immediately adjacent to the –35 region has a significant effect on promoter strength. For example, substitution of the –37 C/G base pair in the *lac* promoter for a T/A base pair (changing the –35 region from CTTTACA to TTTTACA) reduces expression by 10- to 20-fold, possibly by introducing a stronger "A-tract bend." Other substitutions at this site also have deleterious effects (124; W. S. Reznikoff, unpublished data). Similar behavior was observed with the *rrnB* P1 promoter (102). In this case the mutant effect was attributed to the introduction of an unfavorable DNA bend immediately adjacent to the –35 region (64). These observations are important because they suggest that effects of changes in DNA sequence may in some cases result from changes in DNA structure instead of only from changes in DNA functional groups which interact with polymerase.

Downstream Elements. The region immediately downstream of the transcription start site influences promoter strength (27). A series of downstream regions (DSR; also called initial transcribed regions [ITR]) were introduced to change the sequence from position +1 to +20 of a consensus core promoter and of the N25 phage promoter (105). Different sequences affected in vivo promoter strength by more than 10-fold, indicating that the DSR can be important in determining promoter strength. However, in measurements performed at a single polymerase concentration, no effect of these changes on the rate of formation of RP_{o2} was observed. It appears that the DSR does not play a role in open complex formation at these promoters, but rather affects steps involving ternary RP_{init} complexes. Investigation of the in vitro abortive initiation recycling properties of those P_{N25} constructs shown previously to exhibit DSR effects in vivo (Hsu and Chamberlin, in preparation) showed that downstream sequence affects the distribution and amounts of abortive products generated. DSR yielding a high level of in vivo expression yielded a low level of abortive initiation in vitro, and vice versa.

Finding Promoters in DNA Sequence Information

Soon, the entire DNA sequence of the genomes of *E. coli* and other related organisms will be known. Do we know enough about promoter structure to identify promoters in regions of the sequence that lack genetic and/or biochemical analysis? A matrix search protocol based on quantitative homology scores derived from the analysis of P22 P_{ant} promoter mutant data (155) was used to search for promoters (41). The test derived a numerical score for core promoter elements that was a sum of scores at each of the 12 positions in the two hexamers plus a score for the length of the spacer between the –35 and –10 region. Scores were assigned to individual bases at each position based on their effect on P22 P_{ant} promoter activity. Possible promoters were identified upstream of open reading frames. However, this type of analysis is certain to suffer from two types of limitations. First, it does not take into consideration non-core sequences that affect promoter strength. Second, the scoring system assumes that RNA polymerase recognizes the bases at different positions independently; e.g., the scoring system does not include context-dependent effects. The obvious challenge is to develop a predictive understanding of the contribution of non-core sequences and to compensate for context-dependent effects.

EFFECTS OF SOME EXTRINSIC VARIABLES ON TRANSCRIPTION INITIATION

Supercoiling

Negative supercoiling of closed circular or otherwise topologically restrained DNA provides a large thermodynamic driving force for processes that unwind the DNA double helix (e.g., intercalation of dyes; denaturation) or otherwise reduce the amount of negative supercoiling (e.g., wrapping the DNA, as on a nucleosome). Since local denaturation (opening) and bending or wrapping of the promoter DNA appear to be integral parts of the mechanism of open complex promotion, it is reasonable to expect that both the thermodynamics and at least some aspects of the kinetics of open complex formation/dissociation should be strongly affected by the extent of supercoiling of a closed circular promoter DNA. The total unwinding accompanying open complex formation at two strong TAC promoter constructs on supercoiled minicircles was found to be 1.0 to 1.2 turns (216); previous determinations for other promoters ranged from 0.7 to 1.7 superhelical turns. These results indicate that the unwinding that results from opening of 1 to 1.5 turns of B-DNA in forming RP_{o2} is not topologically restrained by the contacts of RNA polymerase with DNA on both sides of the open promoter region, and that DNA opening, as expected, and not wrapping or

twisting of helical DNA, makes the dominant contribution to superhelical unwinding.

The effects of changes in supercoiling on initiation velocity (promoter strength) in vivo and on the kinetics of open complex formation in vitro are sometimes large, often complicated, and not well understood. Effects in vivo are promoter specific; promoter strength can be increased, reduced, or unaffected by changes in supercoiling (see reference 53). In vitro, promoter-specific increases in the association rate constant k_a of up to 160-fold (*lac*UV5 [24]), ~50-fold (*lac*P^s [207]), ~20-fold (TAC 16 [216]), and ~4-fold (TAC 17 [216]) have been reported. These association rate constants typically exhibit biphasic responses (first increasing, then leveling off or decreasing) to increases in negative superhelical density. Maximal rates of association on different promoters (as cited above) appear to occur at different superhelix densities, which may explain the different directions of the effects of changes in supercoiling on promoter strength at different promoters in vivo (see reference 216 and references therein).

Analysis of the effects of changes in superhelix density (in the range 0 to –0.04, which is roughly physiological) on k_a for the TAC 16 and TAC 17 promoters led to the proposal that approximately 20% (TAC 17) to 40% (TAC 16) of the total unwinding at these promoters occurs in or prior to the rate-limiting step in open complex formation (216). For example, if $k_a = K_1k_2$ (equation 13), then this unwinding (equivalent to opening 2 to 4 bp) occurs prior to the transition state separating RP_{c2} from RP_{c1}. If this unwinding occurs relatively early in the opening mechanism, before the formation of open complexes (RP_{o1} or RP_{o2}), it may result from wrapping or other events unrelated to DNA opening, potentially including any untwisting involved in alignment of the –10 and –35 hexamers (see reference 46).

Examples of Repression or Activation of Transcription Initiation by Proteins

In this section, some quantitative studies of the mechanisms by which repressors and activators affect transcription are reviewed. Other aspects of these topics are discussed elsewhere in this volume (see chapters 82 and 83 of this book). Both repressor and activator proteins exert large effects on transcription initiation and could, in principle, affect any of the steps in the mechanism of initiation. Indeed, different repressor and activator proteins exert their effects on initial binding and/or isomerization and/or initial transcript production.

Repression by Competition. The DNA-binding sites (called operators) of many repressors overlap the promoters which they regulate. The simplest mechanism of repression in these cases is that repressor and polymerase cannot simultaneously occupy these overlapping sequences and hence compete (either thermodynamically or kinetically) for the promoter-operator region. A bound repressor reduces the available promoter concentration [*P*], thereby reducing the rate of formation of the initial closed complex.

lexA. The *lexA* operator site (Fig. 11A) overlaps the –35 region of the *uvrA* promoter. An elegant quantitative study was used to determine the in vitro mechanism of repression of the *uvrA* promoter by binding of *lexA* or the weaker binding *lexA* N-terminal DNA-binding fragment to this operator (13). Measurements of the rate of formation of open complexes at a variety of RNA polymerase and *lexA* concentrations were used to show that *lexA* repressor acts as a thermodynamic competitor which reduces the available promoter concentration and thereby reduces the rate of initial binding of RNA polymerase.

lacR. The primary *lac* operator site, *lac* O_1 (Fig. 11B), overlaps the start site (+1) of the *lac* promoter. Does a direct competitive binding mechanism describe *lac* repression? Regulation of the *lac* operon by *lac* repressor (L) has been studied both in vivo and in vitro for over 30 years. In vivo, regulation of expression of the *lac* genes by *lac* repressor appears to be under thermodynamic control, where the fraction of maximum expression (here designated *F*) is simply the fraction of promoter-operator sites (*P*) which are free (i.e., not occluded by bound repressor), and is given by an equilibrium binding isotherm of the form:

$$F \equiv \frac{[P]}{[P]_T} = (1 + K_{PL}a_L)^{-1} \qquad (20)$$

where K_{PL} is the repressor-operator equilibrium binding constant and a_L is the thermodynamic activity of repressor (123, 184, 228, 230). Though equation 20 is consistent with a competitive binding model for *lac* repression, an inverse dependence of F on a_L is also at least qualitatively consistent with other mechanisms of repression proposed on the basis of in vitro experiments, as discussed below.

In vitro, the primary effect of *lac* repressor on transcription appears to be on initiation (36), although it may also affect elongation (48, 193). Footprinting, gel mobility shift, cross-linking, transcription, and abortive initiation assays provided support for a variety of models in which repressor blocked initial binding (121, 134, 136), affected conformational changes of RNA polymerase and/or promoter DNA (21, 36, 191, 213), and/or interfered with nucleotide incorporation steps (125) in transcription initiation.

The mechanism of *lac* repression of the *lac*UV5 promoter was recently investigated in vitro for a choice of the repressor-operator binding constants and ranges of concentrations (thermodynamic activities) of *lac* repressor (L) and RNA polymerase (R) which appear to be relevant in vivo (189). Effects of [L] on the extent of formation and the kinetics of association and dissociation of abortively initiating open complexes (RP_{init}) were quantified using fluorescence-detected abortive initiation and $KMnO_4$ chemical probing. For the chosen solution conditions, both the observed velocity of abortive RNA oligomer synthesis and the $KMnO_4$ reactivities of bases in the open region are functions of [L] and [R], demonstrating that formation of both RP_{init} and the repressor-operator complex (PL) are reversible processes under these conditions (189). Thus the use of a relaxation-to-equilibrium analysis (equations 6 through 8) is required to interpret the kinetics. Addition of L to preformed abortively initiating complexes (RP_{init}) results in a loss of $KMnO_4$ reactivity of the bases in the open region and a reduction in the velocity of abortive product synthesis, indicating a reduction in [RP_{init}] with increasing [L]. The kinetics of dissociation of RP_{init} are first order and independent of $[L]_T$, with a dissociation rate constant which agrees within error with that obtained using heparin as a competitor. This result excludes any significant role of a product ternary complex ($RP_{init}L$) (189).

Quantitative analysis of all the kinetic data supports a direct equilibrium competition mechanism in which reversibly bound repressor reduces the available promoter concentration [P] and therefore reduces the rate and extent of formation of RP_{c1}. The

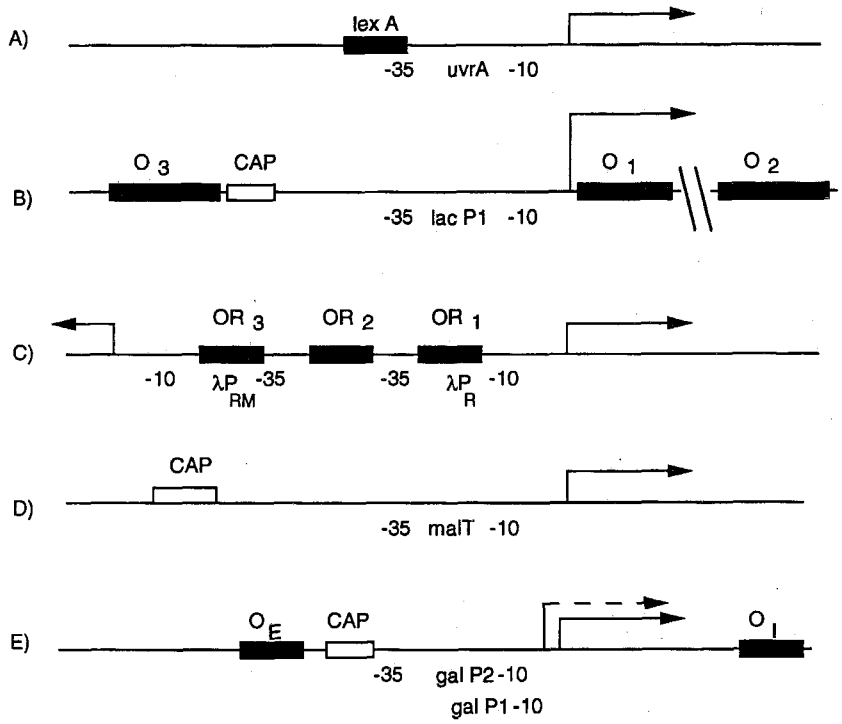

FIGURE 11 Schematic of the regulatory regions of several operons. (A) In the *uvrA* promoter, the binding site for *lexA* overlaps the –35 sequence of the core promoter. (B) In the *lac* operon, the primary binding site for the *lac* repressor (the O_1 operator) overlaps the start site of transcription from the primary P_1 (or mutant UV5) promoter. A binding site for CAP is centered 61.5 bp upstream of this start site. A second promoter (P_2) overlaps P_1 and the CAP site. Two weaker *lac*-repressor-binding (secondary or pseudo-operator) sequences are centered 93 bp upstream (O_3) and 401 bp downstream (O_2) of the center of the O_1 operator sequence. (C) In the λP_R-λP_{RM} regulatory region, the start sites of the two divergent promoters are separated by 82 bp. Three binding sites for λcI repressor are located in this region, centered at 17, 41, and 64 bp upstream of the λP_R start site. (D) At the *malT* promoter, a single CAP site (overlapping a divergent P_X promoter) is centered 70 bp upstream of the transcription start site. (E) In the *gal* operon, two tandem promoters (*gal* P1 and *gal* P2) with start sites separated by 5 bp are regulated by CAP, which binds to a site centered 41.5 bp upstream of the *gal* P1 promoter, and by *galR*, which binds to two operator sequences (O_E and O_I) centered 59.5 bp upstream and 53 bp downstream of the *gal* P1 start site.

agreement between dissociation rate constants of RP_{init} when challenged with either *lac* repressor or heparin, and the dependences on [L] and [R] of abortive synthesis velocities at binding equilibrium and of relaxation rate constants for reversible formation of RP_{init} from PL, all provide evidence for this equilibrium competition mechanism. Without taking the reversibility of open complex formation into account in the analysis of the kinetic data, an incorrect conclusion (i.e., that *lac* repressor might inhibit transcription initiation by affecting an isomerization step) would have been obtained. It therefore appears that, for physiologically relevant choices of binding parameters, *lac* repressor inhibits formation of RP_{init} and hence the observed rate of abortive product synthesis by reducing the equilibrium extent of formation of the first closed complex (RP_{cl}), without affecting either the nature of RP_{init} or steps in formation of RP_{init} from RP_{cl} (189).

λcI Repressor. In the O_R control region of phage lambda, both a direct thermodynamic competitive mechanism of repression and a more complicated indirect mechanism of repression are observed. This regulatory region contains two divergent promoters (λP_R, λP_{RM}) separated by three binding sites for λcI repressor (see Fig. 11C). λcI repressor binds most tightly to the O_R1 operator, located adjacent to the λP_R –35 region. λcI bound to O_R1 effectively represses λP_R in the absence of the other operator sites. λcI binds to O_R2 with second highest affinity. When λcI is bound to both O_R1 and O_R2, λP_R is repressed and λP_{RM} is activated. Binding of λcI reduced k_a but did not affect k_i at the λP_R promoter, indicating that λcI competes at the thermodynamic level with initial binding of RNA polymerase (84).

A second level of modulation of expression, observed in the absence of λcI repressor, results from the proximity of the two promoters, whose divergent start sites are located 82 bp apart. In excess RNA polymerase, the rate of transcription initiation at λP_R is reduced by formation of an open complex at λP_{RM} and vice versa (87). Kinetic studies to determine the effects of changes in the sequence of one promoter on the kinetics of transcription initiation at the other promoter were performed using series of constructs containing wild-type and variant λP_R and λP_{RM} promoters (61, 62, 87, 88, 234). Open complex formation at both λP_{RM} and λP_R (and their variants) is apparently irreversible under the conditions of these studies. RNA polymerase can occupy both promoters simultaneously, so repression by polymerase probably is not caused by competition in the initial binding steps, but rather by an effect of polymerase bound at the one promoter on isomerization steps at the other promoter. Kinetic studies of open complex formation at each

promoter support this proposal by demonstrating that, in all cases studied, k_i is affected by changes in the sequence of the other promoter. Presumably a direct interaction between the two promoter-bound RNA polymerases causes these effects, though other explanations have not been excluded.

Activation by CAP Protein: Evidence for a Direct Protein-Protein Interaction. Catabolite gene activator protein (CAP) is the most extensively studied transcriptional activator (56, 111, 176). CAP is a dimeric protein which, when complexed with cyclic AMP, binds to specific DNA sites upstream of the core promoter sequence and increases promoter strength (see Fig. 11B, D, and E). Two classes of CAP-responsive regulatory systems are known. In class 1 systems, CAP binds to a site centered at position –61.5 relative to the promoter start site. At the *lac* P_1 promoter, CAP increases promoter strength ~50-fold upon binding at –61.5. In class 2 systems (e.g., the *gal* promoter) the CAP-binding site is centered at –41.5. The importance of CAP-binding site location, rather than differences in promoter structure, in distinguishing these two classes of CAP-responsive regulatory systems was confirmed by studies of a synthetic series of CAP-regulated constructs in which the basic promoter sequence was conserved but the CAP-binding site location was varied (68). These differences in binding site location have profound effects on the molecular nature of the activation mechanisms, as described below.

Current evidence, primarily derived from studies of class 1 activation systems (such as the *lac* promoter), suggests that CAP functions by making a protein-protein contact with RNA polymerase. The evidence in support of this model includes the following. (i) CAP and RNA polymerase each have cooperative effects on binding of the other to *lac* regulatory region DNA (111, 174, 212). (ii) CAP binds to RNA polymerase in the absence of core promoter DNA (although this binding appears to be greatly facilitated by DNA fragments containing the CAP-binding site) (90). (iii) CAP activation of transcription initiation is abolished by mutations in a unique loop (containing residue 158) between two β-strands of CAP. These mutants have normal DNA binding and bending activities and they specifically block CAP-RNA polymerase cooperativity in DNA binding and CAP-RNA polymerase binding in solution (10, 59, 166, 237). In addition to this contact location, class 2 promoters demonstrate another possible contact between CAP and RNA polymerase (5). (iv) CAP activation of *lac* transcription initiation is abolished by carboxy-terminal α mutants (95, 112, 223, 242). These mutant RNA polymerases function normally in initiating transcription at many nonregulated promoters. There is a slight ambiguity in this particular point, since the recognition of UP element DNA may involve a nearby portion of the α protein. (v) The proposed CAP-α interaction has been recently confirmed through protein-protein photo-cross-linking experiments (37).

Although the evidence is consistent with a model in which CAP activation occurs by direct interaction with RNA polymerase, other mechanisms have been proposed which may function in addition to the direct contact mechanism. For example, CAP bends DNA (192, 235). Although bending alone is not sufficient to activate *lac* (based upon the properties of the CAP mutants near residue 158), it may play an ancillary role. CAP may also act to prevent RNA polymerase binding to nonproductive sites (136, 168), but genetic evidence suggests that this indirect activation mechanism is not likely to play a major role (52).

The effects of CAP on the kinetics of open complex formation have been studied for both class 1 and class 2 promoters. Studies with the wild-type *lac*P_1 promoter (class 1; CAP site centered at ~ –61.5) demonstrated that the presence of CAP served to increase k_a, while k_i was unaffected (135). Interpretation of these association experiments is complicated by the presence of the overlapping P_2 promoter and the fact that the binding of CAP activates RNA polymerase binding to *lac*P_1 both directly and indirectly (by occluding binding of RNA polymerase at the *lac*P_2 promoter). The dissociation rate constant k_d for RP_{o2} at *lac*P_1 was reduced in the presence of CAP (135).

To avoid the complications of a second overlapping promoter site, the effects of CAP on the kinetics of open complex formation were investigated using a synthetic promoter system in which the distance separating the CAP- and RNA polymerase-binding sites was systematically changed (68). As with the *lac*P_1 promoter, association experiments with a synthetic class 1 promoter demonstrated a significant increase in k_a in the presence of CAP. A small increase in k_i was also observed. The dissociation rate constant k_d was not affected by the presence of CAP. For a synthetic class 2 promoter (CAP site centered at ~ –41.5), CAP significantly increased both k_a and k_i and reduced k_d slightly.

On the basis of these limited studies, it appears possible that the mechanism of CAP activation is different for class 1 and class 2 promoters. For both promoter classes, CAP caused large increases in k_a and only small decreases (*lac*P_1, synthetic class 1) or no effect (synthetic class 2) in k_d. However, CAP binding at –41.5 of the synthetic class 2 promoter dramatically increased k_i, while only a small (or no) increase in k_i was observed for CAP binding at –61.5 in the class 1 promoters. This suggests that for class 1 promoters CAP primarily increases the extent of initial binding (RP_{c1}) by increasing [P] (at *lac*) and/or by decreasing k_{-1}, while the main effect of CAP on open complex formation at class 2 promoters may be to increase k_2 for the $RP_{c1} \rightarrow RP_{c2}$ isomerization step.

While CAP has been implicated in contacting the α subunit of RNA polymerase, the λcI repressor, which activates transcription from the λP_{RM} promoter, apparently contacts the σ^{70} subunit (as described above) (117, 130). Repressor increases the rate of open complex formation at λP_{RM} by increasing both k_a and k_i (85, 87, 234). This results from a direct effect of λcI on open complex formation at λP_{RM} and from an indirect effect resulting from the repression of λP_R by λcI. (As described above, polymerase bound at λP_R represses transcription initiation at λP_{RM}.)

The *malT* promoter is positively regulated by CAP binding to a site centered 70.5 bp upstream of the transcription start site. The mechanism of activation at this CAP site does not appear to correspond to mechanisms of activation at the previously described class 1 or class 2 activated promoters. Effects of CAP on the observed rate of open complex formation at the *malT* promoter were very small at all RNA polymerase concentrations tested (150). Surprisingly, CAP increased the affinity of the open complex for binding to UTP, resulting in production of a higher percentage of productive (full-length) transcripts. This result suggests that in this system CAP affects steps between open complex formation and promoter clearance.

In summary, repressors and activators exert large effects on the individual steps in transcription initiation. Various regulatory proteins affect the kinetics of binding and dissociation of RP_{c1} (13, 68, 85, 135, 189), the kinetics of subsequent isomeriza-

tion steps (68, 85, 87, 234), and the kinetics of steps subsequent to open complex formation but prior to promoter clearance (150).

SUMMARY

In this article, we have discussed the steps of transcription initiation by $E\sigma^{70}$ RNA polymerase in terms of the thermodynamics and kinetics of site-specific binding and of coupled conformational changes, as well as the subsequent enzymology of the NTP-driven processes of chain initiation and downstream movement of polymerase. As for other specific DNA-binding interactions, not only sequence and context but also environmental variables are crucial determinants of the rates and equilibria of the steps involved in forming RP_{o2} and RP_{init}. These strong dependences on sequence, context, and environment of noncovalent interactions between polymerase and promoter DNA allow the cell to regulate the kinetics of initiation in a promoter-specific manner at each forward and reverse step of the process, and they provide the flexibility to adapt to new conditions characteristic of living systems.

Many different levels of control of gene expression during transcription initiation are known, and doubtless many more remain to be identified. Intrinsic regulation is governed by the effects of sequence of the core promoter and its flanking regions on the steps in transcription. Additionally, different sigma factors recognize different classes of core promoter sequences, allowing expression of different families of genes to occur as a response to the condition of the cell. Binding sites for regulatory proteins in the vicinity of the promoter allow repressors and activators to have direct effects (by competition or protein-protein interaction) on the kinetics and equilibria of individual steps in initiation. Looping and the existence of multiple sites for different regulatory proteins allow for a spectrum of mechanisms of regulation (1, 11, 74, 141, 190).

Changes in the extrinsic (regulatory protein concentration and binding strength, supercoiling) and intrinsic (sequence) variables can therefore change the velocity of transcript initiation (V_{TC}) by many orders of magnitude. The assignment of these large effects to individual kinetic steps of the mechanism of initiation is well under way. Relationships between promoter structure and function are being characterized by these kinetic and mechanistic studies.

ACKNOWLEDGMENTS

We thank Peter Schlax for his contributions to the article; Dick Burgess, Pieter deHaseth, Laura Heisler, Ruth Spolar, Rick Gourse, and Carol Gross and her laboratory for discussions and for comments on various drafts; and Sheila Aiello for preparation of the manuscript. Work from the authors' laboratories was supported by grants from the National Institutes of Health (GM 23467 to M.T.R.) and the National Science Foundation (DMB 9020517 to W.S.R.).

LITERATURE CITED

1. **Adhya, S.** 1989. Multipartite genetic control elements: communication by DNA loop. *Annu. Rev. Genet.* **23**:227–250.
2. **Aiyar, S. E., J. D. Helmann, and P. L. deHaseth.** 1994. A mismatch bubble in double-stranded DNA suffices to direct precise transcription initiation by *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **18**:13179–13184.
3. **Aiyar, S. E., Y.-L. Juang, J. D. Helmann, and P. L. deHaseth.** 1994. Mutations in sigma factor that affect the temperature dependence of transcription from a promoter, but not from a mismatch bubble in double stranded DNA. *Biochemistry* **33**:11501–11506.
4. **Allison, L. A., M. Moyle, M. Shales, and C. J. Ingles.** 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**:599–610.
5. **Attey, A., T. Belyaeva, N. Savery, J. Hoggett, N. Fujita, A. Ishihama, and S. Busby.** 1994. Interactions between the cyclic AMP receptor protein and the alpha subunit of RNA polymerase at the *Escherichia coli* galactose operon P1 promoter. *Nucleic Acids Res.* **22**:4375–4380.
6. **Auble, D. T., T. L. Allen, and P. L. deHaseth.** 1987. Promoter recognition by *Escherichia coli* RNA polymerase. Effects of substitutions in the spacer DNA separating the –10 and –35 regions. *J. Biol. Chem.* **261**:11202–11206.
7. **Auble, D. T., and P. L. deHaseth.** 1988. Promoter recognition by *Escherichia coli* RNA polymerase. Influence of DNA structure in the spacer separating the –10 and –35 regions. *J. Mol. Biol.* **202**:471–482.
8. **Ayers, D. G., D. T. Auble, and P. L. deHaseth.** 1989. Promoter recognition by *Escherichia coli* RNA polymerase. Role of the spacer DNA in functional complex formation. *J. Mol. Biol.* **207**:749–756.
9. **Beese, L., V. Derbyshire, and T. Steitz.** 1993. Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**:352–355.
10. **Bell, A., K. Gaston, R. Williams, K. Chapman, A. Kolb, H. Buc, S. Minchin, J. Williams, and S. Busby.** 1990. Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res.* **18**:7243–7250.
11. **Bellomy, G. R., and M. T. Record, Jr.** 1990. Stable DNA loops *in vivo* and *in vitro*: roles in gene regulation at a distance and in biophysical characterization of DNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**:81–128.
12. **Bertrand-Burggraf, E., J. Dunand, R. P. P. Fuchs, and J. F. Lefèvre.** 1990. Kinetic studies of the modulation of ada promoter activity by upstream elements. *EMBO J.* **9**:2265–2271.
13. **Bertrand-Burggraf, E., S. Hurstel, M. Daune, and M. Schnarr.** 1987. Promoter properties and negative regulation of the *uvrA* gene by the LexA repressor and its amino-terminal DNA binding domain. *J. Mol. Biol.* **193**:293–302.
14. **Bertrand-Burggraf, E., J. F. Lefèvre, and M. Daune.** 1984. A new experimental approach for studying the association between RNA polymerase and the tet promoter of pBR322. *Nucleic Acids Res.* **12**:1697–1706.
15. **Bjornson, K. P., M. Amaratunga, K. J. M. Moore, and T. M. Lohman.** 1994. Single-turnover kinetics of helicase-catalyzed DNA unwinding monitored continuously by fluorescence energy transfer. *Biochemistry* **33**:14306–14316.
16. **Blatter, E. E., W. Ross, H. Tang, R. L. Gourse, and R. H. Ebright.** 1994. Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**:889–896.
17. **Bokal, A. J., IV, W. Ross, and R. L. Gourse.** 1995. The transcriptional activator protein FIS: DNA interactions and cooperative interactions with RNA polymerase at the *Escherichia coli rrnB* P1 promoter. *J. Mol. Biol.* **245**:197–207.
18. **Bond, J. P., C. F. Anderson, and M. T. Record, Jr.** 1994. Conformational transitions of duplex and triplex nucleic acid helices: thermodynamic analysis of effects of salt concentration on stability using preferential interaction coefficients. *Biophys. J.* **67**:825–836.
19. **Borukhov, S., V. Sagitov, C. A. Josaitis, R. L. Gourse, and A. Goldfarb.** 1993. Two modes of transcription initiation *in vitro* at the *rrnB* P1 promoter of *Escherichia coli*. *J. Biol. Chem.* **268**:23477–23482.
20. **Bracco, L., D. Kotlarz, A. Kolb, S. Diekman, and H. Buc.** 1989. Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *EMBO J.* **8**:4289–4296.
21. **Brodolin, K. L., V. M. Studitsky, and A. D. Mirzabekov.** 1993. Conformational changes in *E. coli* RNA polymerase during promoter recognition. *Nucleic Acids Res.* **21**:5748–5753.
22. **Brunner, M., and H. Bujard.** 1987. Promoter recognition and promoter strength in the *Escherichia coli* system. *EMBO J.* **6**:3139–3144.
23. **Buc, H.** 1987. Initiation of prokaryotic transcription—kinetic and structural approaches, p. 186–195. *In* F. Eckstein and D. M. J. Lilley (ed.), *Nucleic Acids and Molecular Biology*. Springer-Verlag, Berlin.
24. **Buc, H., and W. R. McClure.** 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.
25. **Buckle, M., A. Fritsch, P. Roux, J. Geiselmann, and H. Buc.** 199J. Kinetic studies on promoter-RNA polymerase complexes. *Methods Enzymol.* **208**:236–257.
26. **Buckle, M., J. Geiselmann, A. Kolb, and H. Buc.** 1991. Protein-DNA crosslinking at the *lac* promoter. *Nucleic Acids Res.* **19**:833–840.
27. **Bujard, H., M. Brunner, U. Deuschle, W. Kammerer, and R. Knaus.** 1987. Structure function relationship of *Escherichia coli* promoters, p. 95–103. *In* W. S. Reznikoff, R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr., and M. P. Wickens, (ed.), *RNA Polymerase and the Regulation of Transcription*. Elsevier, New York.
28. **Burgess, R., B. Erickson, D. Gentry, M. Gribskov, D. Hager, S. Lesley, M. Strickland, and N. Thompson.** 1987. Bacterial RNA polymerase subunits and genes, p. 3–16. *In* Reznikoff, W. S., R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record Jr., and M. P. Wickens (ed.), *RNA Polymerase and the Regulation of Transcription*. Elsevier, New York.
29. **Burgess, R. R., A. A. Travers, J. J. Dunn, and E. K. F. Bautz.** 1969. Factor stimulating transcription by RNA polymerase. *Nature* (London) **221**:43–46.
30. **Carpousis, A. J., and J. D. Gralla.** 1980. Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the *lac*UV5 promoter. *Biochemistry* **19**:3245–3253.
31. **Carpousis, A. J., and J. D. Gralla.** 1985. Interaction of RNA polymerase with *lac*UV5 promoter DNA during mRNA initiation and elongation. *J. Mol. Biol.* **183**:165–177.
32. **Carpousis, A. J., J. E. Stefano, and J. D. Gralla.** 1982. 5′ Nucleotide heterogeneity and altered initiation of transcription at mutant *lac* promoters. *J. Mol. Biol.* **157**:619–633.

33. **Cayley, S., B. A. Lewis, H. J. Guttman, and M. T. Record, Jr.** 1991. Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity: implications for protein-DNA interactions *in vivo*. *J. Mol. Biol.* **222**:281–300.
34. **Chamberlin, M. J.** 1974. The selectivity of transcription. *Annu. Rev. Biochem.* **43**:721–766.
35. **Chamberlin, M. J.** 1976. Interaction of RNA polymerase with the DNA template, p. 159–191. *In* R. Losick and M. Chamberlin (ed.), *RNA Polymerase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Chen, B., B. de Crombrugghe, W. B. Anderson, M. E. Gottesman, I. Pastan, and R. L. Perlman.** 1971. On the mechanism of action of *lac* repressor. *Nature* (London) *New Biol.* **233**:67–70.
37. **Chen, Y., Y. W. Ebright, and R. H. Ebright.** 1994. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. *Science* **265**:90–92.
38. **Chenchik, A., R. Beabealshvilli, and A. D. Mirzabekov.** 1981. Topography of interaction of *Escherichia coli* RNA polymerase subunits with *lac*UV5 promoter. *FEBS Lett.* **128**:46–50.
39. **Chenchik, A. A., R. Beabealshvilli, A. D. Mirzabekov, and V. V. Shik.** 1982. Contact between subunits of *Escherichia coli* RNA polymerase and the nucleotides of the *lac* UV5 promoter. *J. Mol. Biol.* **16**:34–46.
40. **Cowing, D. W., J. Mecsas, M. T. Record, Jr., and C. A. Gross.** 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.

40a. **Craig, M. L., W. C. Suh, and M. T. Record, Jr.** 1995. HO · and DNase I probing of $E\sigma^{70}$ RNA polymerase –λP_E promoter open complexes: Mg^{2+} binding and its structural consquences at the transcription start site. *Biochemistry* **34**:15624–15632.

41. **Daniels, D., G. Punkett III, V. Burland, and F. R. Blattner.** 1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. *Science* **257**:771–778.
42. **Daniels, D., P. Zuber, and R. Losick.** 1990. Two amino acids in an RNA polymerase σ factor involved in the recognition of adjacent base pairs in the –10 region of a cognate promoter. *Proc. Natl. Acad. Sci. USA* **87**:8075–8079.
43. **Darst, S. A., A. M. Edwards, E. W. Kubalek, and R. D. Kornberg.** 1991. Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell* **66**:121–128.
44. **Darst, S. A., E. W. Kubalek, and R. D. Kornberg.** 1989. Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature* (London) **340**:730–732.
45. **Daube, S. S., and P. H. von Hippel.** 1992. Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science* **258**:1320–1324.
46. **deHaseth, P. L., and J. D. Helmann.** 1995. Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol. Microbiol.* **16**:817–824.
47. **deHaseth, P. L., T. M. Lohman, R. R. Burgess, and M. T. Record, Jr.** 1978. Nonspecific interactions of *Escherichia coli* RNA polymerase with native and denatured DNA: differences in the binding behavior of core and holoenzyme. *Biochemistry* **17**:1612–1622.
48. **Deuschle, U., R. Gentz, and H. Bujard.** 1986. *lac* repressor blocks transcribing RNA polymerase and terminates transcription. *Proc. Natl. Acad. Sci. USA* **83**:4134–4137.
49. **Deuschle, U., W. Kammerer, R. Gentz, and H. Bujard.** 1986. Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *EMBO J.* **5**:2987–2994.
50. **Dombroski, A. D., W. A. Walter, and C. A. Gross.** 1993. Amino-terminal amino acids modulate σ factor DNA-binding activity. *Genes Dev.* **7**:2446–2455.
51. **Dombroski, A. J., W. A. Walter, M. T. Record, Jr., D. A. Siegele, and C. A. Gross.** 1992. Polypeptides containing highly conserved regions of transcription initiation factor σ^{70} exhibit specificity of binding to promoter DNA. *Cell* **70**:501–512.
52. **Donnelly, C., and W. S. Reznikoff.** 1987. Mutations in the *lac*P2 promoter. *J. Bacteriol.* **169**:1812–1817.
53. **Drlica, K.** 1984. Biology of bacterial deoxyribonucleic acid topoisomerases. *Microbiol. Rev.* **48**:271–289.
54. **Duval-Valentin, G., and R. Ehrlich.** 1986. Interaction between *E. coli* RNA polymerase and the *tetR* promoter from pSC101: homologies and differences with other *E. coli* promoter systems from close contact point studies. *Nucleic Acids Res.* **14**:1967–1983.
55. **Duval-Valentin, G., and R. Ehrlich.** 1987. Dynamic and structural characterization of multiple steps during complex formation between *E. coli* RNA polymerase and the *tetR* promoter from pSC101. *Nucleic Acids Res.* **15**:575–594.
56. **Ebright, R.** 1993. Transcription activation at class I CAP-dependent promoters. *Mol. Microbiol.* **8**:797–802.
57. **Ellinger, T., D. Behnke, H. Bujard, and J. Gralla.** 1994. Stalling of *Escherichia coli* RNA polymerase in the +6 to +12 region *in vivo* is associated with tight binding to consensus promoter elements. *J. Mol. Biol.* **239**:455–465.
58. **Ellinger, T., D. Behnke, R. Knaus, H. Bujard, and J. D. Gralla.** 1994. Context-dependent effects of upstream A-tracts: stimulation or inhibition of *Escherichia coli* promoter function. *J. Mol. Biol.* **239**:466–475.
59. **Eschenlauer, A., and W. S. Reznikoff.** 1991. *Escherichia coli* catabolite gene activator protein mutants defective in positive control of *lac* operon transcription. *J. Bacteriol.* **173**:5024–5029.
60. **Farrow, L. A., and D. Edelson.** 1974. The steady-state approximation: fact or fiction? *Int. J. Chem. Kinetics* **6**:787–800.
61. **Fong, R. S.-C., S. Woody, and G. N. Gussin.** 1993. Modulation of P_{RM} activity by the lambda Pr promoter both in the presence and absence of repressor. *J. Mol. Biol.* **232**:792–804.
62. **Fong, R. S.-C., S. Woody, and G. N. Gussin.** 1994. Direct and indirect effects of mutation in λP_{RM} on open complex formation at the divergent P_R promoter. *J. Mol. Biol.* **240**:119–126.
63. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505–6525.
64. **Gaal, T., L. Rao, S. T. Estrem, J. Yang, R. M. Wartell, and R. L. Gourse.** 1994. Localization of the intrinsically bent DNA region upstream of the *E. coli* rrnB P1 promoter. *Nucleic Acids Res.* **22**:2344–2350.
65. **Gardella, T., H. Moyle, and M. M. Susskind.** 1989. A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.* **206**:579–590.
66. **Garner, M. M., and A. Revzin.** 1986. The use of gel electrophoresis to detect and study nucleic acid-protein interactions. *Trends Biochem. Sci.* **11**:395–396.
67. **Gartenburg, M.R., and D. M. Crothers.** 1991. Synthetic DNA bending sequences increase the rate of *in vitro* transcription initiation at the *Escherichia coli lac* promoter. *J. Mol. Biol.* **219**:217–230.
68. **Gaston, K., A. Bell, A. Kolb, H. Buc, and S. Busby.** 1990. Stringent spacing requirements for transcription activation by CRP. *Cell* **62**:733–743.
69. **Gilbert, W.** 1976. Starting and stopping sequences for the RNA polymerase, p. 193–206. *In* R. Losick and M. Chamberlin (ed.), *RNA Polymerase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
70. **Gill, S. C., S. E. Weitzel, and P. H. von Hippel.** 1991. *E. coli* Sigma70 and NusA proteins. I. Binding interactions with core RNA polymerase. *J. Mol. Biol.* **220**:307–324.
71. **Grachev, M. A., T. I. Kolocheva, E. A. Lukhtanov, and A. A. Mustaev.** 1987. Studies on the functional topography of *Escherichia coli* RNA polymerase. Highly selective affinity labelling by analogues of initiating substrates. *Eur. J. Biochem.* **163**:113–121.
72. **Grachev, M. A., E. A. Lukhtanov, A. A. Mustaev, E. F. Zaychikov, M. N. Abdukayumov, I. V. Rabinov, V. I. Richter, Y. S. Skoblov, and P. G. Chistyakov.** 1989. Studies on the functional topography of *Escherichia coli* RNA polymerase. A method for localization of the sites of affinity labelling. *Eur. J. Biochem.* **180**:577–585.
73. **Grachev, M. A., and E. F. Zaychikov.** 1980. Initiation by *Escherichia coli* RNA-polymerase: transformation of abortive to productive complex. *FEBS Lett.* **115**:23–26.
74. **Gralla, J. D.** 1989. Bacterial gene regulation from distant DNA sites. *Cell* **57**:193–195.
75. **Gralla, J. D.** 1990. Promoter recognition and mRNA initiation by *Escherichia coli* $E\sigma^{70}$. *Methods Enzymol.* **185**:37–54.
76. **Grana, D., T. Gardella, and M. M. Susskind.** 1988. Effects of mutations in the ant promoter of phage P22 depend on context. *Genetics* **120**:319–327.
77. **Gross, C. A., M. Lonetto, and R. Losick.** 1992. Bacterial sigma factors, p. 129–176. *In* S. L. McKnight and K. R. Yamamoto (ed.), *Transcriptional Regulation*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
78. **Gulland, U., and W. Hillen.** 1992. The Tn10-encoded tetR mRNA has heterogeneous 5′ ends *in vivo* and *in vitro*. *Gene* **114**:97–101.
79. **Ha, J.-H., R. S. Spolar, and M. T. Record, Jr.** 1989. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J. Mol. Biol.* **209**:801–816.
80. **Hansen, U. M., and W. R. McClure.** 1980. Role of the σ subunit of *Escherichia coli* RNA polymerase in initiation. II. Release of σ from ternary complexes. *J. Biol. Chem.* **255**:9564–9570.
81. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343–2361.
82. **Harrison, C., D. Turner, and D. Hinkle.** 1982. Laser crosslinking of *Escherichia coli* RNA polymerase and T7 DNA. *Nucleic Acids Res.* **10**:2399–2414.
83. **Hawley, D., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237–2255.
84. **Hawley, D. K., A. D. Johnson, and W. R. McClure.** 1985. Functional and physical characterization of transcription initiation complexes in the bacteriophage O_R region. *J. Biol. Chem.* **260**:8618–8626.
85. **Hawley, D. K., and W. R. McClure.** 1982. Mechanism of activation of transcription initiation from the λP_{RM} promoter. *J. Mol. Biol.* **157**:493–525.
86. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**:839–872.
87. **Hershberger, P. A., and P. L. deHaseth.** 1991. RNA polymerase bound to the P_R promoter of bacteriophage λ inhibits open complex formation at the divergently transcribed P_{RM} promoter. *J. Mol. Biol.* **222**:479–494.
88. **Hershberger, P. A., B. C. Mita, A. Tripatara, and P. L. deHaseth.** 1993. Interference by P_R bound RNA polymerase with P_{RM} function *in vitro*. *J. Biol. Chem.* **268**:8943–8948.
89. **Heumann, H., H. Lederer, G. Baer, R. P. May, J. K. Kjems, and H. L. Crespi.** 1988. Spatial arrangement of DNA-dependent RNA polymerase of *Escherichia coli* and DNA in the specific complex. *J. Mol. Biol.* **201**:115–125.
90. **Heyduk, T., J. C. Lee, Y. W. Ebright, E. E. Blatter, Y. Zhou, and R. H. Ebright.** 1993. CAP interacts with RNA polymerase in solution in the absence of promoter DNA. *Nature* (London) **364**:548–549.
91. **Hinkle, D. C., and M. J. Chamberlin.** 1972. Studies of the binding of *Escherichia coli* RNA polymerase to DNA. I. The role of sigma in site selection. *J. Mol. Biol.* **70**:157–185.
92. **Hofer, B., D. Muller, and H. Koster.** 1985 The pathway of *E. coli* RNA polymerase-promoter complex formation as visualized by footprinting. *Nucleic Acids Res.* **16**:5995–6013.
93. **Hoopes, B. C., and W. R. McClure.** 1987. Strategies in regulation of transcription initiation, p. 1231–1238. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella*

typhimurium: Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C.
94. **Igarishi, K., N. Fujita, and A. Ishihama.** 1991. Identification of α subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **218**:1–6.
95. **Igarishi, K., N. Fujita, and A. Ishihama.** 1991. Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65**:1015–1022.
96. **Ikeda, R. A., A. C. Lin, and J. Clarke.** 1992. Initiation of transcription by T7 RNA polymerase at its natural promoters. *J. Biol. Chem.* **267**:2640–2649.
97. **Ishihama, A.** 1981. Subunit assembly of *Escherichia coli* RNA polymerase. *Adv. Biophys.* **14**:1–35.
98. **Ishihama, A.** 1993. Protein-protein communication within the transcription apparatus. *J. Bacteriol.* **175**:2483–2489.
99. **Jensen, D. E., and P. H. von Hippel.** 1976. DNA 'melting' proteins. I. Effects of bovine pancreatic ribonuclease binding on the conformation and stability of DNA. *J. Biol. Chem.* **251**:7198–7214.
100. **Jeong, W., and C. Kang.** 1994. Start site selection at *lac*UV5 promoter affected by the sequence context around the initiation sites. *Nucleic Acids Res.* **22**:4667–4672.
101. **Jin, D. J., and C. L. Turnbough.** 1994. An *Escherichia coli* RNA polymerase defective in transcription due to its overproduction of abortive initiation products. *J. Mol. Biol.* **236**:72–80.
102. **Josaitis, C. A., T. Gall, W. Ross, and R. L.Gourse.** 1990. Sequences upstream of the –35 hexamer of *rrnB* P1 affect promoter strength and upstream activation. *Biochim. Biophys. Acta* **1050**:307–311.
103. **Juang, Y.-L., and J. D. Helmann.** 1994. A promoter melting region in the primary σ factor of *Bacillus subtilis*. Identification of functionally important aromatic amino acids. *J. Mol. Biol.* **235**:1470–1488.
104. **Kabata, H., O. Kurosawa, I. Arai, M. Washizu, S. A. Margarson, R. E. Glass, and N. Shimamoto.** 1993. Visualization of single molecules of RNA polymerase sliding along DNA. *Science* **262**:1561–1563.
105. **Kammerer, W., U. Deuschle, R. Gentz, and W. Bujard.** 1986. Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process. *EMBO J.* **5**:2995–3000.
106. **Kashlev, M., J. Lee, K. Zalenskaya, V. Nikiforov, and A. Goldfarb.** 1990. Blocking of the initiation to elongation transition by a transdominant RNA polymerase mutation. *Science* **248**:1006–1009.
107. **Kirkegaard, K., H. Buc, A. Spassky, and J. C. Wang.** 1983. Mapping of single-stranded regions in du*plex* DNA at the sequence level: single-strand-specific cytosine methylation in RNA polymerase-promoter complexes. *Proc. Natl. Acad. Sci. USA* **80**:2544–2548.
108. **Knaus, R., and H. Bujard.** 1990. Principles governing the activity of *E. coli* promoters, p. 110–122. *In* F. Eckstein and D. M. J. Lilley (ed.), *Nucleic Acids and Molecular Biology,* vol. 4. Springer-Verlag, Berlin.
109. **Kobayashi, M., K. Nagata, and A. Ishihama.** 1990. Promoter selectivity of *Escherichia coli* RNA polymerase: effect of base substitutions in the promoter –35 region on promoter strength. *Nucleic Acids Res.* **18**:7367–7372.
110. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz.** 1992. Crystal structure at 3.5Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**:1783–1790.
111. **Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya.** 1993. Transcription regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**:749–795.
112. **Kolb, A., K. Igarashi, A. Ishihama, M. Lavigne, M. Buckle, and H. Buc.** 1993. *E. coli* RNA polymerase, deleted in the C-terminal part of its α-subunit, interacts differently with the cAMP-CRP complex at the *lac*P1 and at the *gal*P1 promoter. *Nucleic Acids Res.* **21**:319–326.
113. **Kovacic, R. T.** 1987. The 0°C closed complexes between *Escherichia coli* RNA polymerase and two promoters, T7-A3 and *lac*UV5. *J. Biol. Chem.* **262**:13654–13661.
114. **Krummel, B., and M. J. Chamberlin.** 1989. RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry* **28**:7829–7842.
115. **Krummel, B., and M. J. Chamberlin.** 1992. Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. *J. Mol. Biol.* **225**:239–250.
116. **Kuhnke, G., H.-J. Fritz, and R. Ehring.** 1987. Unusual properties of promoter-up mutations in the *Escherichia coli* galactose operon and evidence suggesting RNA polymerase-induced bending. *EMBO J.* **6**:507–513.
117. **Kuldell, N., and A. Hochschild.** 1994. Amino acid substitutions in the –35 recognition motif of σ^{70} that result in defects in phage λ repressor-stimulated transcription. *J. Bacteriol.* **176**:2991–2998.
118. **Kumar, A., B. Grimes, N. Fujita, K. Makino, R. A. Malloch, R. S. Hayward, and A. Ishihama.** 1994. Role of the sigma 70 subunit of *Escherichia coli* RNA polymerase in transcription activation. *J. Mol. Biol.* **235**:405–413.
119. **Kumar, A., R. A. Malloch, N. D. Fujita, A. Smillie, A. Ishihama, and R. S. Hayward.** 1993. The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus ten" promoter. *J. Mol. Biol.* **232**:406–418.
120. **Kuwabara, M. D., and D. S. Sigman.** 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10 phenanthroline-copper ion: *Escherichia coli* RNA polymerase-*lac* promoter complexes. *Biochemistry* **26**:7234–7238.
121. **Lanzer, M., and H. Bujard.** 1988. Promoters largely determine the efficiency of repressor action. *Proc. Natl. Acad. Sci. USA* **85**:8973–8977.
122. **Lavigne, M., M. Herbert, A. Kolb, and H. Buc.** 1992. Upstream curved sequences influence the initiation of transcription at the *Escherichia coli* galactose operon. *J. Mol. Biol.* **224**:293–306.
123. **Law, S. M., G. R. Bellomy, P. J. Schlax, and M. T. Record, Jr.** 1993. *In vivo* thermodynamic analysis of repression with and without looping in *lac* constructs. Estimates of free and local *lac* repressor concentrations and of physical properties of a region of supercoiled plasmid DNA *in vivo*. *J. Mol. Biol.* **230**:161–173.
124. **LeClerc, J. E., and N. L. Istock.** 1982. Specificity of UV mutagenesis in the *lac* promoter of M13 *lac* hybrid phage DNA. *Nature* (London) **297**:596–598.
125. **Lee, J., and A. Goldfarb.** 1991. *lac* repressor acts by modifying the initial transcribing complex so that it cannot leave the promoter. *Cell* **66**:793–798.
126. **Leirmo, S., and R. L. Gourse.** 1991. Factor-independent activation of *Escherichia coli rrn*A transcription. I. Kinetic analysis of the roles of the upstream activator region and supercoiling on transcription of the *rrn*B P1 promoter *in vitro*. *J. Mol. Biol.* **220**:555–568.
127. **Leirmo, S., and M. T. Record, Jr.** 1990. Structural, thermodynamic and kinetic studies of the interaction of $E\sigma^{70}$ RNA polymerase with promoter DNA, p. 123–151. *In* D. M. J. Lilley and F. Eckstein (ed.), *Nucleic Acids and Molecular Biology,* vol. 4. Springer-Verlag, Berlin.
128. **Lesley, S. A., and R. R. Burgess.** 1989. Characterization of the *Escherichia coli* transcription factor σ^{70}: localization of a region involved in the interaction with core RNA polymerase. *Biochemistry* **28**:7728–7734.
129. **Lesser, D. R., M. R. Kurpiewski, and L. Jen-Jacobson.** 1990. The energetic basis of specificity in the EcoRI endonuclease-DNA interaction. *Science* **250**:776–786.
130. **Li, M., H. Moyle, and M. M. Susskind.** 1994. Target of the transcriptional activation function of phage λ cI protein. *Science* **263**:75–77.
131. **Lisser, S., and H. Margalit.** 1993. Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.* **21**:1507–1516.
132. **Liu, C., L. S. Heath, and C. L. Turnbough, Jr.** 1994. Regulation of *pyrBI* operon expression in *Escherichia coli* by UTP-sensitive reiterative RNA synthesis during transcriptional initiation. *Genes Dev.* **8**:2904–2912.
133. **Losick, R., and M. Chamberlin (ed.).** 1976. *RNA Polymerase.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
134. **Majors, J.** 1975. Initiation of *in vitro* mRNA synthesis from the wild-type *lac* promoter. *Proc. Natl. Acad. Sci. USA* **72**:4394–4398.
135. **Malan, T. P., A. Kolb, H. Buc, and W. R. McClure.** 1984. Mechanism of CRP-cAMP activation of *lac* operon transcription initiation activation of the P1 promoter. *J. Mol. Biol.* **180**:881–909.
136. **Malan, T. P., and W. R. McClure.** 1984. Dual promoter control of the *Escherichia coli* lactose operon. *Cell* **39**:173–180.
137. **Mangel, W. F., and M. J. Chamberlin.** 1974. Studies of ribonucleic acid chain initiation by *Escherichia coli* ribonucleic acid polymerase bound to T7 deoxyribonucleic acid. *J. Biol. Chem.* **249**:3002–3006.
138. **Manor, H., D. Goodman, and G. S. Stent.** 1969. RNA chain growth rates in *Escherichia coli*. *J. Mol. Biol.* **39**:1–29.
139. **Maslak, M., and C. T. Martin.** 1993. Tests of a model for promoter recognition by T7 RNA polymerase: thymine methyl group contacts. *Biochemistry* **32**:4270–4274.
140. **Maslak, M., and C. T. Martin.** 1993. Kinetic analysis of T7 RNA polymerase transcription initiation from promoters containing single-stranded regions. *Biochemistry* **32**:4281–4285.
141. **Matthews, K. S.** 1992. DNA looping. *Microbiol. Rev.* **56**:123–136.
142. **McClure, W. R.** 1980. Rate-limiting steps in RNA chain initiation. *Proc. Natl. Acad. Sci. USA* **77**:5634–5638.
143. **McClure, W. R.** 1983. A biochemical analysis of the effect of RNA polymerase concentration on the *in vivo* control of RNA chain initiation frequency, p. 207–217. *In* D. L. F. Lennon, F. W. Stratman, and R. N. Zahlten (ed.), *Biochemistry of Metabolic Processes.* Elsevier Scientific Publishers, New York.
144. **McClure, W. R.** 1985. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* **54**:171–204.
145. **McClure, W. R., C. L. Cech, and D. E. Johnston.** 1978. A steady state assay for the RNA polymerase initiation reaction. *J. Biol. Chem.* **253**:8941–8948.
146. **Mecsas, J., D. W. Cowing, and C. A. Gross.** 1991. Development of RNA polymerase-promoter contacts during open complex formation. *J. Mol. Biol.* **220**:585–597.
147. **Meisenberger, O., H. Heumann, and I. Pilz.** 1981. Small-angle x-ray study of DNA-dependent RNA polymerase holoenzyme from *Escherichia coli*. *FEBS Lett.* **123**:22–24.
148. **Melancon, P., R. R. Burgess, and M. T. Record, Jr.** 1982. Nitrocellulose filter binding studies of the interactions of *E. coli* RNA polymerase holoenzyme with deoxyribonucleic acid restriction fragments: evidence for multiple classes of nonpromoter interactions, some of which display promoter-like properties. *Biochemistry* **21**:4318–4331.
149. **Melancon, P., R. R. Burgess, and M. T. Record, Jr.** 1983. Direct evidence for the preferential interaction of *E. coli* RNA polymerase with the ends of DNA restriction fragments. *Biochemistry* **22**:5169–5176.
150. **Menendez, M., A. Kolb, and H. Buc.** 1987. A new target for CRP action at the *mal*T promoter. *EMBO J.* **6**:4227–4234.
151. **Misra, V. K., J. L. Hecht, K. A. Sharp, R. A. Friedman, and B. Honig.** 1994. Salt effects on protein-DNA interactions. The λcI repressor and EcoRI endonuclease. *J. Mol. Biol.* **238**:264–280.
152. **Moore, K. J. M., and T. M. Lohman.** 1994. Kinetic mechanism of adenine nucleotide binding to and hydrolysis by the *Escherichia coli* rep monomer. 1. Use of fluorescent nucleotide analogues. *Biochemistry* **33**:14550–14564.

153. **Moore, K. J. M., and T. M. Lohman.** 1994. Kinetic mechanism of adenine nucleotide binding to and hydrolysis by the *Escherichia coli* rep monomer. 2. Application of a kinetic competition approach. *Biochemistry* **33:**14565–14578.
154. **Mossing, M. C., and M. T. Record, Jr.** 1985. Thermodynamic origins of specificity in the *lac* repressor-operator interaction: adaptability in the recognition of mutant operator sites. *J. Mol. Biol.* **186:**295–305.
155. **Moyle, H., C. Waldburger, and M. M. Susskind.** 1991. Hierarchies of base pair preference in the P22 *ant* promoter. *J. Bacteriol.* **173:**1944–1950.
156. **Mulligan, M.E., J. Brosius, and W. R. McClure.** 1985. Characterization of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter. *J. Biol. Chem.* **260:**3529–3538.
157. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vivo* RNA polymerase selectivity. *Nucleic Acids Res.* **12:**789–800.
158. **Munson, L. M., and W. S. Reznikoff.** 1981. Abortive initiation and long ribonucleic acid synthesis. *Biochemistry* **20:**2081–2085.
159. **Mustaev, A., M. Kashlev, J. Y. Lee, A. Polyakov, A. Lebedev, K. Zalenskaya, M. Grachev, A. Goldfarb, and V. Nikiforov.** 1991. Mapping of the priming substrate contacts in the active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **266:**23927–23931.
160. **Newlands, J. T., W. Ross, K. K. Gosink, and R. L. Gourse.** 1991. Factor-independent activation of *Escherichia coli rrn*A transcription. *J. Mol. Biol.* **220:**569–583.
161. **Niu, W., Y. Zhou, Q. Dong, Y. W. Ebright, and R. H. Ebright.** 1994. Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP). I. Saturation and alanine-scanning mutagenesis. *J. Mol. Biol.* **243:**595–602.
162. **Okada, M., J. Vergne, and J. Brahms.** 1978. Subunit topography of RNA polymerase (*E. coli*) in the complex with DNA. *Nucleic Acids Res.* **5:**1845–1862.
163. **Ollis, D. L., P. Brick, R. Hamlin, N. G. Xuong, and T. A. Steitz.** 1985. Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* (London) **313:**762–766.
164. **Olmsted, M. C. , J. P. Bond, C. F. Anderson, and M. T. Record, Jr.** 1995. GCMC molecular and thermodynamic predictions of ion effects on binding of an oligocation (L^{8+}) to the center of DNA oligomers. *Biophys. J.* **68:**634–647.
165. **Park, C., Z. Hillel, and C.-W. Wu.** 1980. DNA strand specificity in promoter recognition by RNA polymerase. *Nucleic Acids Res.* **8:**5895–5912.
166. **Pelletier, H., M. R. Sawaya, A. Kumar, S. H. Wilson, and J. Kraut.** 1994. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science* **264:**1891–1903.
167. **Perez-Martin, J., F. Rojo, and V. de Lorenzo.** 1994. Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.* **58:**268–290.
168. **Peterson, M. L., and W. S. Reznikoff.** 1985. Properties of *lac*P2 *in vivo* and *in vitro:* an overlapping RNA polymerase binding site within the lactose control region. *J. Mol. Biol.* **185:**525–533.
169. **Rao, L., W. Ross, J. A. Appleman, T. Gaal, S. Leirmo, P. Schlax, M. T. Record, Jr., and R. L. Gourse.** 1994. Factor independent activation of *rrn*BP1: an "extended" promoter element that dramatically increases promoter strength. *J. Mol. Biol.* **235:**1421–1435.
170. **Record, M. T., Jr., J.-H. Ha, and M. Fisher.** 1991. Use of equilibrium and kinetic measurements to determine the thermodynamic origins of stability and specificity and mechanism of formation of site specific complexes between proteins and helical DNA. *Methods Enzymol.* **208:**291–343.
171. **Record, M. T., Jr., T. M. Lohman, and P. L. deHaseth.** 1976. Ion effects on ligand-nucleic acid interactions. *J. Mol. Biol.* **107:**145–158.
172. **Record, M. T., Jr., and R. S. Spolar.** 1990. Some thermodynamic principles of nonspecific and site-specific protein-DNA interactions, p. 33–69. *In* A. Revzin (ed.), *The Biology of Nonspecific DNA-Protein Interactions.* CRC Press, Boca Raton, Fla.
173. **Rees, W. A., R. W. Keller, J. P. Vesenka, G. Yang, and C. Bustamante.** 1993. Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science* **260:**1646–1649.
174. **Ren, Y. L., S. Garges, S. Adhya, and J. S. Krakow.** 1988. Cooperative DNA binding of heterologous proteins: evidence for contact between the cyclic AMP receptor protein and RNA polymerase. *Proc. Natl. Acad. Sci. USA* **85:**4138–4142.
175. **Revzin, A. (ed.).** 1993. *Footprinting of Nucleic Acid-Protein Complexes.* Academic Press, Inc., San Diego, Calif.
176. **Reznikoff, W. S.** 1992. CAP activation of *lac* transcription. *J. Bacteriol.* **174:**655–658.
177. **Reznikoff, W. S., R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr., and M. P. Wickens.** 1986. *RNA Polymerase and the Regulation of Transcription.* Elsevier, New York.
178. **Reznikoff, W. S., D. A. Siegele, D. W. Cowing, and C. A. Gross.** 1985. The regulation of transcription initiation in bacteria. *Annu. Rev. Genet.* **19:**355–387.
179. **Rodriguez, R. L., and M. J. Chamberlin.** 1982. *Promoters: Structure and Function.* Praeger Publishers, New York.
180. **Roe, J.-H., R. R. Burgess, and M. T. Record, Jr.** 1984. Kinetics and mechanism of the interaction of *E. coli* RNA polymerase with the λP_R promoter. *J. Mol. Biol.* **176:**495–521.
181. **Roe, J.-H., R. R. Burgess, and M. T. Record, Jr.** 1985. Temperature dependence of the rate constants of the *E. 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.
182. **Roe, J.-H., and M. T. Record, Jr.** 1985. Regulation of the kinetics of the interaction of *Escherichia coli* RNA polymerase with the λP_R promoter by salt concentration. *Biochemistry* **24:**4721–4726.
183. **Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R. L. Gourse.** 1993. A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262:**1407–1413.
184. **Sadler, J. R., and A. Novick.** 1965. The properties of repressor and the kinetics of its action. *J. Mol. Biol.* **12:**305–327.
185. **Sagitov, V., V. Nikiforov, and A. Goldfarb.** 1993. Dominant lethal mutations near the 5′ substrate binding site affect RNA polymerase propagation. *J. Biol. Chem.* **268:**2195–2202.
186. **Sasse-Dwight, S., and J. D. Gralla.** 1989. $KMnO_4$ as a probe for *lac* promoter DNA melting and mechanism *in vivo*. *J. Biol. Chem.* **264:**8074–8081.
187. **Schick, C., and C. T. Martin.** 1993. Identification of specific contacts in T3 RNA polymerase-promoter interactions: kinetic analysis using small synthetic promoters. *Biochemistry* **32:**4275–4280.
188. **Schickor, P., W. Metzger, W. Werel, H. Lederer, and H. Heumann.** 1990. Topography of intermediates in transcription initiation of *E. coli*. *EMBO J.* **9:**2215–2220.
189. **Schlax, P. J., M. W. Capp, and M. T. Record, Jr.** 1995. Inhibition of transcription initiation by *lac* repressor. *J. Mol. Biol.* **245:**331–350.
190. **Schleif, R.** 1988. DNA looping. *Science* **240:**127–128.
191. **Schmitz, A., and D. J. Galas.** 1979. The interaction of RNA polymerase and *lac* repressor with the *lac* control region. *Nucleic Acids Res.* **6:**111–137.
192. **Schultz, S. C., G. C. Shields, and T. A. Steitz.** 1991. Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. *Science* **253:**1001–1007.
193. **Sellitti, M. A., P. A. Pavco, and D. A. Steege.** 1987. *lac* repressor blocks *in vivo* transcription of *lac* control region DNA. *Proc. Natl. Acad. Sci. USA* **84:**3199–3203.
194. **Severinov, K., D. Fenyo, E. Severinova, A. Mustaev, B. T. Chait, A. Goldfarb, and S. A. Darst.** 1994. The sigma subunit conserved region 3 is part of the 5′ face of active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **269:**20826–20828.
195. **Shaner, S. L., P. Melancon, K. S. Lee, R. R. Burgess, and M. T. Record, Jr.** 1983. Ion effects on the aggregation and DNA binding reactions of *E. coli* RNA polymerase. *Cold Spring Harbor Symp. Quant. Biol.* **47:**463–472.
196. **Siebenlist, U.** 1979. RNA polymerase unwinds an 11 base pair segment of a phage T7 promoter. *Nature* (London) **279:**651–652.
197. **Siebenlist, U., R. B. Simpson, and W. Gilbert.** 1980. *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* **20:**269–281.
198. **Siegele, D. A., J. C. Hu, W. A. Walter, and C. A. Gross.** 1989. Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **206:**591–603.
199. **Simpson, R. B.** 1979. Molecular topography of RNA polymerase interaction. *Cell* **18:**277–285.
200. **Sørensen, K. I., K. E. Baker, R. A. Kelln, and J. Neuhard.** 1993. Nucleotide pool-sensitive selection of the transcriptional start site in vivo at the *Salmonella typhimurium pyrC* and *pyrD* promoters. *J. Bacteriol.* **13:**4137–4144.
201. **Sousa, R., Y. J. Chung, J. P. Rose, and B. C. Wang.** 1993. Crystal structure of bacteriophage T7 RNA polymerase at 3.3Å resolution. *Nature* (London) **364:**593–599.
202. **Sousa, R., J. P. Rose, and B. C. Wang.** 1994. The thumb's knuckle: flexibility in the thumb subdomain of T7 RNA polymerase is revealed by the structure of a chimeric T7/T3 RNA polymerase. *J. Mol. Biol.* **244:**6–12.
203. **Spassky, A.** 1986. Visualization of the movement of the *Escherichia coli* RNA polymerase along the *lac* UV5 promoter during the initiation of the transcription. *J. Mol. Biol.* **188:**99–103.
204. **Spassky, A., K. Kirkegaard, and H. Buc.** 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.
205. **Spolar, R. S., and M. T. Record, Jr.** 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263:**777–784.
206. **Stefano, J. E., and J. Gralla.** 1979. *lac*UV5 transcription *in vitro.* Rate limitation subsequent to formation of an RNA polymerase-DNA complex. *Biochemistry* **18:**1063–1067.
207. **Stefano, J. E., and J. Gralla.** 1982. Spacer mutations in the *lac*Ps promoter. *Proc. Natl. Acad. Sci. USA* **79:**1069–1072.
208. **Steitz, T. A., S. J. Smerdon, J. Jäger, and C. M. Joyce.** 1994. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266:**2022–2025.
209. **Straney, D. C., and D. M. Crothers.** 1985. Intermediates in transcription initiation from the *E. coli lac* UV5 promoter. *Cell* **43:**449–459.
210. **Straney, D. C., and D. M. Crothers.** 1987. A stressed intermediate in the formation of stably initiated RNA chains at the *Escherichia coli lac* UV5 promoter. *J. Mol. Biol.* **193:**267–278.
211. **Straney, D. C., and D. M. Crothers.** 1987. Comparison of the open complexes formed by RNA polymerase at the *Escherichia coli lac* UV5 promoter. *J. Mol. Biol.* **193:**279–292.
212. **Straney, D. C., S. B. Straney, and D. M. Crothers.** 1989. Synergy between *Escherichia coli* CAP protein and RNA polymerase in the *lac* promoter open complex. *J. Mol. Biol.* **206:**41–57.
213. **Straney, S. B., and D. M. Crothers.** 1987. *lac* repressor is a transient gene-activating protein. *Cell* **51:**699–707.
214. **Strauss, H. S., R. R. Burgess, and M. T. Record, Jr.** 1980. Binding of *Escherichia coli* ribonucleic acid polymerase holoenzyme to a bacteriophage T7 promoter-containing fragment: evaluation of promoter binding constants as a function of solution conditions. *Biochemistry* **19:**3504–3515.
215. **Strauss, H. S., R. R. Burgess, and M. T. Record, Jr.** 1980. Binding of *Escherichia coli* ribonucleic acid polymerase holoenzyme to a bacteriophage T7 promoter-containing fragment: selectivity exists over a wide range of solution conditions. *Biochemistry* **19:**3496–3504.

216. **Su, T. T., and W. R. McClure.** 1994. Selective binding of *Escherichia coli* RNA polymerase to topoisomers of minicircles carrying the TAC16 and TAC17 promoters. *J. Biol. Chem.* **269**:13511–13521.
217. **Suh, W. C., S. Leirmo, and M. T. Record Jr.** 1992. Roles of Mg^{2+} in the mechanism of formation and dissociation of open complexes between *E. coli* RNA polymerase and the λP_R promoter: kinetic evidence for a second open complex requiring Mg^{2+}. *Biochemistry* **31**:7815–7825.
218. **Suh, W. C., W. Ross, and M. T. Record, Jr.** 1993. Two open complexes and a requirement for Mg^{2+} to open the λP_R transcription start site. *Science* **259**:358–361.
219. **Sweetser, D., M. Nonet, and R. A. Young.** 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc. Natl. Acad. Sci. USA* **84**:1192–1196.
220. **Szoke, P. A., T. L. Allen, and P. L. deHaseth.** 1987. Promoter recognition by *Escherichia coli* RNA polymerase: effects of base substitutions in the –10 and –35 regions. *Biochemistry* **26**:6188–6194.
221. **Takeya, T., and M. Takanami.** 1974. Isolation of DNA segments containing promoters from bacteriophage T3 DNA. *Biochemistry* **13**:5388–5394.
222. **Tanaka, K., S. Kusano, N. Fujita, A. Ishihama, and H. Takahashi.** 1995. Promoter determinants for *Escherichia coli* RNA polymerase holoenzyme containing σ^{38} (the *rpoS* gene product). *Nucleic Acids Res.* **23**:827–834.
223. **Tang, H., K. Severinov, A. Goldfarb, D. Fenyo, B. Chait, and R. H. Ebright.** 1994. Location, structure, and function of the target of a transcriptional activator protein. *Genes Dev.* **8**:3058–3067.
224. **Taylor, W. E., and R. R. Burgess.** 1979. *Escherichia coli* RNA polymerase binding and initiation of transcription on fragments of λrif^d 18 DNA containing promoters for λ genes and for rrnB, tufB, rplK,A, rplJ,L, and rpoB,C genes. *Gene* **6**:331–365.
225. **Travers, A. A.** 1987. Structure and function of *E. coli* promoter DNA. *Crit. Rev. Biochem.* **22**:181–219.
226. **Tullius, T.** 1989. Physical studies of protein-DNA complexes by footprinting. *Annu. Rev. Biophys. Biophys. Chem.* **18**:213–237.
227. **Tullius, T., B. A. Dombroski, M. E. A. Churchill, and L. Kam.** 1987. Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155:537–558.
228. **von Hippel, P. H.** 1979. On the molecular bases of the specificity of interaction of transcriptional proteins with genome DNA, p. 279–347. *In* R. F. Goldberger (ed.), *Biological Regulation and Development.* Plenum Publishing Corp., New York.
229. **von Hippel, P. H., and O. G. Berg.** 1989. Facilitated target location in biological systems. *J. Biol. Chem.* **264**:675–678.
230. **von Hippel, P. H., A. Revzin, C. A. Gross, and A. C. Wang.** 1974. Non-specific DNA binding of genome regulating proteins as a biological control mechanism. 1. The *lac* operon: equilibrium aspects. *Proc. Natl. Acad. Sci. USA* **71**:4808–4812.
231. **Waldburger, C., T. Gardella, R. Wong, and M. M. Susskind.** 1990. Changes in conserved region 2 of *Escherichia coli* σ70 affecting promoter recognition. *J. Mol. Biol.* 215:267–276.
232. **Waldburger, C., and M. M. Susskind.** 1994. Probing the informational content of *Escherichia coli* σ70 region 2.3 by combinatorial cassette mutagenesis. *J. Mol. Biol.* **235**:1489–1500.
233. **Warne, S. E., and P. L. deHaseth.** 1993. Promoter recognition by *Escherichia coli* RNA polymerase. Effects of single base pair deletions and insertions in the spacer DNA separating the –10 and –35 regions are dependent on spacer DNA sequence. *Biochemistry* **32**:6134–6140.
234. **Woody, S. T., R. S.-C. Fong, and G. N. Gussin.** 1993. Effects of a single base pair deletion in the bacteriophage λP_{RM} promoter. *J. Mol. Biol.* **229**:37–51.
235. **Wu, H. M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein induced DNA bending. *Nature* (London) **308**:509–513.
236. **Xiong, X.F., and W. S. Reznikoff.** 1993. Transcriptional slippage during the transcription initiation process at a mutant *lac* promoter *in vivo*. *J. Mol. Biol.* **231**:569–580.
237. **Zhou, Y., X. Zhang, and R. H. Ebright.** 1993. Identification of the activating region of CAP: isolation and characterization of mutants of CAP specifically defective in transcription initiation. *Proc. Natl. Acad. Sci. USA* **90**:6081–6085.
238. **Zillig, W., P. Palm, and A. Heil.** 1976. Function and reassembly of subunits of DNA-dependent RNA polymerase, p. 101–125. *In* R. Losick and M. Chamberlin (ed.), *RNA Polymerase.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
239. **Zillig, W., K. Zechel, D. Rabussay, M. Schachner, V. S. Sethi, P. Palm, A. Heil, and W. Seifert.** 1970. On the role of different subunits of DNA-dependent RNA polymerase from *E. coli* in the transcription process. *Cold Spring Harbor Symp. Quant. Biol.* **35**:47–58.
240. **Zimmerele, C. T., and C. Frieden.** 1989. Analysis of progress curves by simulations generated by numerical integration. *Biochem. J.* **258**:381–387.
241. **Zimmerele, C. T., K. Patane, and C. Frieden.** 1987. Divalent cation binding to the high- and low-affinity sites on G-actin. *Biochemistry* **26**:6545–6552.
242. **Zou, C., N. Fujita, K. Igarashi, and A. Ishihama.** 1992. Mapping the cAMP receptor protein contact site on the alpha subunit of *Escherichia coli* RNA polymerase. *Mol. Microbiol.* **6**:2599–2605.

APPENDIX
Nomenclature of Sigma Factors from *Escherichia coli* and *Salmonella typhimurium* and Relationships to Sigma Factors from Other Organisms

MICHAEL A. LONETTO AND CAROL A. GROSS

There is some confusion with regard to the nomenclature of sigma factors in *Escherichia coli,* especially when compared to those from *Bacillus subtilis* and other organisms. Similar sigma factors have different names, while different sigma factors have identical names (5). *E. coli* does not contain homologs to the sporulation sigmas σ^E, σ^F, σ^G, and σ^K found in *Bacillus* and other gram-positive organisms. On the other hand, σ^{28} of *E. coli* is sometimes referred to as σ^F (for flagellar synthesis) while it is in fact homologous to σ^D of *Bacillus.* A similar situation is seen with σ^E from *E. coli,* which is not homologous to σ^E of *Bacillus* but belongs to a group which is most similar to σ^H (6). In addition, the stress response sigma factor σ^{32} is unrelated to σ^B, the stress sigma from *Bacillus.*

We have polled a number of workers in this field and have found a nearly uniform dissatisfaction with the present nomenclature. However, the community is divided on whether to base sigma names on molecular weights versus letter designations. Further, among those who prefer letter designations there is disagreement on which letters to use. We therefore endorse the use of molecular weight designations in the table below, while hoping that these problems can be resolved in a more comprehensive manner in the proper forum. The following table lists the sigma factors from *E. coli* and *Salmonella typhimurium,* along with homologous sigma factors, which have similarity of both sequence and function, as well as related sigma factors with either different or unknown function.

LITERATURE CITED

1. **Helmann, J. D.** 1991. Alternative sigma factors and the regulation of flagellar gene expression. *Mol. Microbiol.* 5:2875–2882.
2. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* 57:839–872.
3. **Kustu, S., E. E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* 53:367–376.
4. **Loewen, P. C., and R. Hengge-Aronis.** 1994. The role of the sigma factor sigma S (KatF) in bacterial global regulation. *Annu. Rev. Microbiol.* 48:53–80.
5. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. *J. Bacteriol.* 174:3843–3849.
6. **Lonetto, M. A., K. L. Brown, K. E. Rudd, and M. J. Buttner.** 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase σ factors involved in the regulation of extracytoplasmic functions. *Proc. Natl. Acad. Sci. USA* 91:7573–7577.

Name and synonyms	Gene	Functions	Homologous sigma factors	Sigma factors related by sequence	References
σ^{70}, σ^{D}*	*rpoD*	Major sigma factor during exponential growth	In all eubacteria; approx. 40 cloned; gene usually designated either *rpoD* (σ^{70}, σ^{80}) or *sigA* (σ^{A})	Streptomyces and *Anabaena* species contain families of alternative sigma factors closely related to the primary sigma factor	2, 5
σ^{38}, σ^{S}	*rpoS, katF*	Major sigma factor during stationary phase; response to oxidative and osmotic stress; expression of virulence genes in *Salmonella* and *Yersinia*	RpoS (σ^{S}) (*Salmonella enterica, S. typhi, S. dublin, Y. enterocolitica, Pseudomonas aeruginosa, Shigella flexneri*)		
σ^{32}, σ^{H}*	*rpoH, htpR*	Transcription of heat shock proteins induced by cytoplasmic stress	RpoH (σ^{32}) (*Haemophilus influenzae, P. aeruginosa*), HtpR (*Citrobacter freundii*)	σ^{B}, σ^{C} (*Myxococcus xanthus* and *Stigmatella aurantiaca* fruiting body formation and sporulation)	5
σ^{28}, σ^{F}*	*fliA, flaD* (*E. coli*)	Expression of late flagellar genes, including flagellin	σ^{D} (*B. subtilis*), σ^{F} (*P. aeruginosa*)	WhiG (*Streptomyces coelicolor* early sporulation)	1
σ^{24}, σ^{E}*	*rpoE*	Response to periplasmic stress	Uncertain whether any of the sequence-related sigma factors have similar functions	σ^{E} (*Photobacterium* sp., *H. influenzae, Mycobacterium leprae*), *algT* (*algU*) (*P. aeruginosa,* alginate production), HrpL (*P. syringae,* virulence factors)	6
σ^{54}, σ^{N}	*rpoN, ntrA, glnF*	Pleiotropic functions: nitrogen metabolism, formate degradation, phage shock response	Approx. 20 so far known, usually designated σ^{54}, σ^{55}, or σ^{N}; genes designated *rpoN* or *ntrA*	σ^{L} (*B. subtilis* levansucrase regulation)	3
σ^{19}, FecI	*fecI*	Iron citrate transport	PupI (*P. putida*), PvdS (*P. fluorescens, P. aeruginosa*) (iron siderophore regulation)	σ^{X} (*B. subtilis*), *nccH* (*Alcaligenes xylosoxidans* nickel resistance), *cnrH* (*A. eutrophus* cobalt and nickel resistance)	6

*Conflicts with *B. subtilis* nomenclature.