STRUCTURE AND FUNCTION OF *E. COLI* PROMOTER DNA

Author: **Andrew A. Travers**
Laboratory of Molecular Biology
Medical Research Council
Cambridge, England

Referee: Henri Buc
Interactions Moleculaires et Cellulaires
CNRS
Paris, France

I. INTRODUCTION

The process of RNA chain initiation by a bacterial RNA polymerase requires that this enzyme, either by itself or in concert with other proteins, acts as a sequence specific DNA binding protein, then as a protein that unwinds DNA and ultimately effects strand separation over a limited region, and, finally, as an RNA synthetase. A further facet of this process is that it is regulated both generally and in a promoter specific manner. This regulation requires that the RNA polymerase interacts directly with and responds to both low molecular weight effectors and other regulatory proteins.

In the initiation reaction, the intrinsic physical and chemical properties of the DNA are crucial to promoter function. In this review, the author discusses the active role of DNA structure in determining polymerase function and shows how these considerations complement recent kinetic data.[1-3] A model derived from this discussion can be directly related to mechanisms regulating polymerase-promoter interactions. The area covered by this review complements recent comprehensive reviews of the thermodynamic, kinetic, and genetic aspects of the subject.[4-6]

A. RNA Polymerase

The eubacterial RNA polymerase is a complex oligomeric protein that is structurally and evolutionarily related to the RNA polymerases of archaebacteria and eukaryotes. As normally isolated from *Escherichia coli*, the enzyme can be functionally separated into two macromolecular components, the core polymerase and a σ-factor.[7] The core polymerase, as normally purified, has the subunit composition $\alpha_2\beta\beta'\omega_2$ with individual molecular weights of 155,613, 150,618, 36,512, and ~10,000 daltons, respectively.[8] The genes for α, β, and β′, respectively, *rpoA*, *rpoB*, and *rpoC*, have been cloned and shown to be essential for polymerase function, but the status of the ω-polypeptide remains in doubt. The major σ-factor in *E. coli* is the vegetative σ of molecular weight 70,263 daltons and is encoded by the *rpoD* gene.[9] This polypeptide binds to core polymerase to form the vegetative holoenzyme, the form of RNA polymerase active in the selection of the majority of *E. coli* promoters. The transcription of other sets of genes requires specific recognition by different holoenzymes, normally present in low abundance. Thus, the genes activated by heat shock and those required for nitrogen fixation require, respectively, σ-factors encoded by the *rpoH* and *rpoN* (*ntrA*) genes.[10-13] The structures of different σ-factors show homology with each other. In general, they contain two regions of helix-turn-helix motif characteristic of a particular class of sequence specific DNA binding proteins.[14-16] In addition, there is a further conserved region that has been postulated to be a site of contact between the core polymerase and the σ-polypeptide. Relative to *Bacillus subtilis*, few σ-proteins have been characterized in *E. coli*. However, it seems probable that other regulatory proteins functioning as σ-factors are likely to be identified.

Table 1
CONSERVED SEQUENCES IN *E. COLI* PROMOTERS, ARRANGED ACCORDING TO DISTANCE FROM STARTPOINT

σ-Factor		−35	−25	−16	−10	Ref.	
E. coli		σ_{70}	TTGACa	—	TATGT	TATAAT	
		σ_{32}	T-tC-cCTTGAA	—	—	CCCCATtTa	
		NtrA	—	CTGGcA	TTGCA	TA	
		T4gp55	—	—	—	TATAAAT	
B. subtilis		σ_{43}	TTGACa			TATAAT	24
		σ_{37}	AGG-TT			GG-ATTG-T	24
		σ_{32}	AAATC			TA-TG-TT-TA	24
		σ_{29}	TT-AAA			CATATT	24
		σ_{28}	CTAAA			CCGATAT	24
		SPO1 gp28	T-AGGAGA—A			TTT-TTT	24
		SPO1 gp33-34	CGTTAGA			GATATT	24

Note: Note that the "−16" sequence for the NtrA protein is not necessarily functionally homologous to the corresponding region for σ_{70} and σ_{32}, respectively.[23]

Little is known about the detailed three-dimensional structure of the vegetative holoenzyme. X-ray scattering and electron microscope studies provide data about the overall shape of the molecule in solution,[18] but the precise path of the DNA around the enzyme in the polymerase-promoter complex remains to be established.

II. PROMOTER STRUCTURE

A. Sequence Arrangement

The DNA sequence of a promoter site must reflect both its capacity to act as a specific recognition site for a DNA binding protein and its subsequent enzymatic manipulation resulting in strand separation. The sequences of more than 200 promoters utilized by RNA polymerase from *E. coli* (and the closely related *Salmonella typhimurium*) have now been determined (see Reference 18 for sequences determined up to 1982), and it is clear that the characteristics of different classes of promoter allow both differential recognition and the common requirement for strand separation.

Differential recognition of promoters by holoenzyme species containing different σ-subunits is the principal basis for the separation of promoter sequences into separate classes. Within each class, two regions of strong sequence homology are normally apparent (Table 1). The most abundant class of promoters is that recognized by the vegetative holoenzyme $E\sigma_{70}$. The 2 most highly conserved regions, TTGACa and TATAAT, are centered approximately 33 and 10 bp upstream from the startpoint and are normally termed the "−35" and "−10" boxes.[18] The separation between these 2 regions is variable, but for the great majority of these promoters it is 16 to 18 bp, with 17 bp being of the most frequent occurrence. Similarly the separation between the −10 region and the startpoint varies between 4 and 8 bp. In addition to these regions of strong homology, there is a weakly conserved sequence, the "−16" box, immediately upstream of −10 and an AT-rich region at −45.[18] The sequence arrangements of other promoter classes recognized by different holoenzymes show significant departures from this pattern. Thus, although σ_{32}, responsible for the recognition of promoters activated under conditions of heat shock stress, has a similar spacing between the two conserved major regions,[19] the RpoN protein, which acts as a σ-factor for genes involved in nitrogen metabolism, has 2 conserved regions separated by 1 double helical turn

at −25 and −16.[20,21] By contrast, again, promoters utilized by the phage T4 encoded gp55 σ-factor contain only a single region of conserved sequence similar in position to the −10 region of σ_{70} promoters.[22]

Despite this diversity of promoter structure, one structural motif is apparently common to all classes of promoter, in eukaryotes, eubacteria, and archaebacteria. This motif is the TATA sequence (or in an abbreviated form simply TA), normally located at ∼ −10 in eubacterial promoters. This sequence or closely related sequences are also found at DNA replication origins and at crossover points for site specific recombination.[23]

B. Bending and Bendability

The interaction of a protein with DNA often involves imposing a curvature on the DNA molecule along a defined track on the surface of a protein complex. Examples of such DNA protein complexes are the nucleosome[25,26] and *E. coli* DNA gyrase.[27] The ability of a DNA molecule to assume an appropriate curved configuration is dependent on two factors, bendability and intrinsic bending.[26,28]

The bendability of a DNA molecule is its ability to assume a preferred direction of curvature when constrained in a tight loop on the surface of a protein or in a small circle. When a DNA double helix is bent in this way, the external circumference of the circle or loop will be significantly greater than the internal circumference. This means that the average width of the major and minor grooves must be greater on the outside than on the inside, and, consequently, since the groove width of a DNA molecule is strongly sequence dependent;[29] particular sequences will adopt a preferred orientation relative to the direction of curvature. A further consideration is that in curved DNA the direction of the helical axis must change. Such changes are correlated with the angle between the long axis of adjacent base pair planes termed "roll".[30,31] In the nucleosome, the rotational positioning of DNA about the histone octamer appears to be determined by such sequence-dependent modulations of DNA structure.[26] In particular the sequences ApApA/TpTpT, ApApT/ApTpT, and TpApA/TpTpA are preferentially located where the minor groove faces inward toward the center of curvature, while the sequences RpGpC/GpCpY, RpTpG/CpApY, and CpG are preferentially located where the minor groove faces outward.[26]

To what extent do these considerations apply to promoter DNA? There is no direct estimate of the magnitude and extent of curvature of DNA in the polymerase-promoter complex. However, any preferred direction of curvature of promoter DNA in this complex should be reflected in the preferred occurrence of sequences correlated with rotational orientation at particular positions in the promoter. Analysis of available promoter sequences shows that the sequences associated with an outward facing minor groove are enriched at positions ∼ −16, −26, −38, and −48. The average separation between these positions is ∼10.65 bp, a value close to the helical twist of DNA free in solution.[32,33]

Intrinsic bending, or more correctly the adoption of a highly preferred direction of bending of the DNA double helix, was first apparent experimentally in DNA fragments, isolated from the kinetoplast of *Leishmania*, which migrated anomalously in polyacrylamide gel electrophoresis.[34] Subsequent analysis has shown that this unusual property is dependent upon the presence of homopolymer (dA) · (dT) stretches that are usually ≥4 bp in extent.[35-37] Such a homopolymer stretch assumes in solution a highly characteristic structure with a screw of 10 bp per turn and a relatively narrow minor groove. This structure can be formed by sequences with both AA and AT steps, but is broken both by the substitution of GC base pairs for AT base pairs and by the presence of a TA step. How might such a sequence impose an intrinsic bend on the axis of the double helix? Either the sequence is itself intrinsically curved with each ApA step acting as a "wedge"[38] or the geometry of the junctions between such (dA) · (dT) runs and the flanking base pairs dictate a change in the direction of the helical axis.[39-41] Both these possibilities require that the structures associated

with homopolymer (dA) · (dT) runs be rigid, i.e., lacking bendability. On the nucleosome core, these runs have a strong tendency to be oriented such that the junctions occur in regions of maximum curvature (i.e., where the minor groove points towards or away from the histone octamer), while the centers of the homopolymer sequence occur in regions of minimum curvature.[42] Nuclease digestion studies also provide evidence for a structural discontinuity at one such junction.[43] These observations clearly suggest that runs of (dA) · (dT) ≥4 bp in extent are both rigid and approximately straight and thus are not curved as required by the "wedge" concept. The intrinsic bending associated with such sequences probably results from the stacking of immediately flanking base pairs on highly propeller-twisted (dA) · (dT) base pairs which themselves are maintained in a relatively inflexible stack.

The distribution of dA:dT blocks within promoter sequences can thus both define in part by the intrinsic bend the potential pathway of DNA on the polymerase and also delimit by their rigidity the regions which can be bent. A statistical survey of this distribution reveals first that such sequences occur most frequently in the neighborhood of position −45 and, second, that they are often found in conserved locations further upstream.[28] This distribution correlates with electrophoretic anomalies observed for DNA fragments containing the upstream regions of the *rrnB* P_1 and the *S. typhimurium hisR* promoters.[44,45] Second, (dA) · (dT) blocks ≥5 bp in extent appear to be preferentially excluded from 3 regions sited approximately at −16, −26, and −38 and coinciding with sequences correlated with high curvature.

C. Low-Energy Transitions

The structure of a DNA molecule in solution is not static, but is a dynamic mixture of different conformations. The transitions between such different structures are highly sequence dependent. Although there is substantial evidence that DNA can undergo structural transitions which have a high activation energy, for example, the formation of left-handed Z-DNA, it seems highly probable that low-energy transitions may be more biologically relevant. In the context of promoter structure, two such transitions are likely to be crucial to promoter function. The first is the probability of local unstacking. This is promoted by high temperature, certain solvents, and negative supercoiling.[23,46,47] Experimentally, this local unstacking is detected by enhanced cleavage of the unwound DNA by S1 nuclease[23] and by an increased probability of interaction with the cleavage agent 1,10-phenanthroline-copper ion.[48] The sequences which are most easily unwound are those of lowest thermal stability.[49] Thus, negative supercoiling and low ionic strength preferentially induce unwinding at TA doublets whose intrinsic melting temperature is some 20°C less than that of the next most easily melted doublets. However, the probability of unwinding is clearly dependent on the nature of the flanking sequence.[23] This property of TA may be sufficient to explain its high conservation in the −10 region of bacterial promoter sequences.

The second class of low-energy transition is induced by lowering of water activity, an event which could occur in an interaction with a DNA binding protein. Examples of such transitions include the conversion of the rigid 10 bp per turn dA:dT structure with a narrow minor groove to a ~10.5 bp per turn structure with a minor groove of average width[29] and the B → A transition.

D. The Consensus Structure of a Promoter

To date, the analysis of *E. coli* promoter sequences has been directed towards the establishment of sequence homologies and of the relationship between such homologies and promoter activity.[50] This approach accurately defines the most highly conserved base pairs located in the −10 and −35 regions and also the optimal 17-bp separation between these regions.[18] Since most genetically selected promoter mutations fall within these same conserved regions, it follows that these sequences must define the strongest interactions between

the promoter DNA and RNA polymerase. However, recent experimental data show that promoters which have an exact or very close fit to the consensus sequences in these conserved regions exhibit only moderate activity relative to the most active promoters.[51] This indicates that there are additional sequence constraints that determine promoter strength. These constraints must reflect more general properties of protein-nucleic acid interactions than sequence specific recognition. For example, for functionally significant properties such as bendability, the preferred sequences are both redundant and independent of orientation[42] and thus would not be apparent from simple homology. In addition, the local structure of DNA is strongly dependent on the sequence context. Consequently, any single base change away from a consensus sequence that also alters the conformation or properties of neighboring base pairs may have a much greater effect on promoter function than would be apparent from the change in homology. One example of such a change is the p27 mutation in the *tyrT* promoter in which a T → A transversion converts the wild-type −10 region TATGAT to TATGAA.[52] In this case, the mutation alters the physical properties of the DNA molecule at least 5 bp distant from the substitution.[23]

In addition to the functionally important sequences upstream of the transcription startpoint, recent experiments have shown that sequences within the transcribed region itself can also significantly affect promoter activity.[53] Thus, substitution of downstream regions can alter promoter strength by up to about tenfold both in vivo and in vitro.[53] The molecular mechanism for this effect has not been established. However, it is known that the downstream sequence can affect the timing of the release of the σ-polypeptide.[54] At the λ P_R promoter, σ does not dissociate from the core enzyme until the transcript has reached a length of 13 nucleotides, whereas at the T7 A1 promoter or with poly-(dAT) · (dAT) as template release of σ occurs after the synthesis of a trinucleotide. The authors suggest that this effect reflects the pattern of hydrogen bond donors and acceptors in the major groove of the nascent RNA-DNA hybrid. Such a property would be consistent with the absence of any highly conserved sequences in this region.

II. THE MECHANISM OF TRANSCRIPTION INITIATION

A. General Principles

Classically, the initiation of transcription by RNA polymerase has been considered to be essentially a two-step process described by the equation:

$$R + P \underset{}{\overset{K_B}{\rightleftharpoons}} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} RP_o \rightarrow \text{initiation}$$

This model, initially formulated by Zillig and collaborators in 1967,[55] assumes that the first step in the process is the formation of a complex between RNA polymerase and the double-stranded DNA of the promoter site. This complex, termed a "closed" complex (RP_c), is in rapid equilibrium with free polymerase.[56] A subsequent step requiring DNA strand separation results in the formation of a more stable "open" complex (RP_o) in which one to one-and-a-half turns of the DNA double helix in the vicinity of the transcription initiation site are no longer base paired.

The bipartite model has been extensively investigated, in particular, by the abortive initiation assay devised by McClure.[57] The analysis of the data so obtained assumes that open complex formation can be approximately modeled on a two-step reaction and assigns three parameters to the reaction: K_B, the thermodynamic association constant for closed complex, and k_2 and k_{-2}, kinetic rate constants for the formation of the open complex from the closed complex and for the reverse reaction, respectively. Under standard conditions,

this assay has been successful in predicting the relative in vivo strengths of different promoters from the in vitro-derived parameter $K_B k_2$.

Mutations in promoter sequences affect both K_B and k_2. In general, mutations in the -35 region have a substantial effect on K_B and a smaller, although significant, effect on k_2.[58,59] In contrast, mutations in -10 region affect principally k_2.[3] In general, the relative values of the second-order constant, $K_B k_2$, can be predicted with some accuracy from the promoter sequence.[50]

It has been apparent for some years that the bipartite model is insufficient to explain certain structural and kinetic features of the initiation reactions, and, consequently, various more complex models have been proposed to account for the observed discrepancies. More recently a combination of structural and kinetic studies have resulted in the proposal of a three-step model:

$$R + P \overset{K_B}{\leftrightharpoons} RP_c \leftrightharpoons RP_i \leftrightharpoons RP_o \rightarrow \text{initiation}$$

where an "intermediate" complete (RP_i) is a distinguishable transition between the closed and open complexes.[1-3,60,61] On this model, the k_2 and k_{-2} constants defined for the bipartite model still apply to the overall transitions between RP_c and RP_o.

B. Promoter Location

The initial step in the formation of a closed complex must be the location of the specific promoter site by RNA polymerase. If the rate of this reaction were determined by the rate of diffusion in solution, it would be expected that the forward rate of location should not exceed $10^8\ M^{-1}\ sec^{-1}$.[4] However, the observed rates of complex formation at phage promoters with a high affinity for RNA polymerase are considerably higher and may exceed $10^{10}\ M^{-1}\ sec^{-1}$.[46] This discrepancy suggests that promoter location proceeds not by the direct interactions of the two components free in solution, but through a different process. RNA polymerase, like other DNA binding proteins, has a significant nonspecific affinity for DNA such that the average ratio of affinity between specific (promoter) sites and nonspecific sites K_{Rs}/K_{Rns} is $\sim 10^4$.[63] This means that when DNA and polymerase are mixed in solution, an individual polymerase molecule will be in close association with the DNA for a high proportion of the time. This nonspecific affinity of the enzyme for DNA thus effectively increases the target size for an initial interaction, since on any DNA molecule the number of nonspecific binding sites will greatly exceed specific sites. Once bound nonspecifically, a polymerase molecule can migrate along a DNA molecule in an essentially one-dimensional diffusion limited manner. Such facilitated diffusion can potentially significantly enhance the rate of specific site location and appears to be the mechanism by which the *lac* repressor locates its operator site. For RNA polymerase, the observed rates of location are consistent with the same mechanism. A second prediction of the sliding model is that the target size, i.e., the length of a promoter containing DNA molecule, should be an important determinant of the location rate. Published data indicate that a DNA fragment of 100 bp still exhibits fast association kinetics, suggesting that the required domain for RNA polymerase sliding may be small.[62]

A crucial consideration for all sliding models is the mechanism by which the DNA binding protein distinguishes specific from nonspecific sites. von Hippel and collaborators[4] have proposed that RNA polymerase must exhibit more than one binding conformation, one for nonspecific sites and a second for specific sites. This proposal is based on the premise that sequence specific recognition must involve hydrogen bonding, while sliding can only occur in an electrostatic binding mode. This hypothesis requires that for efficient sampling of the DNA sequence the rate of interconversion of the specific and nonspecific binding confor-

mations must be comparable to the rate of sliding. Since a minimum estimate for the rate of DNA sliding is $\sim 10^3$ bp/sec, the rate of conformation change must be at least 2×10^3 transitions per second. A requirement of this type might be expected to place substantial constraints on the structure of any site specific DNA binding domain in a protein.

An alternative view of promoter location, or indeed the recognition of any specific protein binding site on DNA, depends on the observation that the external structure of the DNA double helix is not uniform, but reflects the internal relationships between the bases.[29,64] This implies that different sequences in a DNA molecule can be distinguished in either of two ways: by contacts with the base pairs or by contacts with sugar-phosphate backbones. The hydrogen bonding positions on the exposed edges of base pairs vary in a discrete and discontinuous fashion according to the sequence of nucleotides. In contrast, the conformation and spacing on the sugar-phosphate chains vary, a smoother and more continuous fashion dependent on the base pair overlap geometries in the core of the helix.[65,66] These two modes of recognition have been termed "digital" and "analog" to emphasize that one process is discrete and the other is continuous.[43]

On this model, the location of a specific recognition site is a two-step process. The protein must first find the appropriate external structure which characterizes its binding site. Such localization may involve sensing, for example, local variations in groove width or, particularly for larger proteins, the global curvature of the double helix. Whatever the precise structures comprising such a recognition site, location requires one-dimensional sliding along the double helix until the protein encounters a feature which is in close proximity to or coincides with a sequence specific binding site. At this point, the protein could adopt a sequence specific binding conformation.[43]

Such considerations predict that the ability of RNA polymerase to locate a promoter site could be strongly influenced by DNA sequences in the close vicinity of the recognized sequences. Such effects could be of two types. Thus, sequences which have a minor groove width different from that of the recognition sequence could, when immediately abutting the latter, change the groove width of the recognition sequence and thereby alter its external structure. Homopolymer runs of the type dA:dT or dG:dC with narrow and wide minor grooves, respectively, would be expected to show this effect and, indeed, their proximity to certain restriction enzyme cleavage sites substantially alters the rate of cutting.[29] A second possibility is that the presence of preferred bends in the DNA helix close to a promoter might serve to direct the polymerase to the specific recognition sequence. Such preferred bends have been located 10 to 40 bp upstream of several bacterial promoters including those for the *Bacillus subtilis spoVG* gene, the *Salmonella typhimurium hisR* gene, and the *E. coli rrnB* P_1 promoter.[44,45] In no case, however, has a specific role for such structures been established, although in all cases the region containing the bend is apparently necessary for optimal promoter activity.

This analog-digital model also predicts that agents which induce changes in the external structure of a DNA molecule might also alter the initial association constant. Although there are no direct measurements of the effect of such perturbations on this constant, it is clear that negative supercoiling, which results in site-specific transient unwinding of the DNA double helix, alters K_B, either decreasing or increasing this constant by factors of up to tenfold. The most frequently observed effect is a decrease, an increase being noted in only a few cases. As discussed in a subsequent section, these increases could result from the recognition of perturbations of the double helical structure induced by the negative supercoiling.

The sequence specific or "digital" recognition of a promoter site would be the consequence upon the initial analog recognition. The details of this recognition process are unknown, but are presumed to involve at least an interaction of the "helix-turn-helix" domain of the σ-subunit with either the -10 or -35 regions, a strong possibility being with the major groove of the latter.

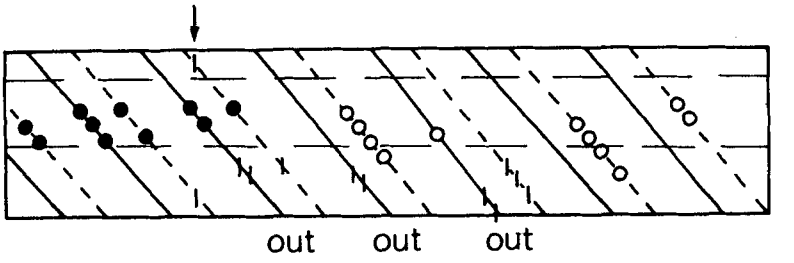

FIGURE 1. Enhanced cleavages (|) by DNAse I in the polymerase — *lacUV5* promoter complex. Drawn as cylindrical projection of DNA at 10.5 bp per turn; ● and ○ indicate phosphate contacts for RNA polymerase and CRP protein, respectively. Note that one enhanced cleavage, indicated by an arrow, is shown twice. Orientation is as for Figures 3 and 6. Transcribed and nontranscribed strands indicated by broken and continuous lines, respectively.

C. RNA Polymerase-Promoter Contacts

A necessary complement to the genetic probing of promoter structure by sequence determination and mutational analysis is the localization of the regions of close contact between RNA polymerase and its binding site. Two major types of experimental approach have been used to investigate this problem. The first, commonly known as footprinting, determines the altered reactivity of the DNA in the polymerase-promoter complex towards chemical or enzymatic modification.[67,68] In the second class of experiment, the promoter DNA is chemically modified prior to exposure to RNA polymerase to determine whether such modifications prevent the binding of the enzyme to the promoter.[69]

A general conclusion from footprint analysis, particularly using the enzyme DNAse I, is that the extent of interaction of RNA polymerase with promoter DNA is highly temperature dependent. The data set is limited to a few selected promoters, and, consequently, although some characteristics common to all promoters are apparent, it is unclear whether certain types of interaction, particularly those upstream of the −35 region, are restricted to particular classes of promoter.

At low temperatures, complexes on the fd PV111 promoter[70] and the T7 A3 promoter[71,72] have the following structural characteristics: (1) The downstream bound of the protected region extends to ~ −5, i.e., upstream of the startpoint, while the upstream bound is placed up to 2 double helical turns upstream of the −35 region at −56 for T7A3 promoter and −61 for the fd PV111 promoter; and (2) particularly in the region between −36 and −56 a periodic arrangement of enhanced DNAse I cleavage sites is observed, spaced at approximately 10-bp intervals. This regular spacing of DNAse I cleavage sites is consistent with the DNA lying on the surface of the polymerase (Figure 1). One explanation for the enhancement of cleavage relative to naked DNA is that such regions are deformed or bent so that either the conformation of the sugar phosphate backbone or the width of the minor groove is changed to a value closer to the optimum for DNAse I cleavage.[29] An alternative explanation is that in the protein-DNA complex the DNA is constrained so that thermal motions resulting in fluctuations of minor groove width are minimized, and, consequently, the probability of cleavage at preferred cleavage sites is further increased. This latter explanation appears unlikely to be correct since enhanced cleavages can occur at bonds which are poor substrates in naked DNA. This periodic pattern of enhanced cleavages in fd PV111 is maintained in polymerase-promoter complexes observed at high temperature[70] and is probably homologous to a similar pattern observed in high-temperature complexes at the *lacUV5* promoter.[73,74] The positions of these enhanced cleavages correlate well with the preferred locations of those sequences associated with an outward-facing minor groove on

Table 2
LIMITS FOR PROTECTION FOR POLYMERASE-PROMOTER COMPLEXES AT *E. COLI* PROMOTERS ARRANGED ACCORDING TO TEMPERATURE DEPENDENCE (I → III ARE IN ORDER OF INCREASING TEMPERATURE)

	Complex			
Promoter	**I**	**II**	**III**	**Ref.**
fd VIII	−61 → −45(−33)	−61 → −5	−61 → +20	70
T7 A3	—	−56 → +1	−47 → +14	71,72
lacUV5	—	—	−52 → +20	2,73,74
bla	—	—	−42 → +20	75
tyrT	—	−45 → +5	(−150) −75 → +20	76
tyrTp27	—	−150 → −5	Not observed	
unc	—	—	−75 → +20	77

Note: The different types of complex are subdivided according to the startpoint proximal limit of protection. The observed bounds are not precisely comparable between promoters because of the widely varying conditions of digestion. Note also that the temperature for the transition between different types of complex is promoter specific.

curved DNA (Section III.B), supporting the suggestion that the promoter DNA is bent at the sites of enhanced cleavage.

For one promoter, fd PV111, an additional more limited interaction of RNA polymerase with the promoter is detected at temperatures lower than that at which the −61 to −5 complex is observed. This complex is limited to the region between −61 and −45, although weak interactions may extend downstream to −33, and is detected principally by enhanced DNAse I cleavage at particular sites.

The extent of polymerase-promoter interaction observed at low temperatures correlates with the results of premodification or interference experiments in which phosphate groups are ethylated or purine residues are methylated prior to polymerase binding. Such experiments, which have been carried out principally with the *lacUV5* promoter, set the limits of residues essential for polymerase-promoter complex formation between −39 and −8.[69]

Both the protection and interference experiments thus define a polymerase-promoter complex that does not include the transcription startpoint and thus probably corresponds to the "closed" promoter complex. An increase in temperature converts such "closed" complexes to open complexes capable of rapid initiation. The principal difference between high-temperature and low-temperature complexes is an extension of the protected region from ~ −5 to ~ +20 (Table 2). In both the fd PVIII and T7 A3 promoters, this extension is correlated with minor modifications in the cleavage pattern upstream of −35.[70,71] The difference in extent of the regions protected in these two complexes suggests that the transition between them involves a substantial spatial rearrangement on the interactions within the polymerase-promoter complex possibly involving a conformational change in the enzyme.

For all the promoters discussed so far, the data are compatible with a single RNA polymerase molecule interacting with the promoter site. However, in two cases, *tyrT* and *unc*, RNA polymerase interactions have been observed substantially further upstream than −60, the upstream bounds being ~ −150 and −75, respectively.[75,76] In the former case, it has been suggested that this extended upstream protection is consequent upon more than one polymerase holoenzyme molecule binding to the promoter region. Indeed, in contrast to fd PVIII promoter protection of the upstream region of the *tyrT* promoter is only observed at high temperatures and not at low temperatures.

Analysis of the detailed pattern both of close contacts between the enzyme and the DNA

revealed by interference and protection experiments and also of the position of enhanced DNAse I cleavage sites revealing the location of exposed minor grooves, reveals that for the low-temperature complexes the enzyme is positioned primarily on one face of the DNA double helix.[53] Assuming that polymerase contacts with the −10 and −35 regions are homologous for all promoters, the extent to which this is true now depends on the separation between the −10 and −35 regions. Thus, the difference in the angular separation between these 2 boxes in the *tyrT* (16-bp separation) and the *lacUV5* (18-bp separation) is ~ 68° assuming an average screw of 10.5 base pairs per turn. This means that the relative orientations of the −35 and −10 regions in these 2 promoters must differ such that contacts are restricted to 1 face for *lacUV5*, but for *tyrT* the polymerase is wound around the promoter DNA; put another way, to bring the *lacUV5* contacts spatially into register with those on *tyrT* would require that the DNA between the −35 and −10 regions of *lacUV5* be underwound so that the average screw is ~ 11.9 to 12.5 base pairs per turn.

A possible three-dimensional interpretation of the contacts between polymerase and the *lacUV5* promoter has been proposed by Buc.[78] He has suggested that the promoter DNA is *wrapped* around part of the protein in the form of a negative superhelix. First, from the evidence cited above, it is clear that at least part of the DNA in the complex is located on the surface of the protein. Second, Chenchick et al.[79] identified by covalent cross-linking a suggestive pattern of interactions between the *lacUV5* promoter and polymerase in the extended complex at 37°. They showed that both DNA strands have multiple contacts with each of the β,β′, and σ-subunits, such that each subunit contacted the promoter at least twice. When displayed on linear map, these patches are interspersed, a pattern reminiscent of the map of the interactions of individual histones with nucleosome bound DNA.[25] A possible geometry for the polymerase interaction is the suggestion[78] that each of these polymerase subunits is equivalent to a step in the staircase contacting the DNA at two different regions. If the steps were tilted with respect to the axis of the staircase, the DNA would be wrapped as a left-handed superhelix and would thus possess a local negative writhe. The precise path of the DNA in such a structure would clearly be subject to the constraints on bendability discussed earlier.

A further feature of some, but not all, extended polymerase-promoter complexes is the presence of enhanced DNAse I cleavage sites at ~ −24 on the nontranscribed strand and ~ −26 and −27 on the transcribed strand. These enhanced cleavages are contained within the short region of the promoter where the frequency of occurrence of (dA) · (dT) blocks is relatively low and thus on the arguments advanced previously may be the site of an imposed bend in the promoter DNA. In this context, it should be noted that a point mutation at position −24 in the *gal*P_1 promoter creates a run of 6 A,T residues, lacking a TA step.[80] This mutation increases promoter activity by ~40%, an effect which the authors ascribe to either the intrinsic bend imposed by the dA:dT block being compatible with the normal direction of curvature in the complex or to the small increase in angular separation between the −10 region and the notional −35 box.[80] The former explanation is favored by the observation that substitution of either $(dT)_9 \cdot (dA)_9$ or d(ATATATATA)·d(TATATATAT) for the wild-type sequence in the λ P_{RM} spacer has little effect on promoter activity either in vivo or in vitro.[81] Furthermore, in this case the small increases in activity that are observed are virtually identical even though the two substitutions should differ in average twist. However, the same experiments showed that substitution of $(dC)_9 \cdot (dG)_9$ or $(dG)_9 \cdot (dC)_9$ in the same position reduced activity in vivo by ~40% by decreased k_f in vitro by at least twofold even though the association constant k_B is substantially increased by either substitution. This contrasts with the substitution of both d(CGCGCGCGC) and d(CTCTCTCTC) which increase activity in vivo, but have no significant effect on k_f in vitro. The authors of this study suggest that this striking effect of the substitution of a (dG) · (dC) tract is a consequence of a particular structural characteristic of this sequence. Under negative torsional

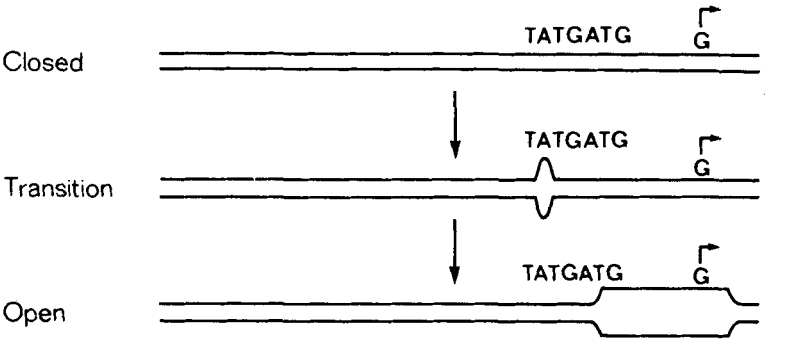

FIGURE 2. Nucleation of unwinding at the *tyrT* promoter. Arrow indicates startpoint of transcription.

stress, short runs of (dG) · (dC) become sensitive to cleavage by S1 nuclease.[82,83] However, this property is also exhibited by (dCT) · (dAG).[84,85] The authors favor the possibility that the effect of (dG) · (dC) substitution is to impede the distortion of DNA in the spacer region that must occur in the transition from the closed to the open complex. This explanation is consistent with the occurrence of mixed asymmetric polypurine-polypyrimidine sequences, but not homopolymeric sequences in tightly bent nucleosome core DNA.[42,86,87]

D. Structural Transitions in Promoter DNA

Prior to RNA chain initiation, the product of the interaction of RNA polymerase with promoter DNA is the separation of the DNA strands over a region ~ 12 bp in extent in the region of the transcription startpoint. This implies that the net angular untwisting of the DNA duplex must amount to at least 410°, a value that is in good agreement with experimental determinations of the extent of topological unwinding during open complex formation at several different promoters.[88,89]

How is the DNA converted from a fully base-paired duplex to a state in which slightly more than one turn is no longer base paired? The extent of the region over which strand separation takes place has been estimated by chemical modification of the open complex on the *lacUV5* promoter at 37°. Such direct measurements suggest that bases in the region from ~ −9 to +3 have the reactivity characteristic of single-stranded DNA,[2] a property which is also observed for *lacUV5* complexes in vivo.[90] This melted region thus does not include the TATA sequence which has been shown in both the *tyrT* and *lacUV5* promoters to be the region of the promoter at which unstacking is most readily induced by negative superhelical stress or thermal motions.[23,48] Indeed, in the stable RNA promoters the melted region contains sequences of the type $(GC)_3$ or GCGCC which would have a substantially higher enthalpy for strand separation than that of the −10 box TATAAT (based on the data of Gotoh and Tagashira[49] and assuming equivalent influences of flanking DNA sequences ΔH for strand separation of $(GC)_3$ and TATAAT are 46.9 and 39.5 kcal/mol, respectively).

Taken together, these two sets of experimental observations suggest the hypothesis that the TATA region serves as the nucleation point for untwisting and that the untwisted region subsequently migrates to the transcription startpoint (Figure 2). This hypothesis requires that untwisting occurs prior to open complex formation and that any mutation which impedes untwisting may trap the polymerase-promoter complex at an early step in the pathway to transcription initiation. Evidence for topological unwinding, as distinct from untwisting, prior to open complex formation depends on the observation that increases in negative superhelicity increase K_B for complex formation at the *lacUV5* promoter. From the magnitude of the change in K_B, Buc and collaborators have calculated that a large unwinding ($\Delta\theta$ = 468° ± 108°) is already apparent in the closed complex.[78] Passage from this complex to

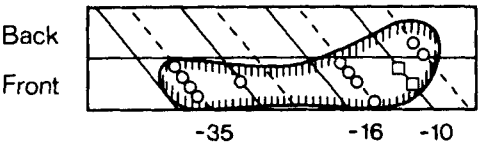

FIGURE 3. Location of TA doublets (□) on transcribed strand (---) relative to polymerase positioning in closed complex.

the intermediate complex is associated with a significant rewinding, while the final transition to the open complex requires a further small unwinding of ~72°. It should, however, be emphasized that the responses of different promoters to supercoiling in respect of K_B and k_2 vary widely. For example, at low temperatures, negative supercoiling increases K_B for complex formation at the *tet* promoter; at high temperatures the converse is true.[9,92] Thus, the topological changes calculated for the *lacUV5* promoter may be a particular characteristic of that promoter.

A second line of evidence consistent with the view that nucleation of untwisting occurs at the TATA sequence is based on the functional consequences of the *tyrTp27* mutation, a T → A transversion converting the *tyrT* −10 region from TATGAT to TATGAA.[52] This mutation results in at least a 50-fold drop in the rate of initiation both in vivo[52] and in vitro.[76] This phenotype may, at least in part, be a direct consequence of a change in the properties of the DNA in the absence of any interactions with RNA polymerase.[23] When unwinding of promoter DNA alone in the supercoiled state is assayed by S1 nuclease, the extent of cleavage in the −10 region is reduced at least tenfold by the mutation.[23] Furthermore, protection of the promoter by the polymerase does not extend beyond −5 even at high temperatures.[76] The clear inference is that the T → A transversion blocks polymerase extension and the polymerase-promoter complex remains in the closed state. This block could result from the lack of a crucial contact with the most highly conserved position of −10 region or alternatively from failure to untwist the −10 region. A corollary of the latter explanation is that in the *tyrT* promoter untwisting must either precede or proceed simultaneously with polymerase extension. The spatial relationship of the −10 region and the polymerase in the closed complex shows that the phosphate backbone of the transcribed strand in the vicinity of the TATA sequence is in close contact with the polymerase (Figure 3). A possible reaction sequence is thus that untwisting of this region allows the DNA to form additional or altered contacts with the polymerase, an event which in turn triggers a conformational change resulting in polymerase extension. Since untwisting by itself need not involve strand separation, we need to ask at what stage of the initiation process does strand separation first become detectable. This point has been addressed by studies on the reactivity of cytosine residues with dimethylsulfate in the complex of polymerase with the *lacUV5* promoter.[2] These experiments show that cytosines close to the transcription startpoint are methylated at 37°C, i.e., in the open complex, but not at 17°C in the intermediate complex. The authors suggest from this data that strand separation occurs between 17 and 37°C on formation of the open complex. In both the intermediate and open complexes on *lacUV5* DNA, the distal limit of protection by RNA polymerase is position ~ +20, but the extent of interaction of the enzyme with this promoter in the initial complex has yet to be established and is not necessarily the same as that observed for the *tyrT* promoter. In a related series of experiments, Straney and Crothers[93] separated two "open" transcription complexes at the *lacUV5* promoter by native gel electrophoresis. These complexes were interconvertible, one predominating at 16°C and the other at higher temperatures. Both complexes protected DNA from DNAse I digestion to position +20 and both were transcriptionally active in vitro, although a tendency for abortive initiation producing a short transcript was much greater for the low-temperature form. The relationship of these com-

plexes to those described by others is not established. In particular it is unclear whether the low-temperature "open" complex of Straney and Crothers[93] is identical to the 17°C complex of Spassky et al.[2] or whether it represents an additional type of complex.

What is the molecular manifestation of the large topological unwinding occurring during closed complex formation at the *lacUV5* promoter? In principle this unwinding could be distributed between twist and writhe, the precise contributions of each depending on the geometry of the complex. At one extreme, the topological unwinding could take the form of "wrapping" the DNA in a tight left-handed supercoil about the polymerase molecule. Another possibility would be that the unwinding of 1.3 turns could be principally accommodated by untwisting, distributed either across the whole polymerase binding site or locally in the −10 region. These alternatives would require a reduction in the local twist to 18° in the former case or the formation of a short stretch of left-handed DNA 6 to 8 bp in extent in the latter case.

E. Kinetic and Thermodynamic Studies

The structural studies outlined in the previous section indicate the normal sequence of events in the process of RNA chain initiation. However, transcription initiation is a dynamic process and it is its overall rate that determines the relative extent of the transcription of different genes in vivo.[5,78] To study this aspect, McClure[57] has devised an assay procedure, the "abortive initiation" assay, from which both the initial binding constant, K_B, and rate of conversion of the initial complex to the open complex, k_2 (recently redefined as k_f), can be determined. The second-order constant, K_Bk_2, is then a measure of promoter strength in vitro, which under a limited range of conditions can predict relative promoter strengths in vivo.[50]

Detailed analysis of this type has so far been confined to three promoters, *lacUV5*,[1,2] λ P_R,[3] and T7A1.[60,61] The general conclusion from all these studies is that the bipartite mechanism of transcription initiation is insufficient to explain the kinetic data and instead it is necessary to propose at least a three-step mechanism involving an initial binding step, followed by nucleation, and, finally, by strand separation.[1-3] A major problem in assessing the data for individual promoters is that assay conditions are variable and it is by no means clear that different sets of data are directly comparable. However, where the kinetic data are directly related to structural studies, as is the case for *lacUV5*, the two sets are mutually consistent.

For both *lacUV5* and λ P_R, the rate-limiting step in open complex formation is determined by the temperature. At high temperatures (>25°C for *lacUV5* and >15°C for λ P_R), this step is the conversion of the closed to the intermediate complex, while at lower temperatures the conversion of the intermediate to the open complex becomes rate limiting. For *lacUV5* this latter kinetic step has been identified with strand separation in the region of the startpoint of transcription.[2]

The major difference in the parameters affecting open complex formation at the *lacUV5* and λ P_R promoters is the temperature dependence of the different steps. At *lacUV5* the kinetic data show that the first and third steps are temperature dependent, but that the second is temperature independent. In contrast, at P_R, the first step is temperature independent and the latter two steps are temperature dependent. There is thus clear evidence that the final step, strand separation, is temperature dependent. The differences between the first two steps are crucial to the mechanistic interpretation. For *lacUV5* the temperature dependence of closed complex formation is a reflection of DNA unwinding, for λ P_R, and T7A1, the temperature dependence of the transition from the closed to the intermediate complex is interpreted as a conformational change in RNA polymerase which proceeds concurrently with the nucleation of unwinding. Two lines of evidence are adduced to support this conclusion.[35] First, the change in heat capacity $\triangle C_p^\circ$ is large and negative in the direction of

Table 3
REACTION SEQUENCE LEADING TO OPEN COMPLEX FORMATION DEFINED KINETICALLY AT DIFFERENT PROMOTERS[1-3,60,61,91,92]

	$E + P$	$\rightleftharpoons$	EP_c	$\rightleftharpoons$	EP_i	$\rightleftharpoons EP_o$
Salt dependence		+ λP_R, T7A1		o		+
Temperature dependence		+ *lacUV5* o λP_R + *tet* (relaxed) − *tet* (supercoiled)		o *lacUV5* o λP_R		+ *lacUV5* + λP_R
Rate-limiting step				>23° *lacUV5* >15° λP_R		<23° *lacUV5* <15° λP_R
Dependence on supercoiling		+ + *lacUV5* + or − *tet*		o *lacUV5* o *tet*		+ *lacUV5*

open complex formation and is interpreted as consistent with a protein conformational change involving a reduction in the amount of hydrophobic surface exposed to water. Second, this conversion is the least salt-sensitive step in open complex formation and thus is assumed not to involve a major change in protein-DNA ionic interactions. The third argument that the corresponding step in the *lacUV5* promoter is relatively insensitive to supercoiling may not be relevant to λP_R. This is because open complex formation at *lacUV5* does not exhibit the same temperature dependence at this second step, and, consequently, may not undergo a comparable conformational change at this stage.[13]

F. Effects of Negative Superhelicity

A key difference between the proposed mechanisms of open complex formation at *lacUV5* and λP_R is thus that for *lacUV5* nucleation of untwisting is thought to occur during closed complex formation, but for λP_R the nucleation event is associated with the transition from the closed to the intermediate complex. In this respect, λP_R may be comparable to the wild-type *lac* promoter. The nucleation event, require an unwinding of the promoter DNA, would be expected to be sensitive to the superhelical density of the template, as would the final unwinding of promoter DNA leading to open complex formation. However, since this latter step is not rate limiting under certain conditions, enhancement of this step by negative torsional stress would not necessarily increase k_2. Studies of the *lac*P_s and eubacterial-like chloroplast promoters[94,95] show that the *overall* activity of a promoter (i.e., $\sim K_B k_2$) increases with superhelical density until a maximum value is reached. In the case of the *lac*P_s promoter, the increase in the rate of transcription initiation is proportional to the square of the superhelical density, θ, and thus to the energy required for supercoiling. After the maximum is attained, the rate of initiation either plateaus or falls, often dramatically, with further increases in superhelical density. The precise value of θ at which the maximum rate occurs is characteristic of individual promoters.

Any differences in the step at which nucleation occurs should thus be reflected in the effect of negative supercoiling on K_B and k_2. The available data (Table 4) on such effects show first that K_B and k_2 (as determined assuming a bipartite mechanism) are apparently linked. Thus, in all cases where negative supercoiling results in a significant increase in K_B, k_2 is either unaffected or reduced by up to sevenfold (for *lacUV5*). Conversely when K_B is unaffected or is significantly reduced, k_2 is increased by factors ranging from $\sim$ 2 to $\sim$ 100. The other conclusion from this set of data is that the effect of negative supercoiling is not necessarily dependent on the precise sequence of a promoter, but rather is dependent on the precise conditions under which the assay is performed. The most striking example is the *tet* promoter for which negative torsional stress either increases or decreases K_B depending on the temperature.[91,92]

Table 4
EFFECT OF NEGATIVE SUPERCOILING ON K_B, k_2, AND K_Bk_2 AT VARIOUS PROMOTERS

	K_B ($\times 10^{-7} M^{-1}$)		k_2 ($\times 10^2$ sec^{-1})		K_Bk_2 ($\times 10^5 M^{-1}$ sec^{-1})		
Promoter	**L**	**SC**	**L**	**S**	**L**	**S**	**Ref.**
TAC 16 (AUU)	29.0	26.0	6.0	20.0	174.0	520.0	96
TAC 17	196.0	82.0	4.3	9.0	840.0	740.0	96
TAC 18	40.0	17.0	8.7	2.7	350.0	461.0	96
TAC 16 (UGU)	28.0	250.0*	4.3	1.5	120.0	375.0	96
lacUV5	0.9	15.0*	11.0	1.6	10.0	24.0	97
lac wt	1.5	0.9.0	0.3	2.5	0.45	2.3	97
tet 25°	1.4	8.5*	0.55	0.54	0.8	4.6	90,91
29°	2.1	7.1*	0.85	0.78	1.8	5.4	90,91
tet 32°	3.3	6.0*	1.1	1.0	3.6	6.0	90,91
tet 37°	8.6	1.3	1.3	2.21	11.2	2.9	90,91
tet 37°	2.9	0.24	2.5	33.0	7.2	7.9	98
bla (GA)	9.9	1.5	0.8	1.9	7.9	2.9	98
bla (ACAA)	20.0	2.9	0.6	2.9	12.0	8.4	98
RNA1	5.6	0.46	1.0	17.0	5.6	7.8	98
	16.0	17.0	0.5	49.0	8.0	833.0	99

Note: * indicates cases where negative superhelicity increases K_B. Values listed are obtained under experimentally diverse conditions and are only directly comparable when taken from a common source.

What is the molecular mechanism that is responsible for these different effects of supercoiling on K_B? The differences in K_B must clearly reflect changes in the initial interaction between RNA polymerase and the template. In terms of DNA structure, the introduction of negative superhelical stress would be expected to affect promoter structure in two ways. First, untwisting would reduce the total angular separation between the -35 and -10 regions, the extent of the reduction being dependent on the precise sequence in the intervening region and the amount by which the topological unwinding was distributed between twist and writhe. Second, untwisting would induce local structural distortions, particularly in the -10 region.[23] Again, such distortions would be strongly dependent on the precise sequences involved. Two lines of argument suggest that at least the latter change may be responsible for the observed variations in K_B. A systematic study of the effects of supercoiling on TAC promoters with spacer lengths of 16, 17, and 18 bp showed only small differences in the extent of change of K_B and k_2 with spacer length,[96] although it could be that the optimal -10 and -35 sequences in TAC promoters would minimize spacer effects. In all these cases, the distance of the startpoint from the -10 region was unchanged. By contrast, a second startpoint in the TAC16 promoter 3 bp downstream from the normal startpoint showed essentially the opposite response to supercoiling from the other TAC promoters. The authors of the study suggest that this alternative startpoint results from the use of a different -10 region 2 bp downstream from the normal -10 region.[96] The consequence of such a shift would be to site the sequence most susceptible to unwinding, in this case TATA, in a slightly different position relative to the polymerase and moreover a position that would be structurally equivalent to the most unwound region of the *tyrT* promoter. A similar conclusion is apparent from the comparison of the effects of supercoiling on the wild-type *lac* and mutant *lacUV5*

promoters which differ only in the −10 region. The difference in the enthalpy of strand separation for the 2 sequences, TATGTT and TATAAT, respectively, is ~ 1.6 kcal/mol.[49] The extent of untwisting at the −10 region consequent upon supercoiling would thus be expected to be greater for *lacUV5* than for the wild-type *lac* promoter. However, since the −10 and −16 regions of the TAC promoter are identical to *lacUV5*,[100] the effect of supercoiling on K_B cannot be solely related to the extent of untwisting at −10, but rather must also be dependent on other characteristics of the promoter.

Of particular interest is the observation that the dependence of K_B on temperature for the *tet* promoter shows a reciprocal relationship between the relaxed and supercoiled states of the templates.[90,91] For both, a substantial change takes place over the same narrow temperature range (32 to 37°C). This result suggests that the differences in the response of K_B and k_2 to supercoiling at the *tet* promoter are a consequence of a cooperative structural change in either the polymerase or the template. These mechanistic differences are probably intimately related to a set of phenomena not normally discussed in this context. These phenomena constitute "the optimum problem" and are, the author believes, central to the elucidation of the regulation and mechanism of eubacterial transcription initiation.

G. Ion Effects and the Optimum Problem

The interaction of RNA polymerase with promoter DNA, in common with the interaction of other DNA binding proteins with their respective target sites, involves the uptake or displacement of ions from both the protein and the DNA components of the reaction.[101] Such effects may be direct when the ions normally constitute part of the binding sites or indirect when the particular structure of either component is dependent on its ionic environment. The experimental problem is to distinguish between direct and indirect effects.

A comprehensive study by Record and co-workers shows that direct effects can account in large part for the general observation that high salt concentrations inhibit promoter specific initiation, anion release occurring both on initial binding and on open complex formation.[101] The general reaction is

$$R^{n+} X_n^- + P^{n-} \leftrightharpoons RP_o + nX^- \qquad \text{(Reference 101)}$$

in which RNA polymerase acts as a polycation associated with ~ 10 to 12 displaceable monovalent anions. This anion release dependent on promoter binding and opening implies an entropic activation. This effect coupled with the high positive enthalpy of limited DNA melting and possible conformational transitions in RNA polymerase is sufficient to explain the substantial enhancement of open complex formation of elevated temperatures.

In accord with this model, Shaner et al.[101] show that open complex formation is strongly dependent on the nature of the anion present such that the effect of a given anion can be directly related to its position in the Hofmeister series. It would also be expected that complex formation should be cation dependent. In particular, Mg^{2+}, which has a higher affinity for the DNA phosphate backbone than Na^+ or K^+, should compete with RNA polymerase for phosphate binding sites. The available data suggest that the dominant effect of Mg^{2+} ions is indeed a competitor for DNA binding. However, there are small but systematic deviations of the observed effects of Mg^{2+} from the predicted effects (Figures 6 and 7 in Reference 101). These deviations suggest that the relative effects of monovalent and divalent cations on polymerase binding are not wholly understood.

A further indication that the ionic interactions involved in transcription initiation are more complex is that the model in its simplest form fails to explain certain observed variations of promoter specific initiation with temperature and ionic composition. In particular, both variables are characterized *experimentally* by an *optimum* value at which initiation is most

efficient. Typically at high temperatures or low ionic strength, both the extent and the rate of initiation are substantially reduced relative to the maximum observed value.[102-104] This behavior is not predicted and is indeed contrary to the standard formulation of the initiation reaction and thus must arise from some additional cause. The major question to be addressed is therefore whether any such additional cause is biologically relevant or is simply a trivial artifact of in vitro systems.

The principal observations pertinent to this problem are

1. Observed salt and temperature optima for transcription are promoter dependent. Thus, the *lacUV5* promoter typically has a higher salt optimum than, for example, the *tyrT* promoter.[105]
2. For a given promoter, the temperature and salt optima are interdependent. In general, increasing the ionic strength raises the temperature optimum and vice versa.
3. The optimum conditions for initiation depend strongly on the conformation of the DNA template. Thus, the introduction of negative superhelical stress both raises the salt optimum and lowers the temperature optimum. Similarly, agents such as glycerol, which lower water activity and promote DNA unwinding, affect the optima in an analogous manner to negative torsional stress.[106,107]
4. Optima are substantially influenced by low molecular weight effector molecules such as ppGpp which interact with RNA polymerase and directly affect its affinity for promoter sites.
5. The interaction of RNA polymerase with the *tyrT* promoter has a bimodal character in certain experimental systems. In particular at a given ionic strength, increasing ppGpp concentration defines two separate optimal conditions for efficient polymerase-promoter complex formation.[108] Furthermore, this bimodality is itself dependent on ionic strength.
6. For some promoters, increasing RNA polymerase concentration above a certain value results in a decrease in transcriptional activity.[109]

Taken together, these observations strongly implicate the interaction of RNA polymerase with the DNA template as a major determinant of the position of an optimum. How then could such a phenomenon arise? One possibility is that the occurrence of an optimum is a trivial consequence of polymerase denaturation or aggregation at high temperatures and/or low salt. This explanation appears unlikely to be correct. In the case of the *tyrT* promoter, a progressive lowering of the KCl concentration from 0.15 *M* reduces the extent of transcription from a supercoiled template, but at the same time greatly increases transcription from a relaxed template.[104] Again, some temperature optima are as low as 25 to 30°C,[102] temperatures at which general denaturation is improbable. A second possible explanation for the occurrence of optima is that the optimum is generated by enhanced nonspecific interactions of RNA polymerase with nonpromoter DNA at high temperatures or low ionic strength. Again for the *tyrT* promoter, which has a particularly sharp salt optimum when relaxed,[104] this explanation is not valid. Thus, the footprint of polymerase bound to this promoter depends on salt concentration in a very similar manner to transcription initiation.[110] In the presence of a 100-fold molar excess of polymerase, the enzyme neither initiates nor protects the promoter from DNAse I digestion at very low salt concentrations. Under the same conditions there is no indication on nonspecific polymerase interactions with DNA sequences flanking the promoter. This result suggests that the occurrence of an optimum directly reflects a variation in the specific interactions between RNA polymerase and the promoter.

This conclusion implies that either K_B or k_2, or possibly both, can show either a positive or a negative dependency on salt concentration and temperature depending on the precise

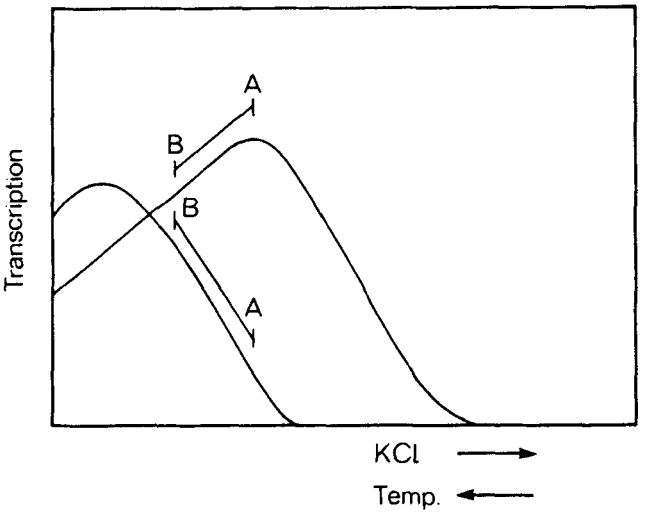

FIGURE 4. Effect of lowering salt concentration on raising temperature (A → B) on transcription in systems with high and low optima.

range of conditions assayed. In no case so far studied does k_2 behave anomalously. However, as already discussed, measurements of K_B for the *tet* promoter reveal a paradox. On a relaxed template, K_B increases fivefold over the temperature range 25 to 37°C. In contrast, when the same template is supercoiled, K_B *decreases* by six- to ninefold over the same temperature range. This apparent paradox can be simply resolved if it is assumed in this case that at 25°C the salt concentration used was above the optimum value for the linear template, but *below* the optimum value for the supercoiled template. (It is well established that supercoiling can raise the salt optimum substantially, see e.g., Reference 104.) Since salt and temperature optima are interdependent, a rise in temperature is equivalent (approximately) to a lowering of the salt concentration. On this argument, the effect of raising the temperature will be to move closer to the optimum for the relaxed template, but further away from the optimum for the supercoiled template (Figure 4). The same data also show that, as would be expected, the temperature optimum for K_Bk_2 differs from that for K_B alone.

The existence of temperature and salt optima has substantial experimental and mechanistic implications. First, it is clear that use of standard assay conditions for all promoters without reference to the optimum conditions for the particular promoter under study could result in different promoters showing radically different dependencies of, for example, K_B on temperature. As a consequence, the process of open complex formation at these promoters might be judged to follow different pathways. Second, the available data suggest that in certain cases optima may arise primarily from variation in K_B.

H. The Regulation of Promoter Selection by ppGpp

In *E. coli* promoter utilization is globally regulated,[111] the activity of particular classes of promoter, such as stable RNA promoters, being regulated in concert. The archetype of this regulatory mode is the stringent response.[112,113] When bacterial cells are starved for an amino acid, the rate of synthesis of rRNA and tRNA together with the mRNA for ribosomal proteins falls abruptly by some tenfold. In contrast, the rate of total mRNA synthesis is only modestly reduced after starvation, while that of certain species, e.g., *his* mRNA and *trp* mRNA, is actually stimulated. There is now substantial evidence that a major effector of this selective regulation is the nucleotide ppGpp which rapidly accumulates to millimolar levels after amino acid starvation. This accumulation is strongly correlated with the reduction

in stable RNA synthesis. Both genetic and biochemical studies indicate that a major target of ppGpp is RNA polymerase itself and that the regulated process is the initiation of transcription.[112,113]

Mutations in the *rpoB* gene can render stable RNA synthesis either more or less sensitive to ppGpp accumulation in vivo.[114,115] In the latter case, the mutant polymerases also synthesize elevated levels of stable RNA in vitro at physiological ppGpp concentrations.[116] Similarly, RNA polymerase containing a mutant σ-subunit shows an altered response to ppGpp in vitro,[117] an effect which is quantitatively reproduced in vivo.[118] These data confirm that the interactions of ppGpp with RNA polymerase observed in vitro are at least qualitatively relevant to the physiology of the living cell.

The strong inhibition of transcription induced by ppGpp is promoter specific, being exhibited only by stable RNA promoters and those directing the synthesis of r-protein mRNA. Those promoters that are sensitive to ppGpp in vitro and are also preferentially inactivated by an amino acid starvation in vivo have sequence characteristics normally associated with weak promoters. Thus, the −35 box deviates from the canonical sequence TTGA (the most highly conserved base pairs) and many promoters in this class, including those supporting the highest rates of initiation in vivo, have a suboptimal separation of 16 bp between the −35 and −10 bases.[112] An additional prominent conserved feature is a GC-rich sequence, the "discriminator", placed immediately downstream of the −10 box.[119,120] In contrast, the structure of the −10 region itself does not appear to correlate with stringent regulation and in some promoters, e.g., $rrnP_1$, is an exact fit to the consensus. In the *tyrT* promoter a 4-bp mutation[121] in the discriminator region which changes the sequences from GCGCCCCG*CT to GCGTTAAG*CT (* indicates transcription startpoint) abolishes both the stringent regulation of *tyrT* transcription in vivo[122] and the inhibition by ppGpp in vitro.[123] This shows that at least for *tyrTp* the GC-rich discriminator is a necessary determinant for stringency. Similar in vitro studies on the discriminator of *thrU-tufB* promoter, GCGCGCTA*CT, show that mutation of any one of GC base pairs in positions −7 to −4 to an AT base pair is sufficient to abolish ppGpp sensitivity.[124] However, in this case, comparable in vivo studies to assay the effect of the mutations on the stringent response are not available.

The correlation of the in vivo response to stringency and the in vitro response to ppGpp is of some importance since contradictory results have been obtained in the past. Thus, transcription from the *rrnAB* P_2 promoter whose activity is not significantly reduced by amino acid starvation in vivo[125] has been shown to be sensitive to ppGpp in vitro.[126,127] Furthermore, base substitutions analogous to those in the *tyrT* and *thrU-tufB* discriminators reduce this inhibition substantially.[104] Again, transcription from another starvation insensitive promoter, *lacUV5*, is inhibited in vitro by high yet physiological concentrations of ppGpp.[128] It is clear therefore that the response of promoters to ppGpp in vitro is sequence dependent, but that the precise relationship to the in vivo response must be carefully controlled relative to a standard nominally insensitive promoter.[128]

At what step in the initiation process does ppGpp act? There is as yet little known about the detailed mechanism of open complex formation of stable RNA promoters. Indeed, there is no evidence that strand separation has actually occurred in complexes described as "open" on the basis of stability or temperature dependence. It is, however, clear from footprinting experiments that such complexes are relatively unstable (in the absence of ppGpp) when compared with polymerase-promoter complexes formed under comparable conditions at the *lacUV5* promoter. Thus, the residence time of RNA polymerase on the *tyrT* or *rrnA* P_1 promoter under optimal conditions is ~ 10 to 30 sec.[76,129] In the presence of ppGpp, no polymerase footprint is observed on either the wild-type *tyrT* promoter or on the mutant *tyrTp27* promoter which fails to permit polymerase extension and thus presumably open complex formation.[76] This result indicates that ppGpp can act at a step prior to polymerase

extension and thus must presumably directly affect the rate of formation or stability of the closed complex.

This conclusion raises the paradox that sequences known to be involved in the control of initiation by ppGpp lie downstream of those sequences known to be covered by RNA polymerase in the initial complex. This conclusion also fails to explain why ppGpp *increases* transcription from both the *tyrT* SSU2 discriminator mutant and the *lacUV5* promoter in vitro. However, although transcription from the wild-type *tyrT* and the SSU2 mutant promoters responds in opposite ways to ppGpp, "stable" complex formation on both promoters shows an essentially identical response over the concentration range 0 to 1 m*M* ppGpp.[108] This result clearly suggests that ppGpp may affect more than one step in the initiation process and that the sequence in the discriminator region may be crucial in the transition from the closed to open complex. In this context, it should be noted that in the *lacUV5* promoter the nontranscribed strand of this region is in close proximity to the σ-subunit.[130] At this juncture, it seems clear that the elucidation of the effects of ppGpp on polymerase-promoter interactions requires a detailed study of the mechanism of complex formation at stable RNA promoters as well as a comparison of the effects of ppGpp on different promoters.

More generally, the effects of ppGpp on transcription initiation are dependent on other factors which facilitate or inhibit open complex formation. Thus, agents which promote DNA unwinding such as glycerol[46] or negative superhelicity[29] decrease the sensitivity of stable RNA promoters to ppGpp. Similarly, high polymerase concentrations overcome the inhibitory effect of ppGpp.[131] Related to these effects ppGpp lowers the KCl optimum[132,133] and raises the temperature optimum[133,134] for transcription. This has the consequence that for a given promoter the sensitivity to ppGpp is a function of KCl and temperature, the greatest inhibitory effects for low ppGpp concentration being observed at low temperatures or high KCl concentrations. In contrast, transcription at low magnesium concentration is more sensitive to ppGpp than at higher concentrations.[135] A further indication of the link between the optimum phenomenon and the action of ppGpp is apparent from a study of "stable" complex formation at the *tyrT* promoter.[108] Measuring stable complex formation by filtration, it was observed that low concentrations of ppGpp (80 to 100 μ*M*) strongly inhibited binding. However, further increase in ppGpp concentration resulted in an increase in binding with a maximum being attained at 200 μ*M*. At ppGpp concentrations >200 μ*M*, binding was again inhibited. Transcription from the *tyrT* promoter also exhibited this bimodal response to ppGpp without any change in the transcription startpoint. Thus, the binding curve must reflect changes in the polymerase-promoter interaction. The nature of these changes is indicated by a difference in the characteristics of the dependence binding on enzyme concentrations. Thus, the concentration of polymerase required for half maximal retention at the maxima is ~ five- to eightfold less than that at the minima. A further characteristic of the bimodal binding response is its dependence on KCl concentration. As the KCl concentration is raised, the positions of both maxima shift to lower ppGpp concentrations.

Taken together, these experiments show that the apparent affinity of RNA polymerase for the *tyrT* promoter can change in a bimodal fashion as a function of ppGpp concentration and at a given ppGpp concentration a similar KCl-dependent variation is observed. The data do not distinguish between variation in the binding constant K_B or the second-order constant $K_B k_2$.

What is the molecular basis for this functional variation? Physical studies on the interaction of ppGpp with RNA polymerase holoenzyme show that the nucleotide reduces the sedimentation coefficient from ~ 13.5 to 14 S to ~ 12.5 S under conditions where the enzyme normally sediments as a monomer.[131] In addition, the nucleotide quenches the tryptophan fluorescence of the enzyme in a similar manner to ATP and GTP.[136,137] The extent of this quenching is again KCl dependent being observed at 50 m*M* KCl, but not at 100 m*M*. At

this latter salt concentration, the fluorescence in the absence of the nucleotide is similar to that in the presence of the nucleotide at the lower salt concentration.[137] Although the nucleotide fails to quench at 100 m*M*, KCl equilibrium dialysis measurements indicate that binding of ppGpp to the enzyme is not decreased. Both of these sets of observations are consistent with the hypothesis that ppGpp induces a substantial structural change in RNA polymerase. They further suggest that this same structural transition can be triggered by a change in the ionic environment. Direct evidence for one salt-dependent conformational transition of the free enzyme has recently been obtained.[138] This transition is anion-linked and facilitates the dimerization of the polymerase holoenzyme. However, dimerization per se is not affected by physiological concentrations of ppGpp[137] and therefore may not be related to the changes in polymerase function induced by the nucleotide. An alternative would be a cation-linked conformational change, since both ppGpp sensitivity and holoenzyme structure[135,139] are known to be correlated with Mg^{2+} concentration.

One possible interpretation of both the functional and structural data is that there exist at least two structural forms of RNA polymerase which interact strongly with the *tyrT* promoter and at least two others which interact weakly. The strong and weak interacting forms would correspond to the maxima and minima of the bimodal binding curve. Independent evidence for the evidence of such forms has been obtained from zone sedimentation of purified polymerase holoenzyme.[140] This procedure effected the partial separation of two enzyme fractions, a fast and a slow sedimenting form, with the ability to transcribe *tyrT* efficiently. There is no direct evidence linking the binding data and the sedimentation data. However, if the observations are related, the most probable correlation is that the first maximum in the binding curve corresponds to a fast-sedimenting form of the enzyme, while the second maximum corresponds to a slow-sedimenting form since ppGpp favors conversion to a slow-sedimenting form.

The clear inference from the available data is thus that the existence of optimal conditions for transcription initiation and regulation of initiation by ppGpp are facets of a single phenomenon. A probable explanation is that both observations reflect structural transitions between different functional forms of RNA polymerase holoenzyme.

I. Mechanistic Models for Transcription Initiation

In the open complex, the close contacts of RNA polymerase holoenzyme with promoters of disparate function are apparently homologous.[69] Yet, in contrast, kinetic and thermodynamic studies on different promoters reveal substantial apparent differences in the pathway to open complex formation. These differences can be understood on the premise that this process involves structural transitions in both the DNA template and RNA polymerase and that these transitions can occur to a certain extent independently of any direct interaction between the enzyme and its substrate.

For promoter DNA, the author proposes that in the transcription complex untwisting is initiated at the TATA sequence and that this region of untwisted DNA subsequently migrates to the transcription startpoint where strand separation takes place. In the absence of polymerase, untwisting can also occur preferentially at the TATA sequence, a process which is favored by very low ionic strength or by negative supercoiling.[23] For RNA polymerase, two states of the enzyme in the transcription complex can be broadly distinguished, a "compact" form which contacts only sequences upstream of the transcription startpoint and an "extended" form which contacts sequences both upstream and downstream of the startpoint. When observed, these forms are associated with the closed and the intermediate/open complexes, respectively. The free holoenzyme can also exist in different structural states. In addition to its ability to dimerize,[141] two forms of the monomer have been proposed to exist, a fast- (~ 14 S) and a slow-sedimenting form (~ 13 S). The relative proportions of these two forms can be shifted to favor the fast form by adenine nucleotides or the slow form by guanine nucleotides.[146]

To what extent can these two sets of observations be related? First, the in vitro conditions which favor the transcription of negatively supercoiled DNA relative to relaxed DNA, i.e., high KCl, low $MgCl_2$, and low temperatures, are the same conditions that favor ppGpp sensitive transcription, i.e., by the slow-sedimenting form of polymerase hoenzyme. This suggests that the different structural forms of free polymerase can discriminate between different conformations of the promoter. In one configuration, the DNA would be an unstressed double helix and would be the preferred template for the fast-sedimenting form of holoenzyme; in an alternative configuration, the promoter DNA would be partially untwisted and be the preferred template for the slow-sedimenting form. Such an interpretation is fully consistent with the observation that K_B for initial complex formation can be either decreased or increased by increasing negative superhelicity. Also of possible relevance is the observation that the *tyrT*p27 mutation hinders both the untwisting of the TATA region in the absence of polymerase and also the transition of polymerase from the compact to the extended form in the transcription complex.

Taken together with the kinetic and thermodynamic data for different promoters, these considerations suggest that there may be at least two mechanistically distinguishable pathways to open complex formations exemplified by the *lacUV5* and λ P_R promoters, respectively, (Figure 5). In the first, RNA polymerase would interact with promoter DNA that was already partially unwound in the −10 region to form an initial complex. This state is then succeeded by an intermediate complex in which the DNA remains unwound, but no strand separation is detectable. In this complex the polymerase covers 60 to 70 bp from positions +20 to −40 or −50. This transition corresponds to the rate-limiting step observed at high temperature for the *lacUV5* promoter. In the final step, strand separation occurs in the vicinity of the transcription startpoint, i.e., the untwisted region has migrated from the −10 box and the polymerase remains extended. This step is presumably driven by the topological unwinding previously generated and becomes rate limiting at lower temperatures. At this juncture, the polymerase becomes competent to initiate RNA synthesis.

The crucial difference between this pathway and the second is that in the latter, the initial interaction of RNA polymerase is with unstressed DNA. In this initial complex, the enzyme remains compact and the DNA is not unwound. This step is then followed by a concerted transformation to the intermediate complex in which both DNA unwinding and polymerase extension takes place so that the complex thus formed is then equivalent to that observed on the *lacUV5* pathway. The subsequent transition to the open complex would then be common to both pathways.

At present it is not known whether the observed differences in the kinetic and thermodynamic parameters for open complex formation on different promoters relate specifically to an intrinsic variability in promoter function or simply reflect assay conditions which are not functionally comparable. If, as I have suggested, the structure of the free polymerase is a determinant of the pathway to open complex formation, then appropriate variation in the reaction conditions would be expected to reveal alternative pathways to open complex formation at the same promoter.

The reaction sequence is assumed to be driven by anion release occurring at different steps triggering conformational changes in RNA polymerase. A particular aspect of this phenomenon is the proposal that K_B is a bimodal function of anion release such that above a critical value the polymerase undergoes a cooperative transition from a high affinity state to a low affinity state thereby generating the characteristic salt optimum curve.

The extent of anion release and thereby the position of the optimum would be a function of the sequence specific and conformation specific recognition of the promoter by RNA polymerase. Thus, for example, an up-mutation in the −35 region would be expected to increase the salt optimum. Similarly, if the probability of the nucleation of untwisting is increased by mutation, the salt optimum should also increase. In this context, it should be

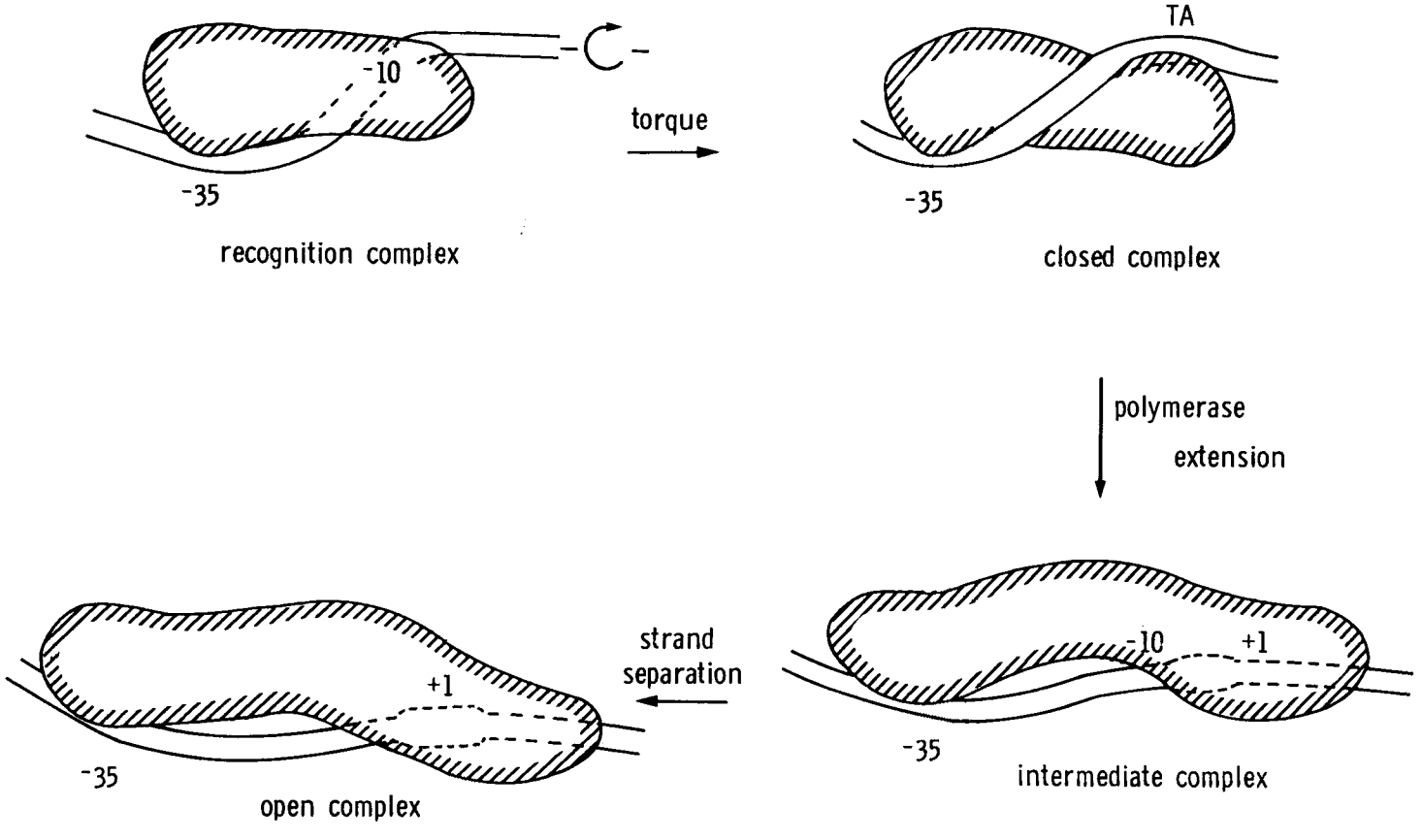

FIGURE 5. Schematic of proposed steps leading to open complex formation. The scheme shown, which is one of a number of possibilities, assumes that a change in the sense of wrapping DNA around the polymerase precedes strand separation (for discussion see text).

Table 5
HIERARCHY OF REGULATION OF TRANSCRIPTION INITIATION IN *E. COLI*

Regulatory mode	Target	Function	Examples
Substitution of RNA polymerase	—	Switch in sequence specificity of promoter recognition	Infection by phases T7, N4
Substitution of σ-factors	Core RNA polymerase	Switch in sequence specificity of promoter recognition	Heat shock, nitrogen fixation, infection by phage T4
Regulatory effectors	RNA polymerase holoenzyme	Alteration in affinity for particular classes of promoter, but retaining sequence selection	ppGpp control
		Facilitation of σ- substitution (?)	
Promoter specific control	DNA/polymerase	Positive control of initiation	CRP-cAMP, λC_I, λC_{II}, AraC RNA polymerase (?)
		Repression	LacI λC_I AraC LexA

noted that mutations in the λ P_{RE} promoter which result in the juxtaposition of a consensus −16 region with a consensus −10 region remove the functional requirement for a −35 region.[142] Mutations with a similar phenotype have also been isolated in the *galP*$_1$ promoter.[143] The sequences generated by these mutations particularly in the −10 to −16 region would be expected to be particularly susceptible to unwinding and thus might increase the rate of nucleation to the extent that specific contacts in the −35 region are no longer necessary. A second possibility, not exclusive of the first, is that the particular sequence arrangements in the mutant *galP*$_1$ and λ P_{RE} promoter allow optimal contacts of the polymerase in the −10 to −16 region which again would minimize any requirement for strong −35 contacts.

IV. REGULATION OF TRANSCRIPTION INITIATION

A. General Principles

In *E. coli*, the rate of transcription initiation at the majority of bacterial promoters is regulated in accordance with the cellular requirement for the spectrum of gene expression under the prevailing environmental conditions. In the growing bacterium, this regulation operates at two levels: first, a direct modulation of the specificity and overall affinity of RNA polymerase for particular classes of promoter affected by σ-factor substitution and by interactions with regulatory nucleotides such as ppGpp and, second, control by regulators which interact primarily with additional DNA binding sites normally in close proximity to a binding site for RNA polymerase.[144] The first regulatory mode acts as coarse global control directly affecting the activity of many promoters, whereas the latter mode, which can act either positively or negatively, functions primarily as a control on particular transcription units (Table 5).

B. Positive Control of Initiation

For many *E. coli* transcription units, initiation is positively regulated, i.e., the binding of a regulatory protein close to the transcription startpoint increases the rate of initiation. For any particular regulated system two aspects are crucial, which step leading to open complex formation is increased and what is the molecular mechanism by which this is accomplished.

As a group, positively regulated promoters are necessarily weak promoters for the interaction with RNA polymerase by itself. The sequences of these promoters normally differ significantly from the consensus configuration and often mutation to closer homology with the consensus results in a higher constitutive rate of initiation and reduction in the degree of regulation. Examples of such mutations include the *lac*p^s and *lac*UV5 mutations in the -10 region of the *lac* promoter and an A $\rightarrow$ T change in the first position of the -10 region of the *ompF* promoter.[146] However, an up-mutation in the -35 region of the λ P_{RM}, although raising the constitutive rate of transcription does not alter the degree of regulation.[59]

The extent to which such mutations result in loss of regulation depends on the precise mechanism of activation. Thus, the activation of the λ P_{RM} promoter by the C_I repressor protein proceeds by increasing k_2 with little effect on K_B.[59] Mutations in the -35 box which increase K_B thus remain susceptible to regulated enhancement of k_2. By contrast, the activation of the *lac* promoter by the CRP-cAMP is somewhat more complex. At this promoter, RNA polymerase by itself can form two alternative complexes, one at site P_1, directing transcription from the normal startpoint, and a second at partially overlapping site P_2, directing initiation in the same direction at -22. In vitro in the absence of the activator protein, RNA polymerase preferentially binds to P_2. However, in the presence of CRP-cAMP, polymerase binding at P_2 is prevented, while that occupancy at P_1 is enhanced. This enhancement results both from a direct increase in K_B (but not k_2) and from lack of exclusion by polymerase bound at P_2.[73,97]

Promoters such as *lac*P_1 and λ P_{RM} have weak yet functional -35 and -10 regions.[152] There exists, however, a class of positively regulated promoters which lack a recognizable -35 region at an appropriate distance from the -10 region. Such promoters include the λ promoters P_I and P_{RE} activated by the λ C_{II} protein[147,148] and the *gal*P_1 promoter again activated by the CRP-cAMP complex.[73] In the former case, the regulatory protein increases both K_B and k_2[148,149] and thus is mechanistically distinguishable from positive regulation at the λ P_{RM} and *lac*P_1 promoters.

What is the molecular mechanism by which activation is effected? In principle, an activator could alter the structure of the promoter DNA, by, for example, potentiating unwinding, or alternatively could directly contact RNA polymerase, thereby altering the interaction of the enzyme with the DNA template. There is now some evidence that for some regulatory systems protein-protein interactions rather than alterations in DNA structure are responsible for activation. Thus, the binding sites of the two auxiliary proteins, λ C_I and P22 C_I, which increase k_2 at their respective P_{RM} promoters are both positioned such that at least one of their phosphate contacts overlaps with a phosphate contact of the σ-subunit in the -35 region (Figure 6A).[144] However, this overlapping contact point has a different relationship to the DNA binding site of the two C_I proteins. For both λ C_I and P22 C_I, mutants exist which bind to DNA normally, but are unable to activate transcription.[144,150] This shows that the DNA binding function is separable from the positive control regulation. Further, the mutations in both cases change amino acids that are spatially distant from the DNA binding site yet have the same positioning relative to RNA polymerase. In addition, a positive control mutant of P22 C_I can be partially suppressed by a mutation in the β-subunit of RNA polymerase, again suggesting that specific protein-protein contacts are necessary for activation.[151]

Once more the mechanism of activation of *lac*P_1 by the CRP-cAMP complex shows

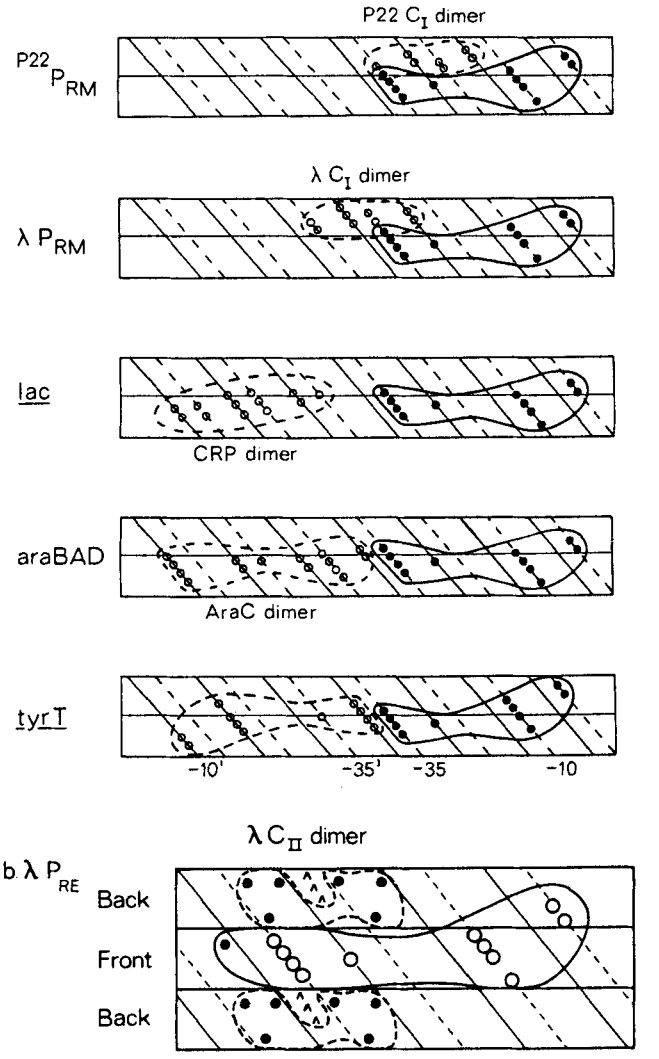

FIGURE 6. Relative orientations of RNA polymerase and activator proteins at promoter site. (A) ○ and ● indicate PO_4 contacts for activator protein and RNA polymerase, respectively; (B) ○, notional phosphate contacts for RNA polymerase; ●, guanine residues protected from methylation in polymerase - C_{II} — promoter complex; ∧, adenine methylation enhanced in C_{II} - DNA complex in absence of polymerase. Note C_{II} protein is shown twice.

additional complexities. This protein differs from the C_I proteins in binding in this case to the same face of the DNA as RNA polymerase. Again, there is evidence for direct protein-protein contacts between the activator and the polymerase. First, in solution, the two proteins interact weakly, an interaction which is enhanced by the addition of DNA.[152] Second, a mutant of CRP has been selected which binds normally at the *lac* promoter, but is unable to position RNA polymerase at *lac*P_1.[153] This mutation was targeted on the basis of assuming approximately homologous contacts between the activator and polymerase for both CRP and λ C_I protein. However, in addition to the evidence favoring protein-protein contacts, it is clear that the CRP-cAMP complex induces a local bend in the DNA.[154] This bending could allow the appropriate juxtaposition of CRP and polymerase molecules for effective contacts

and could also possibly favor the "wrapping" of DNA around the polymerase to facilitate local unwinding.[78]

In two other cases, that of the *araC* protein binding to the *araBAD* control region and the *ntrC* protein at the *glnA*, the orientation of the regulatory protein relative to the polymerase is similar to that of CRP at *lac*P_1.[12,155] For *araBAD*, assuming that the phosphate contacts for RNA polymerase in the −35 region are homologous for all promoters, the polymerase and activator protein would face each other across the minor groove immediately upstream of the −35 region (Figure 6). A similar spatial relationship could also occur between two polymerase molecules binding in opposed orientations as has been proposed for the *tyrT* promoter.[76]

In the examples cited so far, the regulatory activator increases either K_B or k_2, but not both. For the λ C_{II} protein which at λ P_{RE} has this property, the pattern of close approaches of the activator to the polymerase shows parallels with both other classes of activator (Figure 6B). If it is assumed that polymerase bound to λ P_{RE} is oriented such that the notional −35 to −10 separation is the optimal 17 bp the pattern of C_{II} contacts suggests close approaches of the 2 proteins both in the same region as that of the C_I proteins and also on the opposite side of the RNA polymerase molecule.[144] A similar pattern of close approaches would be apparent between the CRP-cAMP complex at *gal*P_1 (where the activator is placed differently from *lac*P_1) if the notional separation between the −10 and −35 regions were 16 instead of 17 bp. For both *gal*P_1 and λ P_{RE}, the precise contacts between the two bound proteins remain to be established.

In most described cases of positive transcriptional control, the regulator protein binds in close proximity to the polymerase binding site. However, recent evidence shows that the *nifA* gene product, a DNA binding protein which acts as a positive activator of promoters for genes required for nitrogen fixation and metabolism, can still stimulate transcription when placed up to 1 kb upstream of the transcription startpoint.[156-158] Thus, as in eukaryotic transcriptional regulation, both activation and repression can be mediated by proteins acting at a distance.[159,160]

The positioning of positive activator proteins relative to RNA polymerase is also relevant to the precise mechanics of DNA manipulation during open complex formation. For those proteins that increase k_2, i.e., the C_I and C_{II} proteins, as well as for CRP at *gal*P_1, their alignment overlapping the polymerase binding site from a downstream limit of at least ~ −27 must place constraints on the path of the DNA around the polymerase at the promoters unless the act of open complex formation is accompanied by the release of the regulatory protein. Whether such displacement occurs in any case remains to be established.

C. Stable RNA Promoters

The stable RNA promoters of *E. coli* are paradoxical. They are among the most active promoters in the bacterial cell, yet in vitro they behave typically as weak promoters. This latter properly is in accord with the sequences of the primary promoter site. In general, stable RNA promoters have a −35 region which deviates in at least 1 position from the consensus and also contain a GC-rich sequence in the region in which strand separation takes place. A further characteristic of the most active rRNA promoters is a suboptimal 16-bp separation between the −35 and −10 boxes.[112,161]

Deletion studies have shown that the sequences conferring high activity on stable RNA promoters lie upstream from the −35 region. In the *tyrT* promoter, removal of sequences between −98 and −40 results in a tenfold loss of promoter activity, while sequences upstream of −98 are required for maximal activity.[162] In the absence of this upstream element, the activity *tyrT* promoter is comparable to *lacUV5*. Similar data have been reported for the *rrnB* P_1 promoter and in the *hisR* promoter of *Salmonella typhimurium* a deletion of 3 bp at ~ −70 results in a threefold loss of activity.[44,45]

How do these upstream activator sequences promote efficient transcription? For both the *tyrT* and *hisR* promoters, deletion mutants in the upstream region show a mutant phenotype in in vitro transcription systems containing only RNA polymerase and the DNA template, although at least for *tyrT* the effect is less marked than the in vivo phenotype.[44,163] This result strongly suggests that RNA polymerase interacts with the upstream region, a conclusion that is confirmed by footprinting studies on the *tyrT* promoter where protection by polymerase extends to at least −150.

Within the upstream region of many stable RNA promoters are sequences corresponding to polymerase binding sites, which can be in either the same or the opposite orientation to the primary binding sites.[120] In the *tyrT* promoter, there is a secondary site in the opposite orientation with the two −35 regions separated by 8 bp. This arrangement is conserved in other *E. coli* tRNA promoters. A second feature of these upstream regions is the frequent occurrence of blocks of (dA) · (dT), at least 5 bp in extent.[120] These blocks are often positioned at −55 and are possibly responsible for the intrinsic bending noted in both the *rrnB* P_1 and *hisR* promoters.

On this evidence there are two possible models for the mechanism of activation of stable RNA promoters. The first is that the structure of DNA in the upstream region allows the template to bend around a single polymerase molecule bound at the normal promoter site and thus increase the rate of initiation. A second model is that a further polymerase molecule (or molecules) binds to the upstream region and activates transcription by the polymerase bound at primary site by protein-protein contact in a manner analogous to the λ C_I repressor or the AraC protein. The available data do not distinguish between these models. The 3-bp deletion in *hisR* at −70 both alters the intrinsic bending and changes the −10 region of the putative upstream polymerase binding site. Similarly a 2-bp substitution at −44 to −45 in the *tyrT* promoter which changes the putative upstream −35 box from TTGAGA to CCGAGA and is a strong down mutation both in vivo and in vitro[164] could reduce activity either by directly affecting polymerase recognition or by altering the bendability of the DNA around a single polymerase molecule.

The in vitro protection and functional data on the *hisR* and *tyrT* promoter imply that RNA polymerase must interact with at least 3 turns of double helix upstream of the normal −35 box. This is substantially more than the presumed interactions for a single polymerase on the *lacUV5* or fd VIII promoters. Further, it should be noted that the bending of DNA in this region of stable RNA promoters will be constrained and directed by the (dA) · (dT) blocks. There is also additional circumstantial evidence that RNA polymerase can act as a transcriptional activator. Thus, the concentration dependence of filterable polymerase-promoter complex formation at the *tyrT* promoter on a DNA fragment which terminates 56 bp upstream of the startpoint suggests that 2 polymerase molecules are necessary for the formation of such complexes.[165] In addition, a second polymerase molecule upstream of the *tyrT* promoter could potentially contact the polymerase bound at primary site across the minor groove immediately upstream of the primary −35 box. Such a contact would be analogous to that proposed between the AraC protein and RNA polymerase (Figure 6).

V. THE IN VIVO REGULATION OF POLYMERASE-PROMOTER INTERACTIONS

A. In Vivo and In Vitro Comparisons and Caveats

A necessary complement to the in vitro dissection of transcription initiation is the correlation of such functional analysis with the actual regulation of transcription in the intracellular environment. Although there is a relatively good qualitative agreement between in vitro and in vivo results, it is nonetheless apparent that most in vitro transcription systems are relatively primitive approximations to the intracellular conditions. Such approximations

are potentially capable of altering the quantitative parameters of the initiation reaction to a substantial degree. One such approximation is the ion composition of in vitro systems, which usually — and possibly inevitably — differs substantially from the intracellular composition, where anions such as glutamate and aspartate are present at significant concentrations.[166] Since such organic anions are low in the Hofmeister series, their presence is likely to directly influence the interactions of DNA binding proteins with their recognition sites. An important consideration in this context is the observation that the intracellular ionic environment of *E. coli* undergoes substantial variation as a function of growth conditions.[167,168] Such changes, in the absence of compensating mechanisms to maintain the internal ionic balance, would be expected again to affect DNA-protein interactions in general and RNA polymerase-promoter interactions in particular. In the case of the AraC DNA binding protein, it has been suggested that the effects of varying ion concentrations are minimized by the phenomenon of *ion compensation*.[169] In this process, the net release of anions on DNA binding is reduced by a compensating uptake at sites on the protein distant from the DNA binding domain. The magnitude of this uptake is itself dependent on salt concentration. The net effect is to reduce the dependence of the binding affinity on ionic strength. For RNA polymerase this problem is particularly acute, since it would be expected that the extent of open complex formation would be inversely related to the 10th to the 12th power of the anion concentration[101] over the appropriate concentration range. Such extreme sensitivity would seem inevitably to be incompatible with the maintenance of balanced transcriptional activity during small or even moderate changes in the intracellular ionic composition. One possibility, favored by this author, is in that the bimodal response of polymerase activity to ionic strength is a mechanism for ion compensation.

There are also other parameters in in vitro systems which differ from their in vivo counterparts. For example, studies on conformational transitions in DNA facilitated by negative torsional stress strongly suggest that the effective superhelical density of DNA in *E. coli* is likely to be in the range -0.024 to -0.027 rather than the value of -0.065 characteristic of most supercoiled templates used in vitro.[47,170] Similarly, the concentrations of RNA polymerase and of DNA used in in vitro systems are 10- to 100-fold lower than the apparent corresponding in vivo concentrations. Despite such gross differences, in vitro data when cautiously interpreted are both consistent with and have at times correctly anticipated conclusions based on in vivo experiments. Such qualitative agreement suggests that it is reasonable, within limits, to base models for transcriptional regulation in vivo upon in vitro results.

B. Regulation by Polymerase Availability

During normal exponential growth of *E. coli*, up to 60% of the actively transcribing polymerase molecules are engaged in the transcription of stable RNA genes. These genes are subject to at least two modes of regulation: stringent control, which is triggered by effective starvation for an amino acid, and growth rate control. The former is a stress response and results in an immediate cessation of stable RNA accumulation caused principally by a 10- to 20-fold drop in the rate of RNA chain initiation at stable RNA promoters. In contrast, the growth rate control acts to balance the production of ribosomal RNA — and hence of ribosomes — to the required cellular protein synthetic capacity. To meet this requirement, the rate of initiation at individual stable RNA promoters increases with increasing growth rate.

The stringest response is strongly correlated with the accumulation of ppGpp to millimolar levels.[171] Since RNA polymerase is one, and possibly the major, target of ppGpp, we have recently proposed that the regulation of stable RNA initiation depends primarily on polymerase availability, a quantity which is a function of both the absolute concentration of the enzyme and its binding affinity.[112] On this model, stringently controlled stable RNA pro-

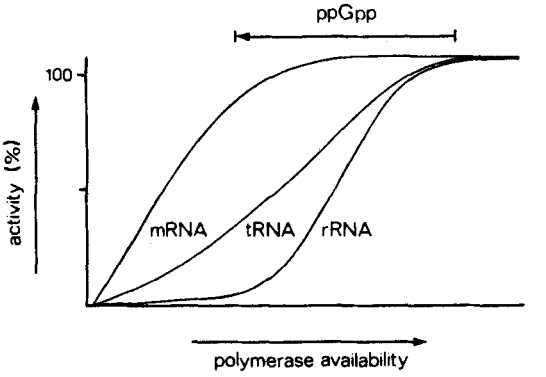

FIGURE 7. Regulation of transcription by polymerase availability. Activity for any promoter is plotted on ordinate. Effect of increasing ppGpp concentration is shown by arrow. The mRNA plot represents the response of, for example, r-protein mRNA.

moters have a low affinity for RNA polymerase, and, consequently, require high polymerase concentrations or affinities for optimal activity. Thus, amino acid starvation would selectively inhibit the synthesis of stable RNA species by reducing, possibly through the mediation of ppGpp, the affinity of polymerase for its binding sites so that those promoters for which polymerase availability is limiting would no longer support efficient initiation. Conversely, initiation would not be impaired at promoters with a high affinity for the enzyme (Figure 7).

Recent evidence suggests that the growth rate response of stable RNA genes may be explained in similar terms. In particular, a mutation in the discriminator of the *tyrT* gene which abolishes stringent control also changes the growth rate dependence of *tyrT* expression from a positive to a negative correlation.[164] This suggests that the molecular mechanisms responsible for stringent control and growth rate control share a common target. However, in addition to the discriminator, sequences in the upstream element of the *tyrT* promoter modulate the positive correlation of expression with growth rate. The only known molecule which interacts with this upstream element is RNA polymerase, again consistent with the proposal that this enzyme is the ultimate target of growth rate control. If this is the case, the problem is to explain how regulation of the enzyme can generate an apparently smooth and characteristic variation in the rate of expression of a particular stable RNA promoter. In such an analysis, it is necessary to consider:

1. The variation in the absolute concentration of RNA polymerase
2. Possibly variations in the affinity of RNA polymerase for its binding sites and whether such variations occur in discrete steps or are continuous
3. The effect of promoter competition

It is well established that the rate of synthesis of RNA polymerase is positively correlated with growth rate. Thus, the *total* number of core polymerase molecules per cell increases ~ fivefold from 1600 to 8000 molecules per cell as the growth rate increases from 0.7 to 2 generations per hour.[172] Over the same range of growth rate, cell mass increases 4.4-fold and DNA content increases about 2.3-fold.[173] This means that the concentration of polymerase varies little with growth rate, since cell mass is directly related to volume. The total concentration of polymerase per cell is thus ~ 1 to 3 μM. However, of these polymerase molecules, only a certain fraction, ~ 30%, will be complexed with the vegetative σ-factor, σ_{70}, and thus be competent for promoter recognition.

Such an estimate of 300 to 900 n*M* for the overall concentration of RNA polymerase holoenzyme in the bacterial cell is substantially higher than the estimate of the effective concentration as 20 to 40 n*M*, i.e., the concentration of polymerase available to interact with promoter sites.[174] This latter estimate is based on two arguments. First, the promoter activity determined at an in vitro polymerase concentration of 30 n*M* directly correlates with the in vivo expression of various *lac* and λ promoters present in single copy. Second, the total polymerase concentration in minicells lacking DNA is 200 to 400 n*M*. This second argument assumes that the free polymerase is equally distributed on division between minicells and DNA-containing cells. However, since RNA polymerase holoenzyme has a significant nonspecific affinity for DNA ($K_d \sim 10^{-6}$ *M*), most polymerase molecules not engaged in transcription will be in close association with DNA, and, consequently, will partition preferentially into the DNA-containing cells. It therefore follows that the cellular polymerase/DNA ratio is a more appropriate measure of effective polymerase concentration than its absolute cellular concentration (for a more extensive discussion of this point with reference to RNA polymerase and DNA binding proteins in general, see Reference 175). If the polymerase/DNA ratio is a crucial parameter, the differences in absolute concentrations between in vitro and in vivo transcription would be of less significance. More importantly, in the context of the present discussion, it is clear this ratio varies by no more than two- to threefold over a wide range of growth rates. Thus, any modulation of polymerase availability must arise primarily from changes in promoter affinity rather than from changes in absolute concentration.

The range of variations in promoter affinity that can be effected by regulators of RNA polymerase, such as ppGpp, has not been extensively quantified. The reported determinations suggest that physiological concentrations of the nucleotide reduce the apparent association constant of polymerase for promoters by ~ 100-fold. Since there is no precedent for effectors of conformational changes in proteins inducing a continuous variation in a particular functional parameter of an enzyme, we must assume that conformational changes effected by ppGpp result in discrete changes in promoter affinity. To reconcile such a discrete, and probably large, change in promoter affinity with the continuous variation in promoter activity observed in vivo, it is only necessary to assume that any equilibrium between a high-affinity state and a low-affinity state is tightly regulated so that the ratio between the two states is continuously variable.

This type of model for polymerase function makes the strong prediction that functionally distinguishable populations of polymerase molecules should be present in vivo. If this were the case, it would be expected that the ability of a promoter to compete for a particular class of polymerase molecules would be determined by sequences which influence the ppGpp dependence of promoter function. Such a differential competitive effect is indeed observed in *E. coli* transformed with plasmids containing different variants of the *tyrT* promoter.[164] Thus, transformants containing a *tyrT* promoter with a wild-type discriminator region fail to increase growth rate significantly on nutritional shift-up. In contrast, transformants containing a mutant *tyrT* promoter in which the discriminator sequence is mutated to a form which abolishes stringent control increase growth rate normally. This inhibition of growth rate also depends critically on the amount of the upstream element that is present. These experiments suggest that the wild-type *tyrT* promoter, but not the discriminator mutant, acts as a sink for a molecule(s) that are necessary for an increased growth rate. In this system, the copy number of the *tyrT* promoter per cell is ~ 30 to 40, thus approximately doubling the effective number of stable RNA promoters per cell. The implications are clearly that the competed molecule(s) is present in limiting concentrations and that competition for factors that bind to promoter sites may be a crucial determinant of promoter activity under normal intracellular conditions. The only molecule that is known both to interact with the upstream region of the *tyrT* promoter and to distinguish functionally between different discriminator sequences is RNA polymerase holoenzyme. Since several polymerase molecules, possibly

Table 6
MODELS FOR GROWTH RATE DEPENDENT CONTROL OF STABLE RNA SYNTHESIS

Model	Passive control	Ribosome feedback	Polymerase availability
RNA polymerase	Limiting under all conditions	Excess under all conditions	Limiting at low growth rates except at highest growth rates or under nutritional shift-up
Sensory of metabolic state of cell	Regulators of biosynthetic operons	Ribosomes	RNA polymerase
Transcriptional effector	Promotor competition	Ribosomes	ppGpp, etc.

up to 7, can bind cooperatively to the promoter region of *tyrT* DNA in vitro, in vivo 30 to 40 copies of a plasmid with a wild-type *tyrT* promoter could potentially sequester 210 to 280 polymerase molecules, a significant fraction of the functional enzyme molecules in the cell. If the observed effects are indeed attributable to competition for RNA polymerase, it follows directly that the in vivo population of polymerase molecules is functionally partitioned between those molecules which interact with the wild-type discriminator and whose removal thereby inhibits growth and those which interact with the mutant discriminator and whose removal should not directly affect transcription of genes necessary for optimal growth. Nevertheless, it should be emphasized that experiments of this type are only consistent with a regulatory role for RNA polymerase in growth rate control; they do not exclude that other factors may be important in determining a precise relationship of transcription initiation with growth rate.

While the discriminator sequence can determine whether a promoter activity correlates positively or negatively with growth rate, the degree of positive correlation may depend on other sequences. Thus, a *tyrT* promoter lacking its upstream element is a weaker promoter in vivo and in vitro and also shows a steeper positive correlation of activity with growth rate. This observation is again consistent with the hypothesis that polymerase availability is a determinant of growth rate control. Since $rrnP_1$ promoters show a steeper positive correlation than many tRNA promoters,[183] this argument would predict that the $rrnP_1$ promoters should require a higher polymerase concentration for activation than some tRNA promoters (Figure 6).

C. Mechanism of Growth Rate Control

A crucial aspect of the differential utilization of promoters at different growth rates is the mechanism by the transcriptional apparatus reacts to both the nutritional environment of the cell and to its capacity for protein synthesis. Any model must account for the seminal observation that except at very low growth rates the bacterial cell produces no more ribosomes than are necessary to maintain protein synthesis at an optimal level.[177] Of the three classes of model proposed, all assume that intracellular signals regulate stable RNA synthesis at the level of transcription initiation (Table 6).

The passive control model of Maaløe[178] assumes that the total rate of RNA transcription is limited by the number of RNA polymerase molecules, that the rRNA promoters are constitutive, and that their activities are altered as a result of competition with other regulated nonribosomal promoters involved in various metabolic pathways. Thus, in this model, the polymerase responds indirectly to changes in environmental conditions. A direct prediction of this model is that the activity of any constitutive promoter should, like a stable RNA promoter, increase at higher growth rates. In contrast, fully induced *lacUV5* expression shows a negative correlation with growth rate.

A second model, the ribosome feedback regulation model, proposed by Nomura and colleagues,[179] assumes that bacterial cells always have an excess capacity for synthesizing all ribosomal components, but that the synthesis rate of rRNA (and therefore ribosomes) and of tRNA is feedback inhibited by free nonfunctioning ribosomes when produced in excess of the amount needed for protein synthesis. Thus, on this model RNA polymerase is always in functional excess and the feedback effect is exerted by a direct interaction of free ribosomes with the transcriptional apparatus. This latter aspect of the model has yet to be experimentally verified.

The third model, that of polymerase availability outlined in a previous section, proposes that the rate of stable RNA synthesis depends directly on the concentration of RNA polymerase and on its affinity for promoter sites. On this model, the availability of polymerase is limiting at low and intermediate growth rates, but may be in moderate excess at the highest growth rates or under conditions of nutritional shift-up. The coupling between transcription and both the nutritional status and the protein synthetic capacity of the cell is achieved by effector molecules such as ppGpp and possibly other macromolecular components of the translational machinery, for example, fmet $tRNA_f^{Met}$,[180,181] which binds to RNA polymerase and alters the equilibrium between high- and low-affinity forms of the enzyme. Thus, the polymerase, rather than the ribosome, is the sensor of the metabolic state of the cell.

There are now considerable experimental data suggesting that the rate of stable RNA synthesis and the protein synthetic capacity of a cell are coupled (this subject is exhaustively reviewed in Reference 178). However, it should be emphasized that none of the available evidence distinguishes between the ribosome feedback and polymerase availability models and it is thus quite conceivable that the actual mechanism could incorporate elements of both models.

ACKNOWLEDGMENTS

I would particularly like to thank Dr. H. Drew for introducing me to the fundamental principles of DNA structure, as well as Dr. H. Buc for many constructive discussions.

REFERENCES

1. **Buc, H. and McClure, W. R.,** 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, 1985.
2. **Spassky, A., Kirkegaard, K., and Buc, H.,** Changes in the DNA structure of the *lac* UV5 promoter during formation of an open complex with *Escherichia coli* RNA polymerase, *Biochemistry*, 24, 2723, 1985.
3. **Roe, Y. H., Burgess, R. R., and Record, M. T., Jr.,** Temperature dependence of the rate constants of the *Escherichia coli* RNA polymerase-λ P_R promoter interaction. Assignment of kinetic steps corresponding to protein conformational change and DNA opening, *J. Mol. Biol.*, 184, 441, 1985.
4. **von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A.,** Protein-nucleic acid interactions in transcription: a molecular analysis, *Annu. Rev. Biochem.*, 53, 389, 1984.
5. **McClure, W. R.,** Mechanism and control of transcription initiation in prokaryotes, *Annu. Rev. Biochem.*, 54, 171, 1985.
6. **Reznikoff, W. S., Siegele, D. A., Cowing, D. W., and Gross, C. A.,** The regulation of transcription initiation in bacteria, *Annu. Rev. Genet.*, 19, 355, 1985.
7. **Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. K. F.,** Factor stimulating transcription by RNA polymerase, *Nature (London)*, 221, 43, 1969.

8. **Chamberlin, M. J.,** Bacterial DNA-dependent RNA polymerases, *Enzymes*, 15, 61, 1982.
9. **Burton, Z. F., Burgess, R. R., Lin, J., Moore, D., Holder, S., and Gross, C. A.,** The nucleotide sequence of the cloned rpoD gene for the RNA polymerase sigma subunit from *E. coli* K12, *Nucleic Acids Res.*, 9, 2889, 1981.
10. **Landick, R., Vaughn, V., Lau, E. T., VanBogelen, R. T., Erickson, J. W., and Neidhardt, F. C.,** Nucleotide sequence of the heat shock regulatory gene of *E. coli* suggests its protein product may be a transcription factor, *Cell*, 38, 175, 1983.
11. **Grossman, A. D., Erickson, J. W., and Gross, C. A.,** The *htpR* gene product of *E. coli* is a sigma factor for heat shock promoters, *Cell*, 38, 383, 1984.
12. **Kustu, S., Sei, K., and Keener, J.,** Nitrogen regulation in enteric bacteria, in *Regulation of Gene Expression*, Booth, I. and Higgins, C. F., Eds., Cambridge University Press, Cambridge, 39, 213, 1986.
13. **Hunt, T. P. and Magasanik, B.,** Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG* and *glnL*, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 8453, 1985.
14. **Stragier, P., Parsot, C., and Bouvier, C.,** Two functional domains conserved in major and alternate bacterial sigma factors, *FEBS Lett.*, 187, 11, 1985.
15. **Merrick, M. J. and Gibbons, J. R.,** The nucleotide sequence of the nitrogen regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial RNA polymerase sigma factors, *Nucleic Acids Res.*, 13, 7607, 1985.
16. **Doi, R. H. and Wang, F.-F.,** Multiple prokaryotic ribonucleic acid polymerase sigma factors, *Microbiol. Rev.*, 50, 227, 1986.
17. **Meisenberger, O., Heumann, H., and Pilz, I.,** Small-angle X-ray study of DNA dependent RNA polymerase holoenzyme from *Escherichia coli*, *FEBS Lett.*, 123, 22, 1981.
18. **Hawley, D. K. and McClure, W. R.,** Compilation and analysis of *Escherichia coli* promoter DNA sequences, *Nucleic Acids Res.*, 11, 2237, 1983.
19. **Cowing, D. W., Bardwell, J. C. A., Craig, E. A., Woolford, C., Hendrix, R. W., and Gross, C. A.,** Consensus sequence of *Escherichia coli* heat shock promoters, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 2679, 1985.
20. **Ausubel, F. M.,** Regulation of nitrogen fixation genes, *Cell*, 37, 5, 1984.
21. **Benyon, J., Cannon, M., Buchanan-Wollaston, V., and Cannon, F.,** The *nif* promoters of *Klebsiella pneumoniae* have a characteristic primary structure, *Cell*, 34, 665, 1983.
22. **Elliott, T. and Geiduschek, E. P.,** Defining a bacteriophage T4 late promoter: absence of a "−35" region, *Cell*, 36, 211, 1984.
23. **Drew, H. R., Weeks, J. R., and Travers, A. A.,** Negative supercoiling induces spontaneous unwinding of a bacterial promoter, *EMBO J.*, 4, 1025, 1985.
24. **Johnson, W. C., Moran, C. P., Jr., and Losick, R.,** Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoter for a developmentally regulated gene, *Nature (London)*, 302, 800, 1983.
25. **Richmond, T. J., Finch, J. T., Rushton, R., Rhodes, D., and Klug, A.,** Structure of the nucleosome core particle at 7 Å resolution, *Nature (London)*, 311, 532, 1984.
26. **Drew, H. R. and Travers, A. A.,** DNA bending and its relation to nucleosome positioning, *J. Mol. Biol.*, 186, 773, 1985.
27. **Kirchhausen, T., Wang, J. C., and Harrison, S. C.,** DNA gyrase and its complexes with DNA: direct observation by electron microscopy, *Cell*, 41, 933, 1985.
28. **Travers, A. A.,** DNA structure and promoter function, *Biochem. Soc. Trans.*, 12, 199, 1986.
29. **Drew, H. R. and Travers, A. A.,** DNA structural variations in the *E. coli tyrT* promoter, *Cell*, 37, 491, 1984.
30. **Fratini, A. V., Kopka, M. L., Drew, H. R., and Dickerson, R. E.,** Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATT^{Br}CGCG, *J. Biol. Chem.*, 257, 14686, 1982.
31. **Calladine, C. R.,** Mechanics of the sequence-dependent stacking of bases in B-DNA, *J. Mol. Biol.*, 161, 343, 1982.
32. **Rhodes, D. and Klug, A.,** Helical periodicity of DNA determined by enzyme digestion, *Nature (London)*, 286, 573, 1980.
33. **Wang, J. C.,** Helical repeat of DNA in solution, *Proc. Natl. Acad. Sci. U.S.A.*, 70, 200, 1979.
34. **Marini, J. C., Levene, S. D., Crothers, D. M., and Englund, P. T.,** A bent helix in kinetoplast DNA, *Cold Spring Harbor Symp. Quant. Biol.*, 47, 279, 1982.
35. **Koo, H-S., Wu, H-M., and Crothers, D. M.,** DNA bending at adenine-thymidine tracts, *Nature (London)*, 320, 501, 1986.
36. **Diekmann, S.,** Sequence specificity of curved DNA, *FEBS Lett.*, 195, 53, 1986.
37. **Hagerman, P. J.,** Sequence dependence of the curvature of DNA: a test of the phasing hypothesis, *Biochemistry*, 24, 7033, 1985.
38. **Trifonov, E. N.,** Curved DNA, *Crit. Rev. Biochem.*, 19, 89, 1986.

39. **Widom, J.,** Bent DNA for gene regulation and DNA packaging, *Bioassays*, 2, 11, 1985.
40. **Hagerman, P. J.,** Evidence for the existence of stable curvature of DNA in solution, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 4632, 1984.
41. **Travers, A. A. and Klug, A.,** The bending of DNA in nucleosomes and its wider implications, *Phil. Trans. Roy. Soc. (London)*, in press.
42. **Satchwell, S. C., Drew, H. R., and Travers, A. A.,** Sequence periodicities in chicken nucleosome core DNA, *J. Mol. Biol.*, 191, 659, 1986.
43. **Drew, H. R. and Travers, A. A.,** Structural junctions in DNA — influence of flanking sequences on nuclease digestion specificities, *Nucleic Acids Res.*, 13, 4445, 1985.
44. **Bossi, L. and Smith, D. M.,** Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of *Salmonella*, *Cell*, 39, 643, 1984.
45. **Gourse, R. L., de Boer, H. A., and Nomura, M.,** DNA determinants of rRNA synthesis in *E. coli*: growth dependent regulation, feedback inhibition, upstream activation antitermination, *Cell*, 44, 197, 1986.
46. **Lee, C-H., Mizusawa, H., and Kakefuda, T.,** Unwinding of double-stranded DNA by dehydration, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 2838, 1981.
47. **Lilley, D. M. J.,** The genetic control of DNA supercoiling and vice versa, in *Regulation of Gene Expression*, Booth, I. and Higgin, C. F., Eds., Cambridge University Press, Cambridge, 1986, 106.
48. **Sigman, D. S., Spassky, A., Rimsky, S., and Buc, H.,** Conformational analysis of *lac* promoters using the nuclease activity of 1,10-phenanthroline-copper ion, *Biopolymers*, 24, 183, 1985.
49. **Gotoh, O. and Tagashira, Y.,** Stabilities of nearest-neighbour doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles, *Biopolymers*, 26, 1033, 1981.
50. **Mulligan, M. E., Hawley, D. K., Entriken, R., and McClure, W. R.,** *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity, *Nucleic Acids Res.*, 12, 789, 1984.
51. **Deuschle, U., Kammerer, W., Gentz, R., and Bujard, H.,** Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures, *EMBO J.*, 5, 2987, 1986.
52. **Berman, M. L. and Landy, A.,** Promoter mutations in the transfer RNA gene *tyrT* of *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A.*, 276, 4303, 1979.
53. **Kammerer, W., Deuschle, U., Gentz, R., and Bujard, H.,** Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process, *EMBO J.*, 5, 2995, 1986.
54. **Bernhard, S. L. and Meares, C. F.,** The σ-subunit of RNA polymerase contacts the leading ends of transcripts 9-13 bases long on the λ P_R promoter but not on T7 A1, *Biochemistry*, 25, 5914, 1986.
55. **Walter, G., Zillig, W., Palm, P., and Fuchs, E.,** Initiation of DNA-dependent RNA synthesis and the effect of heparin on RNA polymerase, *Eur. J. Biochem.*, 3, 194, 1967.
56. **Chamberlin, M. J.,** The selectivity of transcription, *Annu. Rev. Biochem.*, 43, 721, 1974.
57. **McClure, W. R.,** Rate-limiting steps in RNA chain initiation, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 5634, 1980.
58. **Hawley, D. K. and McClure, W. R.,** *In vitro* comparison of initiation properties of bacteriophage λ wild-type P_R and X3 mutant promoters, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6381, 1980.
59. **Hawley, D. K. and McClure, W. R.,** Mechanism of activation of transcription initiation from the λ P_{RM} promoter, *J. Mol. Biol.*, 157, 493, 1982.
60. **Kadesch, T. R., Rosenberg, S., and Chamberlin, M. J.,** Binding of *Escherichia coli* RNA polymerase holoenzyme to bacteriophage T7 DNA. Measurements of binding at bacteriophage T7 promoter A_1 using a template competition assay, *J. Mol. Biol.*, 155, 1, 1982.
61. **Rosenberg, S., Kadesch, T. R., and Chamberlin, M. J.,** Binding of *Escherichia coli* RNA polymerase holoenzyme to bacteriophage T7 DNA. Measurements of the rate of open complex formation at T7 promoter A_1, *J. Mol. Biol.*, 155, 31, 1982.
62. **Bujard, H., Niemann, A., Breunig, K., Roisch, U., Dresel, A., von Gabain, A., Gentz, R., Stuber, D., and Weiher, H.,** *Promoters: Structure and Function*, Rodriquez, R. L. and Chamberlin, M. J., Eds., Praeger, New York, 1982, 121.
63. **Strauss, H. S., Burgess, R. R., and Record, T. M.,** Binding of *Escherichia coli* RNA polymerase holoenzyme to a bacteriophage T7 promoter containing fragment: evaluation of promoter binding constants as a function of solution conditions, *Biochemistry*, 19, 3504, 1980.
64. **Lomonossoff, G. P., Butler, P. J. G., and Klug, A.,** Sequence-dependent variation in the conformation of DNA, *J. Mol. Biol.*, 149, 745, 1981.
65. **Dickerson, R. E.,** Base sequence and helix structure variation in B and A-DNA, *J. Mol. Biol.*, 166, 419, 1983.
66. **Calladine, C. R. and Drew, H. R.,** A base-centered explanation of the B-to-A transition in DNA, *J. Mol. Biol.*, 178, 773, 1984.
67. **Lutter, L. C.,** Deoxyribonuclease I produces staggered cuts in the DNA of chromatin, *J. Mol. Biol.*, 117, 53, 1977.

68. **Schmitz, A. and Galas, D. J.,** The interaction of RNA polymerase and lac repressor with the lac control region, *Nucleic Acids Res.*, 6, 111, 1979.
69. **Siebenlist, U., Simpson, R. B., and Gilbert, W.,** *E. coli* RNA polymerase interacts homologously with two different promoters, *Cell*, 20, 269, 1980.
70. **Hofer, B., Muller, D., and Koster, H.,** The pathway of *E. coli* RNA polymerase-promoter complex formation as visualized by footprinting, *Nucleic Acids Res.*, 13, 5995, 1985.
71. **Wang, J. C.,** Unwinding at the promoter and the modulation of transcription by DNA supercoiling, in *Promoters: Structure and Function*, Rodriguez, R. L. and Chamberlin, M. J., Eds., Praeger, New York, 1982, 229.
72. **Kovacic, T. and Wang, J. C.,** quoted in Reference 78.
73. **Spassky, A., Busby, S., and Buc, H.,** On the action of the cyclic AMP-cyclic AMP receptor protein complex at the *Escherichia coli* lactose and galactose promoter region, *EMBO J.*, 3, 43, 1984.
74. **Gralla, J. D.,** Rapid "footprinting" on supercoiled DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 3078, 1985.
75. **Russell, D. R. and Bennet, G. N.,** Characterisation of the B-lactamase promoter of pBR322, *Nucleic Acids Res.*, 6, 111, 1981.
76. **Travers, A. A., Lamond, A. I., Mace, H. A. F., and Berman, M. L.,** RNA polymerase interactions with the upstream region of the *tyrT* promoter, *Cell*, 35, 265, 1983.
77. **Kanazawa, H., Mabuchi, K., and Futai, M.,** Nucleotide sequence of the promoter region of the gene cluster for the proton-translocating ATPase from *Escherichia coli* and identification of the active promoter, *Biochem. Biophys. Res. Commun.*, 107, 568, 1982.
78. **Buc, H.,** Mechanism of activation of transcription by the complex formed between cyclic AMP and its receptor, *Biochem. Soc. Trans.*, p. 12, 1986.
79. **Chenchick, A., Beabealashvilli, R., and Mirzabekov, A.,** Topography of interaction of *Escherichia coli* RNA polymerase subunits with the *lacUV5* promoter, *FEBS Lett.*, 128, 46, 1981.
80. **Busby, S., Truelle, N., Spassky, A., Drayfus, M., and Buc, H.,** The selection and characterisation of two novel mutations in the overlapping promoters of the *Escherichia coli* galactose operon, *Gene*, 28, 201, 1984.
81. **Auble, D. T., Allen, T. L., and de Haseth, P. L.,** Promoter recognition by *Escherichia coli* RNA polymerase. Effects of substitutions in the spacer DNA separating the −10 and −35 regions, *J. Biol. Chem.*, 261, 11,202, 1986.
82. **Nickol, J. M. and Felsenfeld, G.,** DNA conformation at the 5′ end of the chicken adult β-globin gene, *Cell*, 35, 467, 1983.
83. **Schon, E., Evans, T., Welsh, J., and Efstratiadis, A.,** Conformation of promoter DNA: fine mapping of S1 hypersensitive sites, *Cell*, 35, 837, 1983.
84. **Hertschel, C. C.,** Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease, *Nature (London)*, 295, 714, 1982.
85. **Pulleyblank, D. E., Haniford, D. B., and Morgan, A. R.,** A structural basis for S1 nuclease sensitivity of double stranded DNA, *Cell*, 42, 271, 1985.
86. **Rhodes, D.,** Nucleosome cores reconstituted from poly (dA-dT) and the octamer of histones, *Nucleic Acids Res.*, 6, 1805, 1979.
87. **Simpson, R. T. and Kunzler, P.,** Chromatin and core particles formed from the inner histones and synthetic polydeoxynucleotides of defined sequence, *Nucleic Acids Res.*, 6, 1387, 1979.
88. **Wang, J. C,. Jacobsen, J. H., and Saucier, J. M.,** Physiochemical studies on the interactions between DNA and RNA polymerase. Unwinding of the DNA helix by *Escherichia coli* RNA polymerase, *Nucleic Acids Res.*, 4, 1225, 1977.
89. **Gamper, H. B. and Hearst, J. E.,** A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary initiation and ternary complexes, *Cell*, 29, 81, 1982.
90. **Borowise, J. A. and Gralla, J. D.,** High-resolution analysis of *lac* transcription complexes inside cells, *Biochemistry*, 25, 5051, 1986.
91. **Bertrand-Burggraf, E., Schrarr, M., Lefranc, J. F., and Daune, M.,** Effect of superhelicity on the transcription from the *tet* promoter of pB322. Abortive initiation and unwinding experiments, *Nucleic Acids Res.*, 12, 7741, 1984.
92. **Bertrand-Burggraf, E., Lefevre, J-F., and Daune, M.,** A new experimental approach for studying the association between RNA polymerase and the tet promoter of pB322, *Nucleic Acids Res.*, 12, 1697, 1984.
93. **Straney, D. C. and Crothers, D. M.,** Intermediates in transcription initiation from the *E. coli lac* UV5 promoter, *Cell*, 43, 449, 1985.
94. **Borowiecz, J. A. and Gralla, J. D.,** Supercoiling response of the *lac* p^s promoter *in vitro*, *J. Mol. Biol.*, 184, 587, 1981.
95. **Stirdivant, S. M., Crossland, L. D., and Bogorad, L.,** DNA supercoiling affects *in vitro* transcription of two maize chloroplast genes differently, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 4886, 1985.

96. **Mulligan, M. E., Brosius, J., and McClure, W. R.,** Characterisation *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter, *J. Biol. Chem.*, 260, 3529, 1981.
97. **Malan, T. P., Kolb, A., Buc, H., and McClure, W. R.,** Mechanism of CRP-cAMP activation of *lac* operon transcription initiation activation of the P1 promoter, *J. Mol. Biol.*, 180, 881, 1984.
98. **Ehrlich, R., Larousse, A., Jacques, M. A., Marin, M., and Reiss, C.,** *Eur. J. Biochem.*, 148, 293, 1985.
99. **Wood, D. C. and Lebowitz, J.,** Effect of supercoiling on the abortive initiation kinetics of the RNA-I promoter of Col E1 plasmid DNA, *J. Biol. Chem.*, 259, 11184, 1984.
100. **de Boer, H. A., Comstock, L. J., and Vasser, M.,** The *lac* promoter: a functional hybrid derived from the *trp* and *lac* promoters, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 21, 1983.
101. **Shaner, S. L., Melancon, P., Lee, K. S., Burgess, R. R., and Record, M. T., Jr.,** Ion effects on the aggregation and DNA binding reactions of *Escherichia coli* RNA polymerase, *Cold Spring Harbor Symp. Quant. Biol.*, 47, 463, 1982.
102. **Debenham, P. and Travers, A.,** Selective inhibition of $tRNA^{tyr}$ transcription by guanosine 3′ diphosphate 5′-diphosphate, *Eur. J. Biochem.*, 72, 512, 1977.
103. **Miller, J. S. and Burgess, R. R.,** Selectivity of RNA chain initiation *in vitro*. III. Variables affecting initiation of transcription, *Biochemistry*, 17, 2064, 1978.
104. **Lamond, A. I.,** Supercoiling response of a bacterial tRNA gene, *EMBO J.*, 4, 501, 1985.
105. **Debenham, P. G., Pongs, O., and Travers, A. A.,** Formylmethionyl-tRNA alters RNA polymerase specificity, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 870, 1980.
106. **Travers, A. A.,** On the nature of DNA promoter conformations: the effects of glycerol and dimethylsulphoxide, *Eur. J. Biochem.*, 47, 435, 1974.
107. **Nakanishi, S., Adhya, S., Gottesman, M., and Pastan, I.,** Activation of transcription at specific promoters by glycerol, *J. Biol. Chem.*, 249, 4050, 1974.
108. **Travers, A. A. and Young, H. A. F.,** Bimodal response of *E. coli* RNA polymerase to purine nucleotides, unpublished observations.
109. **Prosen, D. E. and Cech, C. L.,** An *Escherichia coli* RNA polymerase tight-binding site on T7 DNA is a weak promoter subject to substrate inhibition, *Biochemistry*, 25, 5378, 1986.
110. **Travers, A. A.,** unpublished observations.
111. **Gottesman, S.,** Bacterial regulation: global regulatory networks, *Annu. Rev. Genet.*, 18, 415, 1984.
112. **Lamond, A. I. and Travers, A. A.,** Stringent control of bacterial transcription, *Cell*, 41, 6, 1985.
113. **Cashel, M. and Rudd, K. E.,** The stringent response, in *The Biology of E. coli and S. typhimurium*, American Society Microbiology, Washington, D.C., in press.
114. **Little, R., Ryals, J., and Bremer, H.,** *rpoB* mutation in *Escherichia coli* alters control of ribosome synthesis by guanosine tetraphosphate, *J. Bacteriol.*, 154, 787, 1983.
115. **Nene, V. and Glass, R. E.,** Relaxed mutants of *Escherichia coli* RNA polymerase, *FEBS Lett.*, 153, 307, 1983.
116. **Glass, R. K., Jones, S. T., and Ishihama, A.,** Genetic studies on the B subunit of *Escherichia coli* RNA polymerase. VII. RNA polymerase is a target for ppGpp, *Mol. Gen. Genet.*, 203, 265, 1986.
117. **Travers, A. A., Buckland, R., Goman, M., Le Grice, S. S. G., and Scaife, J. G.,** A mutation affecting the subunit of RNA polymerase changes transcriptional specificity, *Nature (London)*, 273, 354, 1978.
118. **Little, R., Ryals, J., and Bremer, H.,** personal communication.
119. **Travers, A. A.,** Promoter sequence for the stringent control of bacterial ribonucleic acid synthesis, *J. Bacteriol.*, 141, 973, 1980.
120. **Travers, A. A.,** Conserved features of coordinately regulated *E. coli* promoters, *Nucleic Acids Res.*, 12, 2605, 1984.
121. **Dunn, R. J., Belagaje, R., Brown, E. L., and Khorana, H. G.,** The synthesis and cloning of two tyrosine suppressor tRNA genes with altered promoter sequences, *J. Biol. Chem.*, 256, 6109, 1981.
122. **Lamond, A. I. and Travers, A. A.,** Genetically separable functional elements mediate the optimal transcription and stringent regulation of a bacterial tRNA gene, *Cell*, 40, 319, 1985.
123. **Travers, A. A.,** A $tRNA^{tyr}$ promoter with an altered response to ppGpp, *J. Mol. Biol.*, 141, 91, 1980.
124. **Mizushima-Sugano, J. and Kaziro, Y.,** Regulation of the expression of the *tufB* operon: DNA sequences directly involved in the stringent control, *EMBO J.*, 4, 1053, 1985.
125. **Sarmientos, P., Sylvester, J. E., Contente, S., and Cashel, M.,** Differential stringent control of the tandem *E. coli* ribosomal RNA promoters from the *rrnA* operon expressed *in vivo* in multicopy plasmids, *Cell*, 32, 1337, 1983.
126. **Gilbert, S. F., de Boer, H. A., and Nomura, M.,** Identification of initiation sites for the *in vitro* transcription of rRNA operons *rrnE* and *rrnA* in *E. coli*, *Cell*, 17, 211, 1979.
127. **Kingston, R. E.,** Effect of deletions near *Escherichia coli rrn* promoter P_2 on inhibition of *in vitro* transcription by guanosine tetraphosphate, *Biochemistry*, 22, 5249, 1983.

128. **Kajitani, M. and Ishihama, A.,** Promoter selectivity of *Escherichia coli* RNA polymerase: differential stringent control of the multiple promoters from ribosomal RNA and protein operons, *J. Biol. Chem.*, 259, 1951, 1984.
129. **Gourse, R.,** personal communication.
130. **Simpson, R. B.,** The molecular topography of RNA polymerase-promoter interaction, *Cell*, 18, 277, 1979.
131. **Debenham, P. G., Buckland, R., and Travers, A. A.,** The regulation of RNA polymerase by guanosine 5′ diphosphate 3′ diphosphate, in *Regulation of Macromolecular Synthesis by Low Molecular Weight Mediators*, Koch, G. and Richter, D., Eds., Academic Press, New York, 1979, 77.
132. **Van Ooyen, A. J. J., de Boer, H. A., Ab, G., and Gruber, M.,** Specific inhibition of ribosomal RNA synthesis *in vitro* by guanosine 3′ diphosphate 5′ diphosphate, *Nature (London)*, 254, 330, 1975.
133. **Travers, A. A.,** Modulation of RNA polymerase specificity by ppGpp, *Mol. Gen. Genet.*, 147, 225, 1976.
134. **Oostra, B. A., van Ooyen, A. J. J., and Gruber, M.,** *In vitro* transcription of three different ribosomal RNA cistrons of *E. coli*: heterogeneity of control regions, *Mol. Gen. Genet.*, 152, 1, 1977.
135. **Mizushima-Sugano, J., Miyajima, A., and Kaziro, Y.,** Selective inhibition of transcription of the *E. coli tufB* operon by guanosine-5′-diphosphate-3′-diphosphate, *Mol. Gen. Genet.*, 189, 185, 1983.
136. **Cashel, M., Hamel, E., Shapshak, P., and Bouquet, M.,** Interaction of ppGpp structural analogs with RNA polymerase, in *Control of Ribosome Synthesis*, Kjeldgaard, N. O. and Maaløe, O., Eds., Munksgaard, Copenhagen, 1976, 279.
137. **Debenham, P. and Travers, A. A.,** unpublished observations.
138. **Shaner, S. L., Piatt, D. M., Wersley, C. G., Yu, H., Burgess, R. R., and Record, M. T., Jr.,** Aggregation equilibria of *Escherichia coli* RNA polymerase: evidence for anion-linked conformational transitions in the promoters of core and holoenzymes, *Biochemistry*, 21, 5539, 1982.
139. **Fisher, R. and Blumenthal, T.,** Evidence for a specific association between RNA polymerase subunits σ and β involved in the closed to open complex transition, *J. Biol. Chem.*, 255, 11056, 1980.
140. **Travers, A. A., Buckland, R., and Debenham, P.,** Functional heterogeneity of *Escherichia coli* ribonucleic acid polymerase holoenzyme, *Biochemistry*, 19, 1656, 1980.
141. **Richardson, J. P.,** Some physical properties of RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.*, 55, 1616, 1966.
142. **Keilty, S. and Rosenberg, M.,** quoted in Reference 23.
143. **Busby, S.,** personal communication.
144. **Hochschild, A., Irwin, N., and Ptashne, M.,** Repressor structure and the mechanism of positive control, *Cell*, 32, 319, 1983.
145. **Dairi, T., Inokuchi, I., Mizuno, T., and Mizushima, S.,** Positive control of transcription initiation in *Escherichia coli*. A base substitution at the Pribnow box render *ompF* expression independent of a positive regulator, *J. Mol. Biol.*, 184, 1, 1985.
146. **Raibaud, O. and Schwatz, M.,** Positive regulation in bacteria, *Annu. Rev. Genet.*, 18, 173, 1974.
147. **Ho, Y-S., Wulff, D. L., and Rosenberg, M.,** Bacteriophage λ protein cII binds promoters on the opposite face of the DNA helix from RNA polymerase, *Nature (London)*, 304, 703, 1983.
148. **Shih, M-C. and Gussin, G. N.,** Differential effects of mutations on discrete steps in transcription initiation at the λ P_{RE} promoter, *Cell*, 34, 941, 1983.
149. **Hooper, B. C. and McClure, W. R.,** quoted in Reference 5.
150. **Hawley, D. K. and McClure, W. R.,** The effect of a lambda repressor mutation on the activation of transcription initiation from the lambda P_{RM} promoter, *Cell*, 32, 327, 1983.
151. **Irwin, N. and Ptashne, M.,** quoted in Reference 25.
152. **Blazy, B., Takahashi, M., and Baudras, A.,** Binding of CRP to DNA-dependent RNA polymerase from *E. coli*: modulation by cAMP of the interactions with free and DNA bound holo and core enzyme, *Mol. Biol. Rep.*, 6, 39, 1980.
153. **Irwin, N. and Ptashne, M.,** quoted in Reference 73.
154. **Wu, H. and Crothers, D. M.,** The locus of sequence-directed and protein induced DNA bending, *Nature (London)*, 308, 509, 1984.
155. **Hendrickson, W. and Schlief, R.,** A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 3129, 1985.
156. **Brown, S. E. and Ausubel, F. M.,** Mutations affecting regulation of *Klebsiella pneumoniae nifH* (nitrogenase reductase) promoter, *J. Bacteriol.*, 152, 143, 1984.
157. **Buck, M., Miller, S., Drummond, M., and Dixon, R.,** Upstream activator sequences are present in the promoters of nitrogen fixation genes, *Nature (London)*, 320, 374, 1986.
158. **Reitzer, L. J. and Magasanik, B.,** Transcription of *glnA* in *E. coli* is stimulated by an activator bound to sites far from the promoter, *Cell*, 45, 785, 1986.
159. **Dunn, T. M., Hahn, S., Ogden, S., and Schleif, R. F.,** An operator at −280 bp that is required for the repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 5017, 1984.

160. **Ptashne, M.,** Gene regulation by proteins acting nearby and at a distance, *Nature (London)*, 322, 697, 1986.
161. **Lamond, A. I.,** The control of stable RNA synthesis in bacteria, *TIBS*, 10, 271, 1985.
162. **Lamond, A. I. and Travers, A. A.,** Requirement for an upstream element for optimal transcription of a bacterial tRNA gene, *Nature (London)*, 305, 248, 1983.
163. **Lamond, A. I.,** Functional Elements of a Bacterial tRNA Promoter, Ph.D. thesis, University of Cambridge, Cambridge, 1984.
164. **Travers, A. A., Lamond, A. I., and Weeks, J. R.,** Alteration of the growth rate dependent regulation of *Escherichia coli tyrT* expression by promoter mutations, *J. Mol. Biol.*, 191, 251, 1986.
165. **Travers, A. A., Lamond, A. I., and Mace, H. A. F.,** ppGpp regulates the binding of two RNA polymerase molecules to the *tyrT* promoter, *Nucleic Acids Res.*, 10, 5043, 1982.
166. **Roe, J-H. and Record, M. T., Jr.,** Regulation of the kinetics of the interaction of *Escherichia coli* RNA polymerase with the lambda P_R promoter by salt concentration, *Biochemistry*, 24, 4721, 1985.
167. **Munro, G. F., Hercules, K., Morgan, J., and Sauerbier, W.,** Dependence of the putrecine content of *Escherichia coli* on the osmotic strength of the medium, *J. Biol. Chem.*, 247, 1272, 1972.
168. **Measures, J. C.,** Role of amino acids in osmoregulation of non-holophilic bacteria, *Nature (London)*, 257, 398, 1975.
169. **Martin, K. and Schleif, R. F.,** Equilibrium DNA binding of AraC protein: compensation for displaced ions, *J. Mol. Biol.*, in press.
170. **Courey, A. J. and Wang, J. C.,** Cruciform formation in negatively supercoiled DNA may be kinetically forbidden under physiological conditions, *Cell*, 33, 817, 1983.
171. **Ryals, J., Little, R., and Bremer, H.,** Control of rRNA and tRNA synthesis in *Escherichia coli* by guanosine tetraphosphate, *J. Bacteriol.*, 151, 1261, 1982.
172. **Shepherd, N. S., Churchward, G., and Bremer, H.,** Synthesis and activity of ribonucleic acid polymerase in *Escherichia coli*, *J. Bacteriol.*, 141, 1098, 1980.
173. **Brunschede, H., Dove, T. L., and Bremer, H.,** Establishment of exponential growth after a nutritional shift-up in *Escherichia coli* B/r: accumulation of deoxyribonucleic acid, ribonucleic acid and protein, *J. Bacteriol.*, 129, 1020, 1977.
174. **McClure, W. R., Hawley, D. K., and Malan, T. P.,** The mechanism of RNA polymerase activation at the lambda P_{RM} and *lac* P^+ promoters, in *Promoters: Structure and Function*, Rodriguez, R. L. and Chamberlin, M. J., Eds., Praeger, New York, 1982, 111.
175. **Linn, S. and Riggs, A. D.,** The general affinity of *lac* repressor for *E. coli* DNA: implications for gene regulation in procaryotes and eukaryotes, *Cell*, 4, 107, 1975.
176. **Glaser, G., Sarmientos, P., and Cashel, M.,** Functional interrelationship between two tandem *E. coli* ribosomal RNA promoters, *Nature (London)*, 302, 74, 1983.
177. **Schaechter, M., Maaløe, O., and Kjeldgaard, N. O.,** Dependency on medium and temperature on cell size and chemical composition during balanced growth of *Salmonella typhimurium*, *J. Gen. Microbiol.*, 19, 592, 1958.
178. **Maaløe, O.,** An analysis of bacterial growth, *Dev. Biol. Suppl.*, 3, 33, 1969.
179. **Nomura, M., Gourse, R., and Baughman, G.,** Regulation of the synthesis of ribosomes and ribosomal components, *Annu. Rev. Biochem.*, 53, 75, 1984.
180. **Debenham, P. G., Pongs, O., and Travers, A.,** Formyl-methonyl-tRNA alters RNA polymerase specificity, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 870, 1980.
181. **Nomura, T., Fujita, N., and Ishihama, A.,** Promoter selectivity of *Escherichia coli* RNA polymerase: alteration by $fMet\text{-}tRNA_f^{Met}$, *Nucleic Acids Res.*, 14, 6857, 1986.