Annu. Rev. Genet. 2002. 36:175–203
doi: 10.1146/annurev.genet.36.032902.111815


# DNA TOPOLOGY-MEDIATED CONTROL OF GLOBAL GENE EXPRESSION IN *ESCHERICHIA COLI*

G. Wesley Hatfield
*Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, California 92697; e-mail: gwhatfie@uci.edu*

Craig J. Benham
*Department of Mathematics, University of California, Davis, California 95616; cjbenham@ucdavis.edu*



**Key Words** DNA supercoiling, SIDD, transcriptional coupling, energy charge, metabolic regulation

■ **Abstract** Because the level of DNA superhelicity varies with the cellular energy charge, it can change rapidly in response to a wide variety of altered nutritional and environmental conditions. This is a global alteration, affecting the entire chromosome and the expression levels of all operons whose promoters are sensitive to superhelicity. In this way, the global pattern of gene expression may be dynamically tuned to changing needs of the cell under a wide variety of circumstances. In this article, we propose a model in which chromosomal superhelicity serves as a global regulator of gene expression in *Escherichia coli*, tuning expression patterns across multiple operons, regulons, and stimulons to suit the growth state of the cell. This model is illustrated by the DNA supercoiling-dependent mechanisms that coordinate basal expression levels of operons of the *ilv* regulon both with one another and with cellular growth conditions.

CONTENTS

## INTRODUCTION

Gene expression patterns must be stringently regulated according to the nutritional needs of the cell and environmental conditions. For bacteria, these circumstances can change rapidly and drastically. For example, an enteric organism like *Escherichia coli* that has been growing under relatively nutrient-rich, stable conditions can suddenly find itself relocated to a hostile environment that is essentially devoid of nutrients. To survive such a transition, this organism possesses a wide variety of metabolic and genetic regulatory networks that enable it to rapidly adjust to new conditions.

Our current understanding of the regulatory systems that coordinate gene expression in *E. coli* involves several hierarchical levels—local control of individual operons, regional control of multiple operons within a regulon, and of multiple regulons within a stimulon or modulon (59), and global control of overall expression patterns (Figure 1). The most basic and best-understood level is the regulation of individual operons. Many different types of operon-specific control mechanisms have been described, each of which responds to regulatory signals that are closely related to its function. For example, the expression of an operon encoding genes for a biosynthetic pathway commonly is repressed by pathway-specific end-products, while expression of an operon encoding genes for a catabolic pathway often is activated by pathway-specific substrates. The DNA binding proteins that mediate operon-specific regulation commonly are present in small numbers

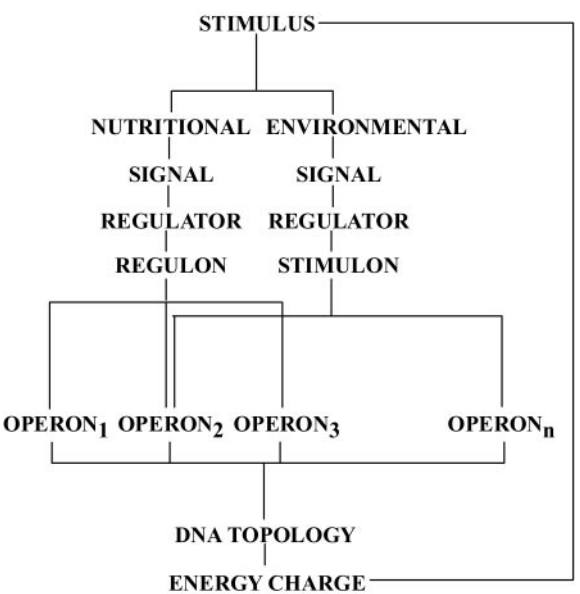


**Figure 1** Hierarchical levels of genetic regulatory circuits in bacteria. See text for discussion.

in the cell and bind in a highly site-specific manner to only a few DNA target sites.

At the next hierarchical level are regulons, which are groups of operons that are coordinately regulated by nutrient or environmental conditions. The operons within a regulon usually participate in a common function, such as nitrogen or carbon utilization, and share a common regulator, usually a protein repressor or activator, that recognizes a DNA target sequence common to all members. The regulatory proteins involved in regulon control are more abundant in the cell, and bind to multiple DNA target sites on the chromosome.

At an even higher hierarchical level, certain environmental changes, such as in the osmolarity or oxygen content of the growth medium, may generate a signal that induces operons contained in multiple regulons. These overlapping networks are referred to as stimulons. The DNA binding proteins that regulate stimulons commonly are even more abundant, and bind to many, often quite degenerate, DNA target sites.

These levels of regulation target specific operons or sets of operons, adjusting their relative expression levels to optimize their activities to the current conditions of the cell. They do not address the highest level of regulation, the global coordination of the expression levels of all the genes in the genome that enables the cell to efficiently accommodate to changing conditions. This global level of gene regulation requires the integration of a variety of nutritional and environmental signals, and the generation of a coordinated response that adjusts the basal levels of expression of all genes so as to optimize cell growth and survival under a broad range of possibly rapidly changing conditions. This arrangement does not override operon-, regulon-, or stimulon-specific controls, but rather tunes the global gene expression patterns to the demands of the cell under the prevailing conditions. For example, a lower basal level of amino acid biosynthesis is needed during stationary phase than during log phase growth. So the expression levels of genes encoding enzymes required for amino acid biosynthesis are lowered in stationary phase, but must rapidly increase during transition to log phase. Throughout this transition, all operon-, regulon-, and stimulon-specific controls on these genes must remain operative, fine-tuning their expression to specific circumstances.

We propose that this highest level of hierarchical gene regulation is mediated by DNA supercoiling whose level is regulated by the energy charge of the cell, which in turn is modulated by nutrient and environmental signals. Here we review the evidence supporting this model.



## DNA SUPERCOILING, ENERGY CHARGE, AND TRANSCRIPTIONAL REGULATION

### Control of DNA Supercoiling

With very few exceptions, DNA extracted from prokaryotic sources has a linking number deficiency, $\Delta Lk < 0$, and is therefore negatively supercoiled [for reviews see (27) and (16)]. This in vivo level of negative supercoiling is tightly controlled

by the combined influences of many factors, including DNA binding proteins, transcription, replication, and the activities of topoisomerase enzymes (21, 27, 51).

Nucleoids that have been carefully removed from *E. coli* are bound by a variety of proteins. Some of these DNA binding proteins stabilize supercoils by wrapping the DNA into stably wound toroidal loops. These constrained supercoils are not lost when the DNA is nicked (66). However, when these binding proteins are removed the constraint on these supercoils is released and the measured superhelix density changes, even though the actual linking number is invariant (75, 76, 91). The portion of the overall supercoiling that is lost when the DNA is nicked is called unconstrained supercoiling, or superhelical tension. About 50% of the supercoiling in *E. coli* is constrained by protein binding (10, 11, 66, 69, 82). A major source of this restraint comes from binding proteins that separate DNA strands, including proteins of the replication apparatus and RNA polymerase (RNAP). Architectural proteins such as HNS, IHF, and the histone-like HU proteins also restrain negative supercoils. It has been observed that HU-deficient mutant strains exhibit levels of supercoiling approximately 15% lower than those of wild-type strains (39). Further reductions are seen in strains that also contain deletions of genes encoding other chromatin organizing proteins such as HNS and IHF (38, 65).

The level of unconstrained superhelical tension within cells is determined by the activities of the enzymes DNA gyrase, and the DNA topoisomerases Topo I and Topo IV (27, 28, 34, 84). DNA gyrase introduces negative supercoils into DNA by a reaction requiring the hydrolysis of ATP. Topo I removes negative supercoils, and thereby relaxes negative superhelical tension, in an ATP-independent reaction (99). Although Topo IV also removes negative supercoils, its primary function appears to be resolving DNA knots and catenanes by a process that requires ATP (24).

Early genetic studies revealed that secondary lesions in the genes encoding DNA gyrase (*gyrA* or *gyrB*) are always observed in strains in which the entire functional *top*A gene that encodes Topo I has been deleted. Null mutations in *gyrB* are lethal in these strains (25, 52, 63, 68, 70). These results suggest that tight control of the interplay between Topo I and gyrase is necessary for optimal growth.

A variety of counterbalancing controls are maintained on gyrase and Topo I, involving regulation both of the activities of these enzymes and of the level of expression of their encoding genes. These controls collectively act as a homeostatic feedback system, maintaining the negative superhelical tension on the *E. coli* chromosome within narrow limits around an optimal level whose value depends on the growth phase of the cell and on environmental conditions (20, 28, 53, 87).

## Cellular Energy Charge and Global DNA Supercoiling Levels

The elegant and extensive studies by Atkinson and co-workers have shown that the energy charge of the adenylate pool, defined as $([ATP] + {}^1\!/_2[ADP])/([ATP] + [ADP] + [AMP])$, is the parameter that correctly describes the amount of metabolically available energy for the cell (1, 2, 18, 19). They also demonstrated that, during states of metabolic adjustment when the energy charge transiently decreases, such

as in the transition from aerobic to anaerobic growth, enzymes involved in ATP-regenerating reactions are activated and enzymes involved in ATP-utilizing reactions are inhibited. In general, decreases in the energy charge induce increases in the rates of enzymes that produce ATP and decreases in the rates of enzymes that consume ATP, while increases in the energy charge have the opposite effect. Together these changes maintain the energy charge in homeostatic balance. These findings explain why, although the absolute levels of adenylate pools differ under different growth conditions, the energy charge of the cell remains constant at a value of $\sim$0.85 during balanced growth under all growth conditions. In stationary phase, however, the energy charge is maintained at a lower level (41). To facilitate the following discussion, we refer to energy charge as the cellular [ATP/ADP] ratio, realizing that this is a simplification of Atkinson's definition.

Many studies have shown that the level of global negative supercoiling is controlled by the cellular energy charge. Because the enzymatic activity of gyrase is controlled by the intracellular [ATP]/[ADP] ratio, not by the free ATP concentration (38, 39), high negative superhelical densities occur when cellular energy charge is high, and low negative superhelical densities occur when it is low (27, 29, 39, 95, 96). It is well known that energy charge and DNA supercoiling play coordinated roles in cellular adaptation and survival, both under suboptimal growth conditions and during growth state transitions. In nongrowing *E. coli* cells in stationary phase where the [ATP/ADP] ratio is low, the superhelical density of a reporter plasmid is $\sigma \approx -0.03$. As cells recover and enter into log phase, the [ATP/ADP] ratio increases and the global negative superhelical density moves into the midphysiological level of $\sigma \approx -0.05$, a value typical during balanced growth (45). Physical stresses alter both cellular energy charge and DNA supercoiling levels. For example, osmotic stress (salt shock) causes the cellular [ATP/ADP] ratio to transiently increase fourfold, and the negative superhelical density of the bacterial chromosome to increase to a value as high as $\sigma = -0.09$ (38). During transitions from aerobic to anaerobic growth the cellular [ATP/ADP] ratio decreases, and the global negative superhelical density transiently falls from $\sigma = -0.05$ to $\sigma = -0.038$ (20).

## Effects of DNA Supercoiling on Gene Expression

Steck et al. used O'Farrell 2-D protein gel electrophoresis techniques to quantify the relative abundances of proteins expressed in *E. coli* strains containing nonlethal mutations in either *gyrB* or *topA* that alter the global superhelical density of the chromosome (84). Of the 88 proteins whose abundances were quantified, 39% showed changes of abundance, and inferentially of cognate gene expression levels, during steady-state growth in oversupercoiling *topA* versus undersupercoiling *gyrB* mutants. Maximal abundances of some proteins occurred at supercoiling levels below that of the wild type, while others were most abundant at elevated negative superhelix densities. A third class exhibited optimum expression at a normal physiological supercoiling level.

There are many ways in which DNA template topology (i.e., an imposed linking difference $\Delta Lk$) can influence gene expression. These could involve changes

of helicity, $\Delta Tw$, and/or changes of tertiary structure, $\Delta Wr$. The former include alterations of the helical twist of the B-form, and/or local transitions to alternative secondary structures having different helicities from that of B-DNA, while the latter require bending deformations. Supercoils that are stabilized by interactions with other molecules can be toroidal, but unstabilized supercoils commonly are plectonemically interwound. [For a recent review of the ways in which DNA topology can influence transcriptional activity see (23)].



REGULATION BY SUPERHELICAL MODULATION OF HELICAL TWIST Drlica and coworkers showed that the level of supercoiling which gives optimal expression for promoters correlates with the length of the spacer region between their −35 and the −10 regions (84). Promoters whose maximum expression occurs at low superhelical densities tend to have spacer regions that are shorter than 17 base pairs. Promoters whose maximal activity occur at high levels of supercoiling commonly have spacers that are longer than 17 base pairs. Promoters with 17 base pair spacers were preferentially optimized for normal physiological levels of supercoiling and showed less sensitivity to changes in supercoiling than did the others.

The relative orientation between the −35 and the −10 regions can strongly affect the ability of the $\sigma^{70}$ subunit of RNA polymerase (RNAP) to locate and bind to a promoter (92). As negative supercoiling untwists DNA (i.e., $\Delta Lk < 0$, so $\Delta Tw < 0$), its effect on the helical twist of the spacer can explain the correlation between promoter spacer length, superhelicity, and activity. A long spacer region, with a larger intrinsic twist than would provide optimal alignment, will in this model become more active at higher negative superhelical densities because this deformation decreases twist. Conversely, a short spacer whose twist is less than optimal would have its activity decreased by negative supercoiling, and hence would be more active at smaller superhelix densities. Optimal recognition occurs with 17-base pair spacing containing the negative helical twist characteristic of a midphysiological superhelical density. Cell growth conditions that alter the global superhelical density of the chromosome, by changing the helical alignment of recognition elements within $\sigma^{70}$ promoters, would affect their activities accordingly. Examples of such a supercoiling-induced realignment mechanism have been demonstrated in promoters involved in processes as diverse as the cold-shock and osmotic-shock responses, amino acid biosynthesis, and carbon utilization (14, 42, 43, 45, 90, 92).

REGULATION BY SUPERHELICAL MODULATION OF TERTIARY STRUCTURE The imposition of a negative linking difference on a topological domain of DNA may alter the tertiary structure through its effect on Wr. In short regions, this can cause looping, while in longer regions it can induce plectonemic interwinding. Looped structures can form microdomains, which act as small independent topological domains. This process has been proposed to be involved in prokaryotic transcriptional activation (56). The formation of an interwound structure brings regions that are remote along the sequence into close physical proximity. If the DNA reptates through such a structure, eventually any site will find itself close in space to any

other site. In this way, plectonemic interwinding can greatly enhance the opportunities for sites remote along the duplex, or molecules bound thereto, to interact. This effect is the basis for the activity of the NtrC-dependent enhancer in *E. coli*, which can act in a supercoil-dependent manner over large distances, on the order of 2000 bp (50).

REGULATION BY SUPERHELICALLY DRIVEN TRANSITIONS TO ALTERNATIVE SECONDARY STRUCTURES DNA supercoiling is known to drive transitions to a wide variety of alternative secondary structures, including local denaturation (44), transitions to Z-form (5) and to H-form (40), and cruciform extrusion (48). The formation of alternative DNA structures can serve regulatory functions, either by forming or modifying a regulatory binding site, or by altering the level of unconstrained supercoiling in the balance of the domain (3, 4, 7, 8). Dai & Rothman-Denes have shown that the bacteriophage N4 virion RNA polymerase (vRNAP) promoters contain short inverted repeat sequences centered at position −12 (22). These sites extrude cruciforms that are required for vRNAP recognition at physiological levels of supercoiling.

Negative superhelicity is known to destabilize the DNA duplex at specific locations (44). This effect, known as stress-induced duplex destabilization (SIDD), has been implicated in a variety of regulatory processes, including many that are involved in transcriptional regulation. SIDD can decrease the energy required for open complex formation and thereby increase transcriptional activity. Travers and coworkers used in vitro transcription and S1 nuclease to probe the structure of the *tyr*T promoter. They found that negative supercoiling increased the rate of initiation of transcription from this promoter by inducing unwinding that assisted open complex formation (26). SIDD also can affect the binding affinities of single strand-specific DNA binding proteins. Levens and coworkers have shown that the FUSE element located 1500 bp upstream from the promoter region of the human *c-myc* gene regulates the initiation of transcription from that gene (54). The mechanism for this activation is formation of a SIDD site at the FUSE element, which then reacts with the single strand-specific regulatory binding protein FPB. Although this eukaryotic system has a large set of other regulatory transcription factors and elements, in the absence of this specific system endogenous *c-myc* expression cannot be sustained.

Local denaturation is the most extreme form of duplex destabilization. By altering the helicity of the region involved, local denaturation diminishes the level of unconstrained superhelicity experienced by the rest of the topological domain. This process can affect any regulatory activity that is attuned to supercoiling levels. The converse process, whereby the binding of a DNA duplex binding protein to a destabilized site can force it back to B-form, also can be important. For example, this binding-induced reassociation can transmit the destabilization from the original SIDD site to the next most susceptible location, which can be a substantial distance away. We describe below cases where this process regulates promoter activity.

## SIDD ANALYSIS

Local transitions to conformations that are less twisted in the right-handed sense than the B-form can be driven by negative superhelicity (i.e., $\Delta Lk < 0$). By accommodating some of the imposed linking difference as a net decrease of the local twist at the transition site(s), they diminish by a corresponding amount the superhelical deformation imposed on the balance of the molecule. Although transition to an alternative conformation requires energy, it also releases energy by this partial relaxation. In a domain containing a single susceptible site, transition will occur when the energy it releases exceeds its energy costs.

Experimentally observed superhelical transitions include cruciform extrusions (48), transitions to Z-form (83), and to H-form (40, 44). While regulatory roles have been suggested for each of these types of transitions, to date only strand separation has been shown to be widely involved in regulation.

The constraint imposed by DNA superhelicity is the constancy of the linking number within a topological domain. Suppose this domain contains $N$ base pairs, and that it is in a state in which $n$ base pairs are denatured. If the helical twist of B-DNA is $A$ bp/turn, then $n/A$ helical turns are untwisted when these $n$ base pairs change from B-form to the unstressed denatured state. If $\Delta Lk = -n/A$, then this transition completely relaxes the imposed superhelicity. However, if $n$ has any other value, complete relaxation does not occur and a residual deformation remains. Because single-stranded DNA is highly flexible, this residual deformation will cause the two single strands within an unpaired region to helically rotate around each other with a twist rate of $\tau$ radians per bp. The residual linking difference $\Delta Lk_{res}$ is the amount of the imposed linking difference that remains after these two twist effects of the transition are accounted for. In this way, the superhelical constraint is expressed by the equation

$$\Delta \mathrm{Lk} = \Delta \mathrm{Lk}_{\mathrm{res}} - (\mathrm{n}/\mathrm{A}) + (\mathrm{n}\tau/2\pi).$$

Although thermal denaturation in linear or nicked DNA involves only near-neighbor interactions, when the transition is driven by superhelicity, this constraint globally couples together the conformational states of all the base pairs within a topological domain. This coupling occurs because transition of any base pair alters its helicity, which by the above equation changes the partitioning of $\Delta Lk$ throughout the domain. So whether transition occurs at a given site depends not just on its local sequence properties, but also on how that transition competes with all others to which the domain is susceptible. For this reason, strictly local methods are not appropriate for analyzing superhelical transitions. Instead, these must be analyzed as global events, including competitions among all local transitions that the base sequence of the domain permits.

The behavior of a superhelical domain that contains multiple sites susceptible to transitions will be determined by a possibly complex competition among them. Whether or not transition occurs at a specific site will depend not just on its local properties, but also on how well that transition competes with others elsewhere in

the domain. For example, Lilley and colleagues have shown that local denaturation of an A + T-rich region contained within a supercoiled DNA plasmid requires a significantly higher level of negative supercoiling when a $(TG)_{12}$ region is inserted in the plasmid (12). This insertion creates a site that can form left-handed Z-DNA under superhelical tension, and this transition competes effectively with denaturation. This illustrates how superhelically induced structural transitions compete with one another for the negative superhelical energy required for their formation, with the presence of a more susceptible site inhibiting less favorable transitions at other positions. However, this competition is not determined exclusively by the energy costs of the competing transitions themselves, but rather by the net energy difference for each between its cost and the relaxation it affords. Thus, even though the B-Z transition may be energetically more expensive than denaturation, it still can be favored because it produces significantly more relaxation per transformed base pair.

In general, a given linking difference imposed on a DNA domain can be accommodated by many combinations of torsional and flexural deformations, and by those conformational transitions to which its base sequence renders it susceptible. All of these responses to superhelicity require energy, and they all are topologically coupled together by the superhelical constraint arising from the requirement that the linking number must remain constant. In consequence, there are very many conformational states accessible to such a molecule. The competition among these alternative conformational states determines a thermodynamic equilibrium distribution. That is, a population of identical molecules at equilibrium distributes itself among all accessible states according to their energies, with low-energy states being exponentially more populated than high-energy states.

Benham and coworkers have developed theoretical techniques to analyze the superhelical stress-induced destabilization of the DNA double helix (5, 32, 86). Although these methods focus explicitly on denaturation, the same formalism can be applied to analyze other types of transitions. The values of the energy parameters used in these calculations have all been taken from experimental measurements. In particular, the free energy required for denaturation is known to depend significantly on base sequence and environmental conditions (15, 85). This is why destabilization is not uniform along a sequence, but instead is concentrated at specific locations.

These methods calculate two SIDD properties at single base-pair resolution in domains of specified base sequence and superhelicity. The transition profile is the graph of the probability of denaturation of each base pair along the DNA sequence. A more sensitive measure of destabilization may be calculated as the incremental free energy $G(x)$ needed to force the base pair at position $x$ to always be separated (7, 25). A value of $G(x)$ near or below zero indicates an essentially completely destabilized base pair, which is predicted to denature with high probability at equilibrium. Positive values of $G(x)$ occur for base pairs where incremental free energy is needed to assure separation. Regions of partial destabilization are indicated by intermediate $G(x)$ values. Stress-induced duplex destabilization (SIDD) profiles are plots of $G(x)$ versus $x$. Figure 2 shows the SIDD profile (*A*) and the

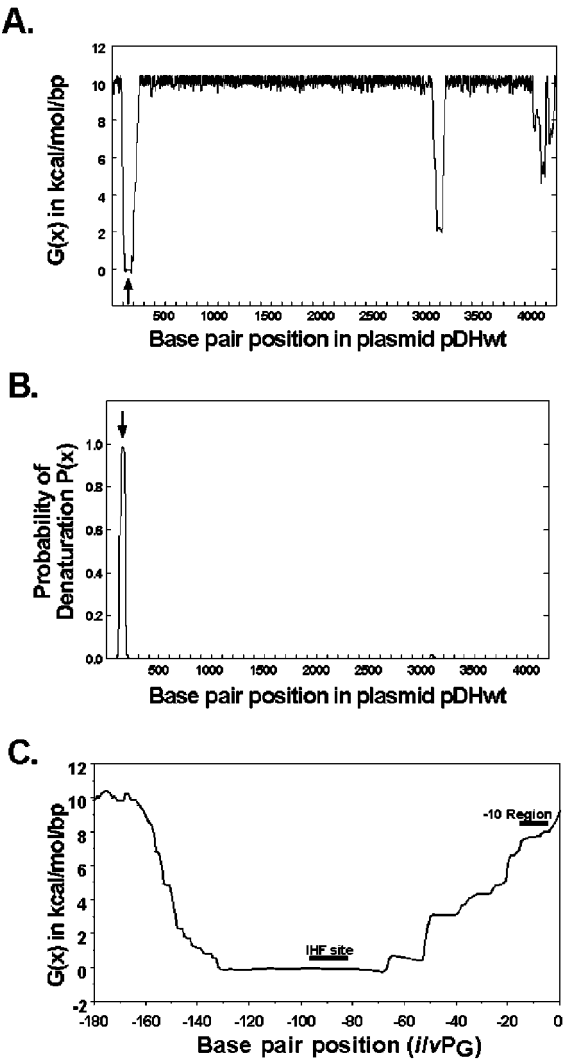


**Figure 2** Stress-induced duplex destabilization (SIDD) profile of plasmid pDHΔwt at $\Delta Lk = \sigma - 0.05$. (*A*) SIDD profile of pDHΔwt indicating the predicted free energy G(x) required for DNA duplex destabilization as a function of base pair location x. The arrow indicates the location of the IHF binding site in the UAS region of the *ilv*$P_G$ promoter insert. (*B*) P(x) is the probability of DNA duplex destabilization as a function of base pair location x. The arrow indicates the location of the IHF binding site in the UAS1 region. (*C*) Closeup SIDD profile [G(x)] of pDHΔwt from bp positions −180 to +1 relative to the *ilv*$P_G$ transcriptional start site (bp positions 76 to 256 in the plasmid).

denaturation probability profile (*B*) for the experimental pDHΔwt plasmid containing the *ilv*$P_G$ promoter that was used in experiments described in the following section.

SIDD profiles are more informative than transition profiles because they also depict sites where the amount of free energy needed to induce denaturation is fractionally decreased, but not enough to denature with a significant probability. This will be important when duplex opening occurs by processes that can provide sufficient free energy to cause local denaturation only if the DNA site involved already is marginally destabilized by stresses. Such partially destabilized regions could be biologically important as facilitators of strand separation by enzymatic or other processes.

Calculations have been performed of the predicted stress-induced destabilization properties of numerous genomic DNA sequences. These calculations commonly assume a superhelix density of $-0.055$, which corresponds to a midphysiological value; however, calculations at any superhelical density are possible. The deformation and transition energy parameters are given their experimentally measured values, so there are no free parameters in the analysis. Yet the predictions of these calculations are in precise quantitative agreement with experimental results in all cases for which experimental data on the locations and extents of superhelical denaturation are available. The sites that denature are predicted exactly, and the calculated extents of transition at each site agree precisely with the experimental measurements. In every case, the predicted superhelicity required to drive a specific amount of separation is within one turn of the observed value over the whole range where experiments were performed (6, 9, 32). This reflects the limit of accuracy with which extents of transition can be experimentally measured. And most importantly, the major changes in the locations of destabilized regions that result from minor sequence alterations are precisely predicted.

## GLOBAL REGULATION OF GENE EXPRESSION BY DNA TOPOLOGY-DEPENDENT MECHANISMS

To enable an organism to be both metabolically efficient and rapidly adaptive, mechanisms must exist to coordinate its global patterns of gene expression to its growth and environmental conditions. We propose that global levels of gene expression from specific promoters are coupled to the growth and nutritional states and environmental conditions of the cell through the regulation of transcriptional initiation by mechanisms that are sensitive to DNA superhelicity. This implies, for example, that the basal level expression of operons encoding structural genes for the biosynthesis of intermediary metabolites should be coordinated above the operon-specific level, so their basal expression levels are low when the chromosomal superhelical density is low, and are high when it is high. Work in our laboratories has shown this to be the case for operons of the *ilv* regulon, which encode the structural genes for the enzymes required for the biosynthesis of the branched chain amino acids, L-isoleucine, L-valine, and L-leucine (Figure 3). We also have

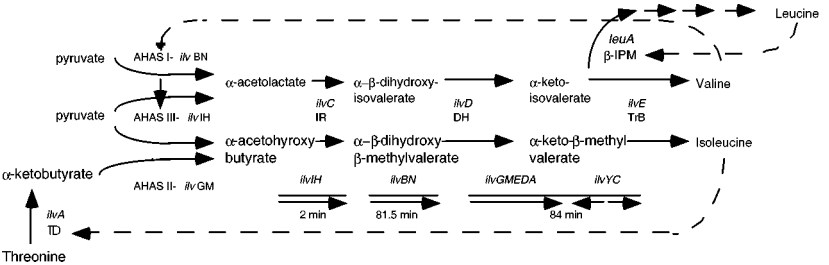


**Figure 3** The biosynthetic pathways for the synthesis of the branched chain amino acids L-isoleucine, L-valine, and L-leucine in *E. coli*. The enzymes involved in the common pathway for branched chain amino acid biosynthesis are abbreviated as follows: AHAS, acetohydroxyacid synthase; IR, acetohydroxyacid isomeroreductase; DH, dihydroxyacid dehydrase; TrB, transaminase B; TD, threonine deaminase; $\beta-$IPM, $\beta-$isopropylmalate synthase. Genes encoding each of these enzymes are indicated in italics. Feedback inhibition patterns are indicated by dashed lines. The genomic organization of the operons of the *ilv* regulon is shown below the metabolic pathway.



shown that genes encoding $tRNA^{Leu}$, which are not in the *ilv* regulon, are also expressed in a manner that is tuned to biosynthetic demand through chromosomal supercoiling.

Unlike genes for catabolic systems that are transcriptionally inactive in the absence of a catabolite-inducer, genes required for the biosynthesis of intermediary metabolites such as amino acids must be continuously expressed at levels tuned to the amounts of their pathway end-products. For example, operons regulated by attenuation, such as the *ilvGMEDA*, *ilvBN*, and *leu* operons of the *ilv* regulon, continuously transcribe a leader RNA whose translation into a leader polypeptide monitors the intracellular levels of their pathway end-products (46). The *ilvY* gene of this regulon also must be continuously expressed to maintain an IlvY protein-DNA complex that continuously monitors cellular levels of the $\alpha$-acetohydroxyacid isomeroreductase substrate, and adjusts the expression of the *ilvC* gene accordingly (71, 72). Because these monitoring activities that are typical of biosynthetic systems present a high energy cost to the cell, one expects global mechanisms to exist that coordinate them both with each other and with cellular demand. These global mechanisms would be expected to respond to the energy charge of the cell in a manner independent of operon-specific controls.

In this section we describe the *ilv* regulon, a well-understood system involving hierarchical levels of global and operon-specific regulation that together coordinate the biosynthesis of branched chain amino acids with the metabolic demands of the cell and its nutritional and environmental growth conditions. We begin with a brief description of the metabolic and operon-specific genetic regulatory mechanisms of this regulon that coordinate rates of carbon flow through these branched chain amino acid biosynthetic pathways with the expression levels of the genes encoding the enzymes of these pathways. Then we describe the energy charge-coupled, DNA

supercoiling-dependent, mechanisms that coordinate the basal levels of expression of operons of the *ilv* regulon with one another, and with the nutritional and environmental growth conditions of the cell, in ways that are independent of the operon-specific controls.

## Metabolic and Operon-Specific Regulatory Mechanisms of the *ilv* Regulon

METABOLIC REGULATION Carbon flow through biosynthetic pathways is regulated by end-product inhibition of the first enzyme specific for each pathway (89), so threonine deaminase is end-product inhibited by L-isoleucine, and $\beta$-isopropylmalate synthase is end-product inhibited by L-leucine (Figure 3). However, because the parallel pathways for L-valine and L-isoleucine biosynthesis are catalyzed by a single set of bifunctional enzymes, L-valine inhibition of the first enzyme specific for its synthesis would compromise the cell for L-isoleucine biosynthesis. This type of a regulatory problem is often solved by using multiple isozymes that are differentially regulated by multiple end-products. In this case, there are three $\alpha$-acetohydoxy acid synthase (AHAS) isozymes that catalyze the first step of the L-valine pathway, which is also the second step of the L-isoleucine pathway. AHAS I and AHAS III have substrate preferences for condensation of the two pyruvate molecules required for L-valine and L-leucine biosynthesis, and are both end-product inhibited by L-valine. AHAS III is also end-product inhibited by L-leucine. The third isozyme, AHAS II, which has a substrate preference for the condensation of pyruvate and $\alpha$-ketobutyrate required for L-isoleucine biosynthesis, is not inhibited by any of the branched chain amino acids. These intricate metabolic circuits respond to substrate inputs and end-product outputs to insure a balanced flow of carbon substrates through these pathways under all growth conditions.

GENETIC REGULATION The ilv regulon contains 15 structural genes organized into five operons, *ilvGMEDA*, *ilvBN*, *ilvIH*, *ilvYC*, and *leuABCD* (Figure 3). A variety of genetic regulatory mechanisms are involved in regulating these operons. The *ilvGMEDA*, *ilvBN*, and *leuABCD* operons are controlled by transcriptional attenuation mechanisms that respond to the levels of aminoacylated-, leucyl-, valyl-, and isoleucyl tRNA in the cell (31, 33, 36, 47, 57). The two remaining operons, *ilvIH* and *ilvYC*, are each regulated by other operon-specific mechanisms that are uniquely suited to the biosynthetic roles of their gene products.

The *ilvIH* structural genes encode the heterodimeric subunits of AHAS III. At the metabolic level, this isozyme is end-product inhibited by L-valine and L-leucine. At the genetic level, *ilvIH* operon expression is repressed by free L-leucine, as mediated by the global regulatory L-leucine-responsive protein, Lrp. In the absence of L-leucine, Lrp cooperatively binds to six highly degenerate sites within a 250-base pair upstream region, which activates transcription from the downstream promoter of the *ilvIH* operon. In the presence of L-leucine, this higher-order Lrp-DNA complex dissociates, and transcription from the promoter decreases (17). The supercoiling dependence of this operon has not yet been investigated.

Each of the four operons of the *ilv* regulon described to this point is regulated at the genetic level by mechanisms that respond to the intracellular levels of their pathway-specific end-products, either a free branched chain amino acid (*ilvIH*) or the cognate branched chain aminoacyl-tRNAs (*ilvGMEDA*, *ilvBN*, and *leu*). In contrast, the *ilvC* gene of the *ilvYC* operon is regulated by its substrates, not by pathway end-products. To understand this unusual situation, consider the biochemical role of the *ilvC* gene product, $\alpha$-acetohydroxyacid isomeroreductase. This enzyme catalyzes the rate-limiting step in parallel pathways, and hence must be responsive to changing concentrations of substrates produced by the three differentially regulated AHAS isozymes. By tuning its gene expression level to the concentration of substrate, it can keep carbon efficiently flowing through any or all of these parallel pathways.

The *ilvY* and *ilvC* genes comprising the *ilvYC* operon are oppositely oriented and divergently transcribed from overlapping promoter sites (Figure 4). Operon-specific regulation is mediated by the IlvY protein, the product of the *ilvY* gene. IlvY is a member of the LysR family of regulatory proteins (78). It is a homodimer that cooperatively binds to adjacent operator sites $O_1$ and $O_2$ in the divergent promoter region in a manner independent of substrate-inducer concentration. The $O_1$ site covers the region between positions $-10$ and $+1$ of the *ilvY* promoter. By modulating the expression of its own gene, IlvY effectively autoregulates its own synthesis. The $O_2$ site is located on the opposite face of the helix from $O_1$, and overlaps the $-35$ region of the *ilvC* promoter (93, 78). In the absence of substrate-inducers, the IlvY protein causes a 60° bend centered at the $-35$ region of the inactive *ilvC* promoter. When substrate-inducers bind to this preformed IlvY protein-DNA complex, the bend is relaxed and the affinity for RNA polymerase is increased 100-fold (72). The IlvY protein is autoregulated at a level that keeps these operators nearly saturated at all times. In this way, the IlvY protein-DNA complex acts as a sensor of the intracellular abundance of the $\alpha$-acetohydroxyacid isomeroreductase substrates, thereby continuously adjusting expression of the *ilvC* gene to the abundance of its substrates, synthesized by the three AHAS isozymes.

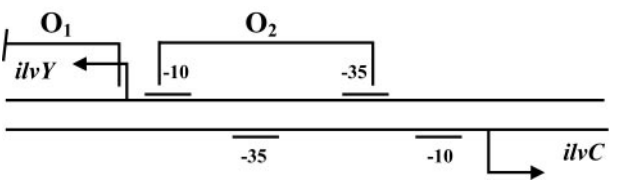


**Figure 4** The divergent promoter region of the *ilvYC* operon. Arrows identify the transcription start sites for the *ilvY* and *ilvC* genes. The $-10$ and $-35$ hexanucleotide regions of the *ilvY* and *ilvC* promoters are identified with horizontal lines. The $O_1$ (truncated) and $O_2$ operator sites are denoted by brackets.

## Global Control of Basal Level Expression of the Operons of the *ilv* Regulon

At this time, three DNA supercoiling-dependent transcriptional regulatory mechanisms related to branched chain amino acid biosynthesis and utilization have been documented (62, 61, 79, 80; M.L. Opel & G.W. Hatfield, unpublished data). These act independently of the operon-specific controls described above to coordinate the expression levels of the operons of the *ilv* regulon with each another, and with the nutritional needs and growth state of the cell in its physical environment. Two of these mechanisms (those governing the *ilvGMEDA* and *LeuV* operons) modulate basal level transcription into their structural genes by a protein-mediated (IHF or Fis) transmission of local superhelical energy from an upstream SIDD site to a downstream promoter site. This influences the rate of RNAP-promoter open complex formation and/or the rate of RNAP binding (79, 80; M. L. Opel & G.W. Hatfield, unpublished data). In the third case, the basal level expression of the *ilvC* gene is enhanced by additional local superhelical energy contributed to the *ilvC* promoter region by divergent transcription of the *ilvY* gene (61, 62, 71). In each case, local superhelical energy is provided to the promoter regions to amplify promoter activity over the entire range of global physiological superhelical densities in a manner that coordinates the basal level expression of these operons within multiple regulons, both with one another and with the energy charge of the cell. Each of these mechanisms is described separately below.

REGULATION OF TRANSCRIPTIONAL INITIATION BY PROTEIN-MEDIATED TRANSLOCATION OF SUPERHELICAL ENERGY As described above, negative superhelicity imposed on a DNA domain can drive local transitions to alternative, nonB-DNA conformations. This transition behavior can be influenced by proteins that bind at or near susceptible sites. For example, if a DNA region that would be favored to form a superhelix-induced alternative structure becomes trapped in the B-form by the binding of a protein, transition may instead occur at the next most favored site, which could be remote along the sequence. In this way, protein binding events can cause the translocation of destabilization to other sites. If this second site is in the −10 region of a promoter, this binding-induced translocation of destabilization can activate transcriptional initiation by facilitating open complex formation. This is the basic mechanism employed by the *ilvGMEDA* and *leuV* operons.

***The ilvGMEDA operon*** Nested 5′-deletions extending into the *ilv*$P_G$ promoter of the *ilvGMEDA* operon identified an upstream activating region (UAS). This UAS contains a high-affinity Integration Host Factor (IHF) target binding site located 92 base pairs upstream from the transcriptional start site. Biochemical and genetic experiments have shown that IHF binding to this site, both in vivo and in vitro on a supercoiled DNA template, causes a fivefold activation of transcription from the downstream *ilv*$P_G$ promoter (97). Several experimental approaches have established that this activation occurs in the absence of specific protein interactions

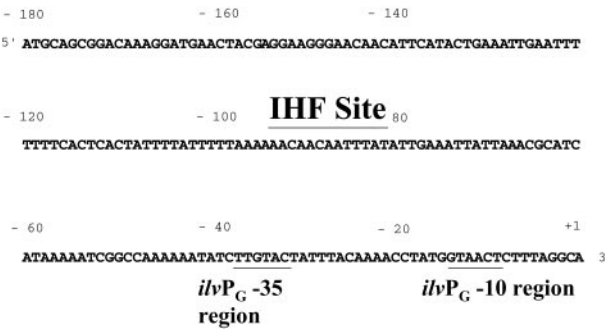


**Figure 5** DNA sequence of the *ilv*$P_G$ UAS-promoter region. See text for discussion.

between IHF and RNA polymerase; it is not the consequence of a DNA looping mechanism; and it requires a superhelical DNA template (64, 65, 79–81).

The possible presence of a SIDD site in the UAS was first suggested by the observation that the base-pair composition of the *ilv*$P_G$ promoter-regulatory region is exceptionally A + T rich, the 80-bp segment from bp positions −67 to −153 being approximately 88% ( A + T) (Figure 5). In order to determine if this region does indeed contain a SIDD site, SIDD profiles were calculated for the pBR322-based plasmid pDHΔwt (79). This plasmid contains the *ilv*$P_G$ promoter region from positions −248 to +6, together with transcriptional terminators located downstream from the *ilv*$P_G$ start site. The results of these calculations are presented as destabilization profiles in Figure 2. Subsequent chemical probing experiments confirmed that, in the absence of IHF, the UAS region was indeed destabilized at the superhelix densities where activation occurs. Moreover, IHF binding in the UAS region of a superhelical DNA template resulted in the transmission of this duplex destabilization into the −10 region of the downstream *ilv*$P_G$ promoter site. Abortive transcription assays showed that this DNA structural change at the downstream promoter site is correlated both with an increase in the rate of open complex formation, and with a concomitant increase in the rate of transcriptional initiation (79).

These results suggested that a novel, protein-mediated, DNA supercoiling-dependent, DNA structural transmission mechanism regulates basal level transcription from the *ilv*$P_G$ promoter (Figure 6). In this mechanism, the binding of IHF prevents superhelical destabilization at the SIDD site in the A + T-rich UAS region, which transfers that destabilization to the −10 region of the nearby *ilv*$P_G$ promoter. This explanation accounts for the DNA supercoiling-dependence of the IHF-mediated activation of transcription from this promoter, and for the fact that this activation occurs in the absence of specific interactions with RNAP.

According to this mechanism, the primary determinant for IHF-mediated activation is predicted to be superhelically induced DNA destabilization. Neither specific DNA sequences nor specific IHF-RNAP interactions are required. To

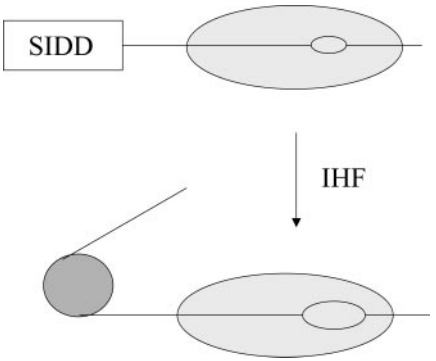


**Figure 6** DNA structural transmission mechanism for protein-mediated DNA superhelical-dependent transcriptional activation of the *ilv*PG promoter. In this model, IHF binding prevents the superhelical destabilization of the SIDD region of the UAS, which transfers the transition to the −10 region of the downstream *ilv*PG promoter, thereby facilitating open complex formation during transcription initiation. Open complex formation is represented by a bubble (*denoted by an oval*) in the duplex DNA (*designated by a single line*). Increased open complex formation is indicated by the increased size of the bubble.



directly test this prediction, the sequence (CG)13AATT(CG)22 was inserted into plasmid pDHΔwt approximately 500 base pairs upstream from the UAS-*ilv*$P_G$ promoter to yield plasmid pSSΔZ (80). This sequence is susceptible to a superhelically induced transition to Z-form DNA (69). Because the B-Z transition at this remote site competes effectively with the SIDD site in the UAS, this insertion can alter the destabilization properties of the UAS without changing the DNA sequence in any part of the *ilv*$P_G$ regulatory-promoter region. This transition was shown to absorb 13 negative superhelical turns, thereby relaxing the global superhelix density of the remainder of the supercoiled DNA template by a corresponding amount. Since this B-Z transition occurs at a lower threshold superhelical density ($\sigma = -0.025$) than does the destabilization of UAS ($\sigma = -0.038$), it inhibits UAS destabilization until approximately 13 additional negative superhelical turns have been added to the DNA template. If the energy required for IHF-mediated transcriptional activation is indeed derived by IHF binding-induced transfer from the destabilized UAS region, then the superhelicity required for IHF activation in the pSSΔZ plasmid should be offset by approximately 13 turns. The results of transcription assays on DNA templates of defined superhelix densities showed this to be the case. The superhelicities required both for half-maximal basal level and for IHF-activated transcription were indeed offset by 13 turns (79). This experiment clearly demonstrated that IHF-mediated transcriptional activation of the *ilv*$P_G$ promoter is solely DNA supercoiling-dependent and confirmed the predictions of the protein-mediated, DNA structural transmission mechanism of transcriptional activation proposed above.

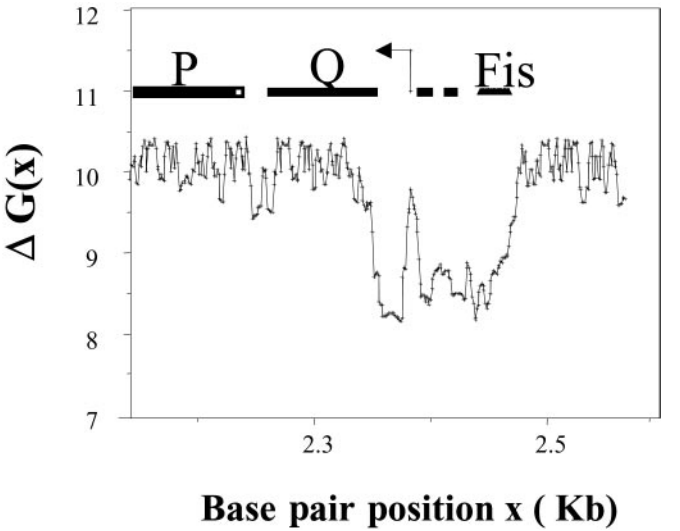


**Figure 7** SIDD profile of the *leuV* promoter region. Predicted free energies G(x) for duplex destabilization at a superhelical density of $\sigma = -0.055$ at base position x are expressed in kcal/mol/bp. The positions of relevant features are indicated above the SIDD plot. The transcriptional start site is indicated by an arrow. The locations of the Fis binding site, the −10 and −35 hexanucleotide regions, and the promoter-proximal structural genes are identified by thick horizontal lines. The discriminator is the peak of duplex stability in the SIDD region near the transcriptional start site. The three identical structural genes for the $tRNA^{Leu1}$ isoacceptor of the *leuV* operon are *leuQ*, *leuP*, and *leuV* (not shown).

***The leuV operon*** The *leuV* operon of *E. coli* encodes three of the five genes for $tRNA^{Leu1}$ isoacceptor. Like other stable RNA-encoding genes, it has a strong promoter, with near-consensus RNAP recognition sequences and a G + C-rich discriminator region located between base pair positions −8 and +1 (Figure 7). Transcription from the *leuV* promoter is enhanced by a third RNAP recognition element located between base pairs −39 to −47. This A + T-rich UP sequence makes contacts with the $\alpha$-subunits of RNAP, stabilizes closed complex formation, and activates *leuV* expression more than tenfold (30, 67, 73). The upstream activating sequence (UAS) of this promoter contains a Fis protein binding site centered at base pair position −71. Fis binding to this site enhances *leuV* expression an additional threefold (74).

A large body of evidence demonstrates that Fis is a class-I activator, enhancing transcription by increasing RNAP binding affinity through direct contacts with the C-terminal domain of its $\alpha$-subunits (12, 13, 15, 60, 98). However, other mechanisms also are involved in Fis activation of stable RNA promoters. For example, Fis binding has been shown to increase the rate of promoter clearance at the *rrnD*

promoter (77). Muskhelishvili & Travers have shown that Fis activates transcription from the *tyrT* promoter by enhancing the rate of open complex formation and promoter clearance, as well as RNA polymerase binding affinity (56). Most recently, Opel et al. have obtained evidence that basal level expression of the *leuV* promoter is also activated by a Fis-mediated translocation of superhelical energy mechanism similar to the IHF-mediated, DNA supercoiling-dependent mechanism previously described for the *ilv*$P_G$ operon (M.L. Opel & G.W. Hatfield, unpublished data).

A SIDD profile of the promoter-regulatory region of the *leuV* operon is shown in Figure 8. At a midphysiological superhelical density of $\sigma = -0.05$, this region is predicted to be destabilized from base pair positions +43 to −94. The Fis protein binding site is located in the upstream region of this SIDD site centered at base-pair position −72. An interesting feature of the SIDD profile is the sharp peak of duplex stability at the G + C-rich discriminator region between positions +1 and −8. The presence of this region of high duplex stability at a site where strand separation must occur predicts a high energy of activation for open complex formation. Indeed, *leuV* transcription is exceptionally sensitive to negative DNA supercoiling, increasing over 100-fold from its lowest level on a relaxed DNA template to its highest level on a more supercoiled DNA template. The SIDD profile further predicts that the upstream region of the SIDD site should be stabilized by Fis binding to its target site, and that this binding should destabilize the downstream portion of this SIDD region containing the *leuV* promoter sequences. This prediction was confirmed with in vitro transcription assays and $KMnO_4$ structural probing experiments performed with supercoiled DNA template topoisomers in the presence or absence of Fis and/or RNAP. These experiments showed that Fis binding enhances the rate of open complex formation in a DNA supercoiling-dependent manner. At subsaturating concentrations of RNAP, Fis activation is facilitated both by protein-protein interactions between Fis and RNAP, and by DNA supercoiling-dependent enhancement of open complex formation. At saturating RNAP concentrations, only the enhancement of open complex is seen. Mutant Fis proteins that do not form contacts with the $\alpha$-subunits of RNAP but bind to the target site with wild-type affinities were used to demonstrate that this activation does not require Fis-RNAP interactions (M.L. Opel & G.W. Hatfield, unpublished data). These mutant proteins maintain their ability to facilitate DNA supercoiling-dependent enhancement of open complex formation. Thus, Fis activation of basal level transcription from the *leuV* promoter involves at least two mechanisms: stabilization of the closed complex by protein interactions with the $\alpha$-subunit of RNAP, and increasing open complex formation by translocation of superhelical energy from the upstream portion of the SIDD region containing the Fis target site to the downstream portion of this region containing the *leuV* promoter sequences.

Unlike the *ilv*$P_G$ promoter that reaches its peak transcriptional activity at a high physiological superhelical density near $\sigma = -0.10$, transcriptional activity of the *leuV* promoter peaks at a superhelical density near $\sigma = -0.07$ and decreases thereafter to the level observed on a relaxed DNA template. This decrease in

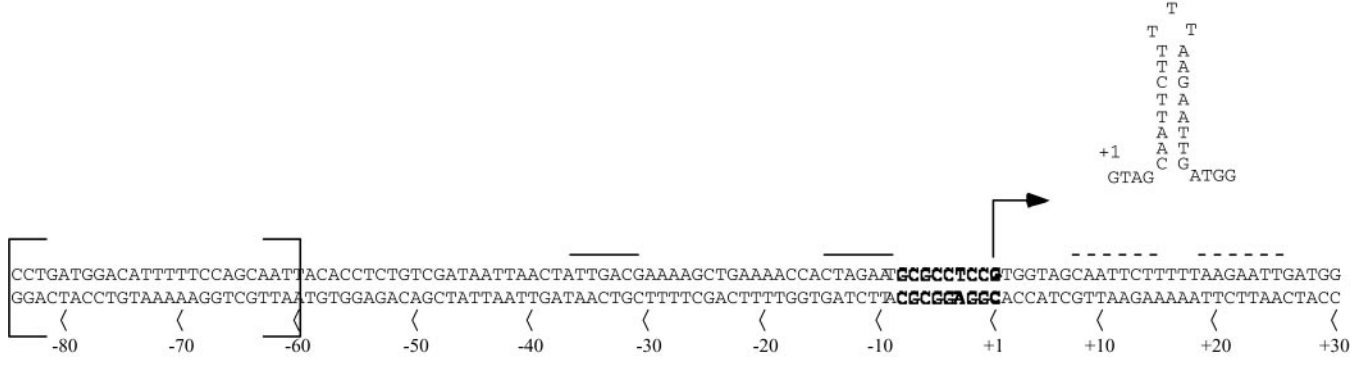


**Figure 8** Nucleotide sequence of the *leuV* promoter region from −84 to +30. The transcriptional start site for *leuV* is identified by an arrow. The −10 and −35 hexanucleotide regions are identified by horizontal lines. The Fis binding site is denoted by brackets. The discriminator is identified by bold-face type. An inverted repeat is identified by a dashed line. Cruciform extrusion at the inverted repeat is depicted above the inverted repeat.

transcriptional activity at high superhelix densities is accompanied by the formation of a cruciform structure located between base-pair positions +8 to +26 relative to the transcriptional start site. Structural probing experiments with DNA topoisomers show that the threshold superhelical density required for extrusion of this cruciform structure is $\sigma = -0.069$, the same superhelical density beyond which transcription from the *leuV* promoter decreases. This suggests that, at superhelical densities beyond this threshold, the global and Fis-transferred local superhelical energy in the promoter region is absorbed by the cruciform. This transition has two effects on *leuV* gene expression: It usurps the local superhelical energy that would otherwise have been transferred to the promoter region by Fis for open complex formation, and it physically blocks RNAP binding.

A schematic diagram illustrating how Fis binding and DNA supercoiling-induced structural transitions regulate transcription from the *leuV* promoter is presented in Figure 9. At the low physiological superhelical densities typical of stationary phase growth, transcription from the *leuV* promoter is very low. This is due to the energy barrier for open complex formation caused by the G + C-rich discriminator near the transcription start site. Under these conditions, Fis can activate transcription about threefold by increasing RNAP binding through interactions with its $\alpha$-subunits. Since this low level of global superhelicity is insufficient for SIDD site formation, no additional activation by Fis-mediated translocation of superhelical energy to the promoter site is possible. As the global superhelicity of the DNA template is increased to the midphysiological range, the energy barrier for open complex formation caused by the discriminator is decreased and transcription

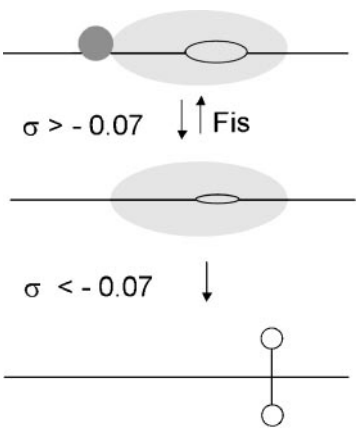


**Figure 9** DNA supercoiling and Fis-mediated regulation of the *leuV* promoter. RNA polymerase is represented by a large oval. Fis is represented by a small circle. Open complex formation is represented by a bubble in the duplex DNA (*designated by a single line*). The strength of open complex formation is indicated by the size of the bubble. A cruciform is indicated by a stem-loop structure. See text for details.

increases up to 100-fold. As the SIDD site now is present, Fis binding can activate transcription both by enhancing RNAP binding and by protein-mediated translocation of superhelical energy to the promoter region. As the global superhelical density passes beyond the midphysiological range, the superhelical energy at the SIDD site is usurped by formation of the cruciform structure near the transcription start site, and RNAP binding is inhibited.

### REGULATION OF TRANSCRIPTIONAL INITIATION BY TRANSCRIPTIONAL COUPLING

LysR-type regulated operons are the largest class of positively regulated operons, and are found in many prokaryotic species (37, 78). The prevalence of a divergent gene arrangement among the LysR-type regulated operons suggests an evolutionary conservation of potential regulatory significance. According to the twin-domain model (49), a local domain having a high level of DNA supercoiling can be generated between, and influence the activities of, divergently transcribed promoters. Mojica & Higgins have used in vivo psoralen cross-linking techniques to demonstrate that localized domains of increased negative DNA supercoiling are indeed generated upstream from an actively transcribed promoter (55). They demonstrated DNA supercoiling-mediated transcriptional coupling between the divergently oriented *tetA* and mutant *leu-500* promoters.

The *ilvYC* operon of *E. coli* K-12 is a prototypic LysR-type regulated system (78, 37). Rhee et al. used double-reporter gene constructs to provide the first in vivo evidence for transcriptional coupling in a naturally occurring system, the *ilvYC* operon of *E. coli* (71). They showed that each of these promoters is intrinsically sensitive to global DNA supercoiling, and that a 13-fold decrease in transcriptional activity from the *ilvY* promoter results in an 11-fold decrease in transcription from the divergent *ilvC* promoter. This transcriptional coupling was shown to be the consequence of transcription-induced negative DNA supercoiling. In this situation, a highly stressed local topological domain is created in the promoter region by divergent transcription, in which the total supercoiling is the sum of the basal, global superhelicity plus the supercoiling arising from divergent transcription. This suggested a strategy to document and characterize transcriptional coupling in a purified in vitro transcription system. When a set of DNA topoisomer templates containing the wild-type, divergently oriented *ilvY* and *ilvC* promoters were transcribed in a purified system, optimal transcriptional activity was observed to occur at superhelical density $\sigma = -0.065$ for the *ilvY* promoter, and $\sigma = -0.11$ for the *ilvC* promoter (71). If the level of negative DNA supercoiling in the divergently transcribed promoter region were in fact proportional to the sum of transcription-induced (local) DNA supercoiling and the global superhelical density of the DNA template, then a decrease in transcription (hence in transcription-induced supercoiling) from either promoter should require a compensating increase in global DNA supercoiling to maintain maximal transcription from its divergently transcribed partner. Conversely, an increase in transcription from either promoter should require a corresponding decrease in global supercoiling to maintain maximal transcription from the divergently oriented other promoter.

Further, the twin-domain model of Liu & Wang (49) also predicts that the levels of transcription-induced negative DNA supercoiling in the divergent promoter domain region should be proportional to the lengths of the transcripts. Opel & Hatfield have performed in vitro experiments using a purified transcription system and DNA topoisomer templates containing the genes of the *ilvYC* operon that confirm both of these predictions (62).

## SUMMARY AND PERSPECTIVES



We have proposed a model in which chromosomal superhelicity serves as a global regulator of gene expression in *E. coli*, tuning expression patterns across multiple operons, regulons, and stimulons to suit the current growth state, nutritional requirements, and environmental conditions of the cell. Because the level of DNA superhelicity varies with the cellular energy charge, it can change rapidly—often in less than a minute—in response to a wide variety of altered cellular and environmental conditions. This is a global alteration, affecting the entire chromosome and the expression levels of all operons whose promoters are sensitive to superhelicity. In this way, the global pattern of gene expression may be dynamically tuned to changing needs of the cell under a wide variety of circumstances.

This model is illustrated by the DNA supercoiling-dependent mechanisms that control basal level expression of the *ilvYC* and *ilvGMEDA* operons. The basal expression levels of these operons are coordinated by distinct, DNA supercoiling-dependent mechanisms. However, both mechanisms have the same effect—to provide additional local superhelical energy to the promoter regions to lower the energy of activation for open complex formation and to increase the rate of transcription initiation from the promoter. This additional energy is important for biosynthetic promoters that must coordinate their transcriptional activity over the entire superhelical density range. Such promoters must possess a high superhelical density-activity optimum. This makes them, by definition, intrinsically weak promoters because of the high energy of activation required for open complex formation. However, by supplementing the global superhelicity of the chromosome with locally generated superhelicity, their promoter activities can be amplified to a level sufficient for their basal level functions, while at the same time maintaining their ability to increase their activity over the entire range of physiological superhelical densities. In the case of the *ilvGMEDA* operon, this is accomplished by IHF-mediated translocation of superhelical energy from an upstream SIDD site to the downstream promoter region. In the case of the *ilvYC* operon, this is accomplished by the additional local superhelical energy contributed to the *ilvC* promoter region by the divergently transcribed *ilvY* promoter. In both cases, promoter activities increase tenfold over the physiological superhelical densities that the chromosome encounters. Therefore, these different mechanisms serve to coordinate the basal level of expression of these two operons of a common biosynthetic pathway with one another and with the nutritional and physical signals that determine energy charge of the cell.

Since free amino acids are penultimate end-products for protein synthesis, it is perhaps not surprising that the operon encoding genes for the major leucyl-tRNA isoacceptors is also regulated by a similar DNA supercoiling-dependent mechanism. However, in this case the global regulatory, abundant, DNA architectural protein that mediates translocation of superhelical energy from the upstream portion of the SIDD site to the downstream promoter is Fis rather than IHF. Since the *leuV* operon is concerned with consumption rather than production, it has evolved a supercoiling-dependent mechanism to augment L-leucine conservation during periods of stress where the superhelical density of the chromosome is high. Under these conditions, *leuV* expression is controlled by a complex competition between an activating duplex destabilization, which is favored at midphysiological superhelical densities, and an inhibiting cruciform extrusion that occurs at more extreme superhelicity levels.

Note that these DNA supercoiling-dependent global regulatory mechanisms described here affect operon expression in a manner independent of the specific controls whereby these operons respond to the in vivo levels of metabolically important, small-molecule coregulators and other cellular signals. Instead, these mechanisms serve to adjust the basal level, or capacity, for operon expression according to changes in DNA supercoiling that reflect the energy charge of the cell.

There are many possible approaches for coupling superhelicity in a physiologically meaningful way. We have documented two distinct types of mechanisms—transcriptional coupling at divergent promoters, and a novel, binding-induced transmission of destabilization. The prevalence of divergently oriented ORFs in *E. coli* suggests that some form of transcriptional coupling could represent a widespread regulatory mechanism. Although to date the transmission-of-destabilization mechanism has only been demonstrated to regulate two operons, other evidence suggests that this too could be a widely used regulatory strategy. Perhaps the most intriguing clue is the fact that both operons are regulated by highly abundant DNA architectural proteins, IHF and Fis, that bind to hundreds of sites on the *E. coli* chromosome. Another intriguing observation is that recently completed calculations of the SIDD profiles of the entire *E. coli* chromosome suggest the presence of approximately 1100 SIDD sites. Furthermore, the great majority of these sites are found in intergenic regions close to promoters (C. J. Benham, unpublished data).

Certainly, much work remains to assess the generality of our proposed model. Careful experimental evaluation of the structure and sequence motifs of IHF and Fis binding sites will enable us to search the *E. coli* genome for binding sites, and cross-linking immunoprecipitation experiments with DNA microarrays will allow an estimation of in vivo occupancy at these sites. These results together with the results of gene expression profiling experiments will allow us identify genes in the vicinity of IHF or Fis binding sites that overlap SIDD sites. This approach will illuminate which genes are regulated by supercoiling-dependent mechanisms, which genes are regulated by IHF or Fis, and which among these are also regulated in a supercoiling-dependent manner.

**The *Annual Review of Genetics* is online at http://genet.annualreviews.org**

## LITERATURE CITED


1. Atkinson DE. 1965. Biological feedback control at the molecular level. *Science* 150:851–57
2. Atkinson DE. 1969. Regulation of enzyme function. *Annu. Rev. Microbiol.* 23:47–68
3. Benham CJ. 1979. Torsional stress and local denaturation in supercoiled DNA. *Proc. Natl. Acad. Sci. USA* 76:3870–74
4. Benham CJ. 1980. Theoretical analysis of transitions between B- and Z-conformations in torsionally stressed DNA. *Nature* 286:637–38
5. Benham CJ. 1990. Transitions in superhelical DNA molecules of specified sequence. *J. Chem. Phys.* 92:6294–305
6. Benham CJ. 1992. Energetics of the strand separation transition in superhelical DNA. *J. Mol. Biol.* 225:835–47
7. Benham CJ. 1993. Sites of predicted stress-induced DNA duplex destabilization occur preferentially at regulatory loci. *Proc. Natl. Acad. Sci. USA* 90:2999–3003
8. Benham CJ. 1996. Duplex destabilization in superhelical DNA is predicted to occur at specific transcriptional regulatory regions. *J. Mol. Biol.* 255:425–34
9. Benham C, Kohwi-Shigematsu T, Bode J. 1997. Stress-induced duplex DNA destabilization in scaffold/matrix attachment regions. *J. Mol. Biol.* 274:181–96
10. Bensaid A, Almeida A, Drlica K, Rouviere J. 1996. Cross-talk between topoisomerase I and HU in *Escherichia coli*. *J. Mol. Biol.* 256:292–300
11. Bliska JB, Cozzarelli NR. 1987. Use of site-specific recombination as a probe of DNA structure and metabolism in vivo. *J. Mol. Biol.* 194:205–18
12. Bokal AJ, Ross W, Gaal T, Johnson RC, Gourse RL. 1997. Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the *Escherichia coli* rrnB P1 promoter. *EMBO J.* 16:154–62
13. Bokal AJ, Ross W, Gourse RL. 1995. The transcriptional activator protein FIS: DNA interactions and cooperative interactions with RNA polymerase at the *Escherichia coli* rrnB P1 promoter. *J. Mol. Biol.* 245:197–207
14. Borowiec JA, Gralla JD. 1985. Supercoiling response of the lac ps promoter in vitro. *J. Mol. Biol.* 184:587–98
15. Breslauer KJ, Frank R, Blocker H, Marky LA. 1986. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* 83:3746–50
16. Calladine CR. 1997. *Understanding DNA: The Molecule and How it Works*. New York: Academic
17. Calvo JM, Matthews RG. 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol. Rev.* 58:466–90
18. Chapman AG, Atkinson DE. 1977. Adenine nucleotide concentrations and turnover rates. Their correlation with biological activity in bacteria and yeast. *Adv. Microb. Physiol.* 15:253–306
19. Chapman AG, Fall L, Atkinson DE. 1971. Adenylate energy charge in *Escherichia coli* during growth and starvation. *J. Bacteriol.* 108:1072–86
20. Cortassa S, Aon MA. 1993. Altered topoisomerase activities may be involved in the regulation of DNA supercoiling in aerobic-anaerobic transitions in *Escherichia coli*. *Mol. Cell Biochem.* 126:115–24
21. Cozzarelli NR. 1980. DNA gyrase and the supercoiling of DNA. *Science* 207:953–60
22. Dai X, Greizerstein MB, Nadas Chinni K, Rothman Denes LB. 1997. Supercoil-induced extrusion of a regulatory DNA hairpin. *Proc. Natl. Acad. Sci. USA* 94: 2174–79
23. Dai X, Rothman-Denes LB. 1999. DNA

structure and transcription. *Curr. Opin. Microbiol.* 2:126–30
24. Deibler RW, Rahmati S, Zechiedrich EL, 2001. Topoisomerase IV, alone, unknots DNA in *E. coli*. *Genes Dev.* 15:748–61
25. DiNardo S, Voelkel KA, Sternglass R, Reynolds AE, Wright A. 1983. *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations at or near DNA gyrase genes. *Cold Spring Harbor Symp. Quant. Biol.* 47:779–84
26. Drew HR, Weeks JR, Travers AA. 1985. Negative supercoiling induces spontaneous unwinding of a bacterial promoter. *EMBO J.* 4:1025–32
27. Drlica K. 1992. Control of bacterial DNA supercoiling. *Mol. Microbiol.* 6:425–33
28. Drlica K, Franco RJ. 1988. Inhibitors of DNA topoisomerases. *Biochemistry* 27: 2253–59
29. Drlica K, Zhao X. 1997. DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiol. Mol. Biol. Rev.* 61:377–92
30. Estrem ST, Gaal T, Ross W, Gourse RL. 1998. Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. USA* 95:9761–66
31. Friden P, Newman T, Freundlich M. 1982. Nucleotide sequence of the ilvB promoter-regulatory region: a biosynthetic operon controlled by attenuation and cyclic AMP. *Proc. Natl. Acad. Sci. USA* 79:6156–60
32. Fye RM, Benham CJ. 1999. Exact method for numerically analyzing a model of local denaturation in superhelically stressed DNA. *Phys. Rev.* 59:3408–26
33. Gemmill RM, Wessler SR, Keller EB, Calvo JM. 1979. leu operon of *Salmonella typhimurium* is controlled by an attenuation mechanism. *Proc. Natl. Acad. Sci. USA* 76:4941–45
34. Giaever GN, Snyder L, Wang JC. 1988. DNA supercoiling in vivo. *Biophys. Chem.* 29:7–15
35. Gosink KK, Ross W, Leirmo S, Osuna R, Finkel SE, et al. 1993. DNA binding and bending are necessary but not sufficient for Fis-dependent activation of rrnB P1. *J. Bacteriol.* 175:1580–89
36. Hauser CA, Hatfield GW. 1983. Nucleotide sequence of the ilvB multivalent attenuator region of *Escherichia coli* K12. *Nucleic Acids Res.* 11:127–39
37. Henikoff S, Haughn GW, Calvo JM, Wallace JC. 1988. A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* 85:6602–6
38. Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, et al. 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. *Cell* 52:569–84
39. Hsieh LS, Rouviere-Yaniv J, Drlica K. 1991. Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. *J. Bacteriol.* 173:3914–17
40. Htun H, Dahlberg JE. 1989. Topology and formation of triple-stranded H-DNA. *Science* 243:1571–76
41. Jensen PR, Loman L, Petra B, Vanderwe C, Westerhouse HV. 1995. Energy buffering of DNA structure fails when *Escherichia coli* runs out of substrate. *J. Bacteriol.* 177:3420–26
42. Jones PG, Inouye M. 1994. The cold-shock response—a hot topic. *Mol. Microbiol.* 11:811–18
43. Jones PG, VanBogelen RA, Neidhardt FC. 1987. Induction of proteins in response to low temperature in *Escherichia coli*. *J. Bacteriol.* 169:2092–95
44. Kowalski D, Natale DA, Eddy MJ. 1988. Stable DNA unwinding, not "breathing," accounts for single-strand-specific nuclease hypersensitivity of specific A+T-rich sequences. *Proc. Natl. Acad. Sci. USA* 85: 9464–68
45. Kusano S, Ding QQ, Fujita N, Ishihama A. 1996. Promoter selectivity of *Escherichia coli* RNA polymerase E sigma 70 and E sigma 38 holoenzymes. Effect of DNA supercoiling. *J. Biol. Chem.* 271:1998–2004
46. Landick R, Yanofsky C. 1987. Transcription attenuation. See Ref. 58, 1:1276–301

47. Lawther RP, Hatfield GW. 1980. Multivalent translational control of transcription termination at attenuator of ilvGEDA operon of *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. USA* 77:1862–66
48. Lilley DM. 1980. The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. *Proc. Natl. Acad. Sci. USA* 77:6468–72
49. Liu LF, Wang JC. 1987. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. USA* 84:7024–27
50. Liu Y, et al. 2001. DNA supercoiling allows enhancer action over a large distance. *Proc. Natl. Acad. Sci. USA* 98:14883–88
51. Luttinger A. 1995. The twisted 'life' of DNA in the cell: bacterial topoisomerases. *Mol. Microbiol.* 15:601–6
52. McEachern F, Fisher LM. 1989. Regulation of DNA supercoiling in *Escherichia coli*: genetic basis of a compensatory mutation in DNA gyrase. *FEBS Lett.* 253:67–70
53. Menzel R, Gellert M. 1983. Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell* 34:105–13
54. Michelotti GA, Michelotti EF, Pullner A, Duncan RC, Eick D. 1996. Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Mol. Cell Biol.* 16:2656–69
55. Mojica FJ, Higgins CF. 1996. Localized domains of DNA supercoiling: topological coupling between promoters. *Mol. Microbiol.* 22:919–28
56. Muskhelishvili G, Buckle M, Heumann H, Kahmann R, Travers AA. 1997. FIS activates sequential steps during transcription initiation at a stable RNA promoter. *EMBO J.* 16:3655–65
57. Nargang FE, Subrahmanyam CS, Umbarger HE. 1980. Nucleotide sequence of ilvGEDA operon attenuator region of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 77:1823–27
58. Neidhardt FC, ed. 1987. *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* Washington, DC: ASM
59. Neidhardt FC, Savageau MA. 1996. Regulation beyond the operon. See Ref. 58, 2:1310–24
60. Newlands JT, Josaitis CA, Ross W, Gourse RL. 1992. Both fis-dependent and factor-independent upstream activation of the rrnB P1 promoter are face of the helix dependent. *Nucleic Acids Res.* 20:719–26
61. Opel ML, Arfin SM, Hatfield GW. 2001. The effects of DNA supercoiling on the expression of operons of the ilv regulon of *Escherichia coli* suggest a physiological rationale for divergently transcribed operons. *Mol. Microbiol.* 39:1109–15
62. Opel ML, Hatfield GW. 2001. DNA supercoiling-dependent transcriptional coupling between the divergently transcribed promoters of the ilvYC operon of *Escherichia coli* is proportional to promoter strengths and transcript lengths. *Mol. Microbiol.* 39:191–98
63. Oram M, Fisher LM. 1992. An *Escherichia coli* DNA topoisomerase I mutant has a compensatory mutation that alters two residues between functional domains of the DNA gyrase A protein. *J. Bacteriol.* 174:4175–78
64. Parekh BS, Hatfield GW. 1996. Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model. *Proc. Natl. Acad. Sci. USA* 93:1173–77
65. Parekh BS, Sheridan SD, Hatfield GW. 1996. Effects of integration host factor and DNA supercoiling on transcription from the ilvPG promoter of *Escherichia coli*. *J. Biol. Chem.* 271:20258–64
66. Pettijohn DE, Pfenninger O. 1980. Supercoils in prokaryotic DNA restrained in vivo. *Proc. Natl. Acad. Sci. USA* 77:1331–35
67. Pokholok DK, Redlak M, Turnboug CL, Dylla S, Holmes WM. 1999. Multiple mechanisms are used for growth rate and stringent control of leuV transcriptional initiation in *Escherichia coli*. *J. Bacteriol.* 181:5771–82
68. Pruss GJ, Manes SH, Drlica K. 1982. *Escherichia coli* DNA topoisomerase I

mutants: increased supercoiling is corrected by mutations near gyrase genes. *Cell* 31:35–42
69. Rahmouni AR, Wells RD. 1989. Stabilization of Z DNA in vivo by localized supercoiling. *Science* 246:358–63
70. Raji A, Zabel DJ, Laufer CS, Depwe RE. 1985. Genetic analysis of mutations that compensate for loss of *Escherichia coli* DNA topoisomerase I. *J. Bacteriol.* 162:1173–79
71. Rhee KY, Opel M, Ito E, Hung SP, Arfin SM, Hatfield GW. 1999. Transcriptional coupling between the divergent promoters of a prototypic LysR-type regulatory system, the ilvYC operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 96:14294–99
72. Rhee KY, Senear DF, Hatfield GW. 1998. Activation of gene expression by a ligand-induced conformational change of a protein-DNA complex. *J. Biol. Chem.* 273:11257–66
73. Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, et al. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 262:1407–13
74. Ross W, Salomon J, Holmes WM, Gourse RL. 1999. Activation of *Escherichia coli* leuV transcription by FIS. *J. Bacteriol.* 181:3864–68
75. Rouviere-Yaniv J. 1978. Localization of the HU protein on the *Escherichia coli* nucleoid. *Cold Spring Harbor Symp. Quant. Biol.* 42 (Pt. 1):439–47
76. Rouviere-Yaniv J, Yaniv M, Germond JE. 1979. *E. coli* DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA. *Cell* 17:265–74
77. Sander P, Langert W, Mueller K. 1993. Mechanisms of upstream activation of the rrnD promoter P1 of *Escherichia coli*. *J. Biol. Chem* 268:16907–16
78. Schell MA. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* 47:597–626
79. Sheridan SD, Benham CJ, Hatfield GW. 1998. Activation of gene expression by a novel DNA structural transmission mechanism that requires supercoiling-induced DNA duplex destabilization in an upstream activating sequence. *J. Biol. Chem.* 273:21298–308
80. Sheridan SD, Benham CJ, Hatfield GW. 1999. Inhibition of DNA supercoiling-dependent transcriptional activation by a distant B-DNA to Z-DNA transition. *J. Biol. Chem.* 274:8169–74
81. Sheridan SD, Opel ML, Hatfield GW. 2001. Activation and repression of transcription initiation by a distant DNA structural transition. *Mol. Microbiol.* 40:684–90
82. Sinden RR, Kochel TJ. 1987. Reduced 4,5′,8-trimethylpsoralen cross-linking of left-handed Z-DNA stabilized by DNA supercoiling. *Biochemistry* 26:1343–50
83. Singleton CK, Klysik J, Stirdiva SM, Wells RD. 1982. Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions. *Nature* 299:312–16
84. Steck TR, Franco RJ, Wang JY, Drlica K. 1993. Topoisomerase mutations affect the relative abundance of many *Escherichia coli* proteins. *Mol. Microbiol.* 10:473–81
85. Steger G. 1994. Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Res.* 22:2760–68
86. Sun HZ, Mezei M, Fye R, Benham CJ. 1995. Monte Carlo analysis of conformational transitions in superhelical DNA. *J. Chem. Phys.* 103:8653–65
87. Tse-Dinh YC. 1985. Regulation of the *Escherichia coli* DNA topoisomerase I gene by DNA supercoiling. *Nucleic Acids Res.* 13:4751–63
88. Tse-Dinh YC, Beran RK. 1988. Multiple promoters for transcription of the *Escherichia coli* DNA topoisomerase I gene and their regulation by DNA supercoiling. *J. Mol. Biol.* 202:735–42
89. Umbarger HE. 1996. Biosynthesis of the branched-chain amino acids. See Ref. 58, 2:442–57
90. Urios A, Herrera G, Aleixand V, Blanco

M. 1990. Expression of the recA gene is reduced in *Escherichia coli* topoisomerase I mutants. *Mutat. Res.* 243:267–72

91. Varshavsky AJ, Nedospa SA, Bakayev VV, Bakayeva TG, Georgiev GP. 1977. Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res.* 4:2725–45
92. Wang JY, Syvanen M. 1992. DNA twist as a transcriptional sensor for environmental changes. *Mol. Microbiol.* 6:1861–66
93. Wek RC, Hatfield GW. 1986. Nucleotide sequence and in vivo expression of the ilvY and ilvC genes in *Escherichia coli* K12. Transcription from divergent overlapping promoters. *J. Biol. Chem.* 261:2441–50
94. Wek RC, Hatfield GW. 1988. Transcriptional activation at adjacent operators in the divergent-overlapping ilvY and ilvC promoters of *Escherichia coli*. *J. Mol. Biol.* 203:643–63
95. Westerhoff HV, van Workum M. 1990. Control of DNA structure and gene expression. *Biomed. Biochim. Acta* 49:839–53
96. Westerhoff HV, Odea MH, Maxwell A, Gellert M. 1988. DNA supercoiling by DNA gyrase. A static head analysis. *Cell Biophys.* 12:157–81
97. Winkelman JW, Hatfield GW. 1990. Characterization of the integration host factor binding site in the ilvPG1 promoter region of the ilvGMEDA operon of *Escherichia coli*. *J. Biol. Chem.* 265:10055–60
98. Zacharias M, Goringer HU, Wagner R. 1992. Analysis of the Fis-dependent and Fis-independent transcription activation mechanisms of the *Escherichia coli* ribosomal RNA P1 promoter. *Biochemistry* 31: 2621–28
99. Zechiedrich EL, Khodursk AB, Bachelli S, Schneider R, Chem DR, et al. 2000. Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J. Biol. Chem* 275:8103–13

*Annual Review of Genetics*
*Volume 36, 2002*

# CONTENTS