Quantitative contribution of the spacer length in the supercoiling-sensitivity of bacterial promoters

Raphaël Forquet, William Nasser, Sylvie Reverchon, Sam Meyer

Quantitative contribution of the spacer length in the supercoiling-sensitivity of bacterial promoters

Raphaël Forquet [ORCID]**, William Nasser** [ORCID]**, Sylvie Reverchon** [ORCID] **and Sam Meyer** [ORCID]*

Université de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, CNRS, UMR5240 MAP, F-69622, France

ABSTRACT

DNA supercoiling acts as a global transcriptional regulator in bacteria, but the promoter sequence or structural determinants controlling its effect remain unclear. It was previously proposed to modulate the torsional angle between the −10 and −35 hexamers, and thereby regulate the formation of the closed-complex depending on the length of the 'spacer' between them. Here, we develop a thermodynamic model of this notion based on DNA elasticity, providing quantitative and parameter-free predictions of the relative activation of promoters containing a short versus long spacer when the DNA supercoiling level is varied. The model is tested through an analysis of *in vitro* and *in vivo* expression assays of mutant promoters with variable spacer lengths, confirming its accuracy for spacers ranging from 15 to 19 nucleotides, except those of 16 nucleotides where other regulatory mechanisms likely overcome the effect of this specific step. An analysis at the whole-genome scale in *Escherichia coli* then demonstrates a significant effect of the spacer length on the genomic expression after transient or inheritable superhelical variations, validating the model's predictions. Altogether, this study shows an example of mechanical constraints associated to promoter binding by RNA Polymerase underpinning a basal and global regulatory mechanism.

INTRODUCTION

DNA supercoiling (SC), the level of over- or underwinding of the double-helix, is a fundamental property of DNA. In bacterial cells, the chromosome is maintained at a negative SC level by a finely controlled balance between the relaxing activity of topoisomerase I (and IV), and the introduction of negative supercoils by the DNA gyrase (1,2). This level is affected by a variety of factors, including growth phase and environmental conditions (3). In addition to playing a key role in the physical organisation of the chromosome (4), SC

was soon discovered to affect the expression of many promoters (5,6). More recently, genome-wide expression profiling showed that SC acts as a global transcriptional regulator in many bacteria, based on studies employing gyrase inhibitors inducing a global DNA relaxation and, in turn, a global and complex response of genes (3). Rapid variations in SC levels may thus play an important and global role in the transcriptional response of bacteria to environmental changes (1,2).

At the mechanistic level, SC affects the transcription process at multiple steps, both indirectly through regulatory proteins, and directly by modulating the interaction of RNA Polymerase (RNAP) with DNA (3). A strong mechanism of the latter interaction is the facilitation of open-complex formation by negative SC during transcription initiation, due to the destabilisation of the double-helix (7). But other steps of transcription also contribute to this response, including closed-complex formation (8), promoter escape (9), transcription elongation and termination (10). Because of this complexity, regulatory models able to predict the response of a given promoter to SC variations are mostly lacking, at the quantitative and even the qualitative levels (3), in contrast to those involving regulatory proteins (11). Considering the widespread relevance of this mechanism in bacterial gene expression, obtaining such quantitative models is an important objective, both for a fundamental understanding of the process and for applications, e.g., in synthetic biology (12). The objective of this paper is to develop such a model, focusing on the specific step of closed-complex formation during transcription initiation.

The latter step involves the binding of RNAP to gene promoters, through specific recognition of the −35 and −10 elements by the 2.4 and 4.2 regions of the σ factor, respectively (13). The efficiency of RNAP binding on the promoter depends primarily on the proximity of the −35 and −10 hexamers to their respective consensus, but also on the spacer element between them. The latter exhibits little or no contact with the transcription machinery, except for promoters with an extended −10 element which interacts with the 3.0 region of σ factors (13). The spacer exhibits a variable length, ranging usually from 15 to 19 nucleotides (nt) for σ70-dependent promoters (13,14). For some pro-

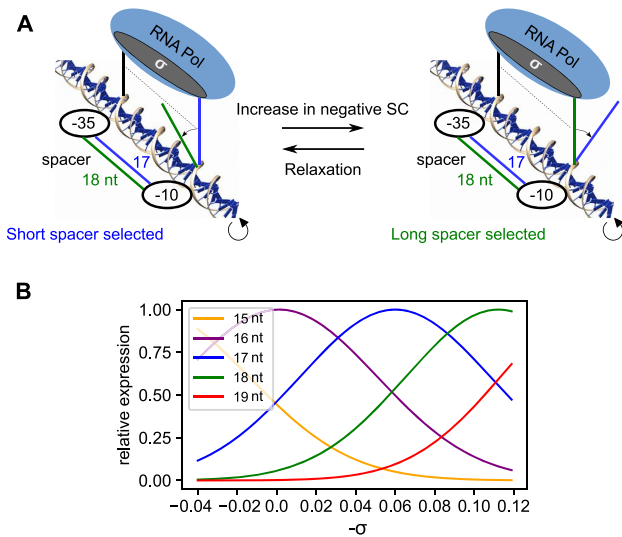*To whom correspondence should be addressed. Tel: +33 4 72 43 85 16; Email: sam.meyer@insa-lyon.fr

Figure 1. Quantitative modelling of the coupling between RNAP binding sites orientation, spacer length and DNA supercoiling. (**A**) Schematic depiction of −35/−10 alignment depending on spacer length and superhelical density. With long spacers (green), RNAP binding sites are out-of-phase at moderate SC levels, and become optimally aligned when DNA is highly negatively supercoiled, and conversely for shorter spacers (17-nt spacer in blue). (**B**) Quantitative relative expression levels predicted by the regulatory model for 15- to 19-nt spacers, depending on σ. Short spacers are more expressed at relaxed levels, while long spacers are favoured at highly negative levels. The only parameter θ$_P$, representing the optimal angle for RNAP binding, was chosen such that 17-nt spacers are maximally expressed at σ = −0.06 (with most predictions being independent of its value, see text).

moters, the spacer is ill-defined because of the weakness of their −35 sequence, which is then compensated by an extended −10 element (15) or by the assistance of regulatory proteins such as CRP (16). The maximal core promoter activity is reached with spacers of 17 nt for σ70-dependent promoters, and the addition or subtraction of nucleotides from this optimal length reduces their expression by several-fold (17,18). While the spacer sequence exhibits no specific requirement, point mutations (19,20) or modifications of its AT richness (21,22) also affect promoter expression (12,23), presumably by altering its 3D conformation (24,25).

In addition to altering the expression level of promoters, the spacer length has been shown to strongly modulate their SC response (17,18). A qualitative model of this regulation mode was proposed around 30 years ago, based on a simple geometric effect (8). Because of the helical nature of DNA, variable spacer lengths are associated to different relative orientations of the −35/−10 binding sites, which might modulate RNAP binding. In turn, the presence of torsional stress in the spacer might rotate the −35/−10 binding sites toward a (un)favourable orientation and thus regulate RNAP activity (Figure 1 A). This qualitative notion was supported by a review of observations from a collection of individual promoters (8,12,26), by the observation that the 18-nt spacer of the *lac* promoter is untwisted prior to open-complex formation (27), and it was also suggested as the mechanism of activation of several TFs (28). According to this mechanism, promoters containing sub-

optimal spacer lengths are not simply under-expressed, but might rather be selectively down- or up-regulated by the cell depending on the global SC level, according to their length. However, the analysis of this mechanism is difficult because the effect of SC on closed-complex formation also depends on other elements of the promoter, either downstream (29) or even at long upstream distance from the binding elements (30), complicating the integration of the accumulated data into a unifying model.

In this study, we wish to (i) translate the proposed qualitative notion into a quantitative regulatory model, (ii) develop a rigorous validation of its predictions, based on *in vitro* and *in vivo* analyses of mutant promoters containing variable spacer lengths and (iii) demonstrate the genome-scale relevance of this mechanism by a statistical analysis of high-throughput expression data, which were not available when the previous qualitative models were developed but are particularly suited to this global regulation mode. In spite of the complicating factors mentioned above, these results consistently show that the variability of spacer lengths underpins a global selectivity of promoter activity depending on the cellular SC level, varying either transiently due to topoisomerase inhibitors, or inheritably in the longest-running evolution experiment. Since this regulatory mechanism affects the basal interaction between promoter and RNAP, independently of promoter-specific TFs, it might play a widespread role in the prokaryotic kingdom.

MATERIALS AND METHODS

Thermodynamic model of supercoiling-dependent transcription

We assume that the formation of the closed-complex is limited by an intermediate state where the spacer DNA must be (un)twisted to a favourable relative orientation of the −10 and −35 elements allowing RNAP binding (Figure 1A). This deformation is treated in the elastic approximation, and the associated orientational free energy depends on the spacer length n and average superhelical level σ:

$$\Delta G_{or}(\sigma, n) = \frac{n}{2} k_\theta \left(\frac{\theta_P}{n} - \alpha_0 (1 + \sigma) \right)^2 \qquad (1)$$

where $k_\theta = 71.4 \, k_B T.rad^{-2}$ is the DNA sequence-averaged twist stiffness (31,32), $\alpha_0 = 34°$ is the average twist angle between adjacent nucleotides (33), $k_B T$ is the Boltzmann factor, and θ_P is a global parameter representing the optimal twist angle between −35 and −10 sites for RNAP binding. In this thermodynamic equilibrium, each of the n uniform basepair steps of the spacer is deformed from its initial angle modulated by the torque associated to the global SC level ($\alpha_0(1 + \sigma)$) to the angle imposed by RNAP (θ_P/n).

To illustrate the model predictions (Figure 1) and subsequent calculations, we assumed that 17-nt spacers achieve this optimal angle at the standard superhelical level σ = −0.06, i.e. $\theta_P = 17 \times \alpha_0 \times (1 − 0.06) = 543°$, but most key predictions are independent of its value (see upcoming paragraphs).

In this model, we assume that the total superhelicity σ can be utilised by RNAP for the twist deformation of spacer DNA (although most of it is stored as writhe at equi-

librium). This unidimensional approximation of DNA is frequently used in analyses of the regulatory effect of SC (8,34), and is relevant here in particular because the spacer DNA is too short to significantly writhe (before RNAP binding). Since RNAP is assumed to impose a fixed angle θ_P, its flexibility is neglected in the calculation.

The total free energy associated to the transcription process is assumed to contain three contributions related, respectively, to the effects of SC, of the spacer element, or of both simultaneously:

$$\Delta G(\sigma, n, s) = \Delta G_{sc}(\sigma, s) + \Delta G_{sp}(n, s) + \Delta G_{or}(\sigma, n) \quad (2)$$

where

- $\Delta G_{or}(\sigma, n)$ is the orientational deformation energy (Eq. 1), and introduces a *coupled* dependence on σ and n under the approximations given above.
- $\Delta G_{sc}(\sigma, s)$ represents all other mechanisms of regulation by SC (e.g. promoter opening, 3D deformations, structural transitions, etc.), which are assumed to be independent of spacer length n, and depend on the specific promoter sequence s (and thus even on elements distant from the spacer). This contribution explains the strong global promoter activation by negative SC in Figures 2 and 3 (even with short spacers), notably because of the facilitated opening of the sequence downstream of the −10 element during open-complex formation.
- $\Delta G_{sp}(n, s)$ represents all other mechanisms of modulation of transcriptional activity by the spacer DNA, which depend on its sequence as well as its length (e.g. 3D conformation, sequence-specific interactions with RNAP within the closed-complex, etc.), but are assumed to be independent of σ. In particular, we assume that the longitudinal distance between −35 and −10 elements may vary depending on the spacer length and sequence (possibly requiring the stretching of the spacer for RNAP binding), but with no SC-dependence, that is, neglecting the twist-stretch coupling of DNA.

Transcription rates are then computed using a thermodynamic framework (11):

$$k(\sigma, n, s) = k_0 \exp\left(-\frac{\Delta G(\sigma, n, s)}{k_B T}\right) \quad (3)$$

where k is the transcription rate, k_0 is the basal rate (which depends, e.g., on the −10/−35 sequence affinities for RNAP), and $k_B T$ is the Boltzmann factor. To simplify the notations, we use the latter as the energy unit in the following. Based on the equations above, the model does not predict the general SC-dependence of a promoter nor the effect of the spacer on the absolute expression level (since both depend on unpredictable terms in Eq. 2), but it does predict how the SC-sensitivity depends on spacer length by Eq. (1). Since most of the parameters involved in the latter have experimentally known values, it is possible to derive several quantitative and parameter-free predictions underpinning all analyses of the manuscript, as follows.

Prediction of relative *in vitro* expression levels depending on spacer length

From *in vitro* expression data of mutant promoters (Figure 2), we isolate the specific effect of the spacer length n by normalising each datapoint by the corresponding value obtained in the reference mutant promoter ($n_0 = 17$), thereby eliminating any other regulatory effect of SC:

$$\log\left(\frac{k(\sigma, n, s)}{k(\sigma, n_0, s)}\right) = (\Delta G_{or}(\sigma, n_0) - \Delta G_{or}(\sigma, n)) +$$
$$(\Delta G_{sp}(n_0, s) - \Delta G_{sp}(n, s)) \quad (4)$$

The second term is independent of σ, and is thus a constant for a promoter of given sequence and spacer size, hereafter quoted $Q_{sp}(n, s)$. Using a linear expansion in $\Delta n/n_0 = (n - n_0)/n_0$ in the first term, the relative expression level of each spacer length simplifies to a linear dependence in σ (Supplementary Eq. S1), as visible in Figure 2 B without approximation. Crucially, the *slope* of each line is defined without any adjustable parameter, and is proportional to the torsional stiffness of DNA and to $\Delta n = n - 17$. The intercept depends on the global parameter θ_P, and may also depend on the spacer sequence and length due to $Q_{sp}(n, s)$.

Prediction of *in vivo* expression fold-changes during superhelical variations

All analysed *in vivo* data (from mutant promoters or transcriptomics data) involve relative expression levels (fold-changes) induced by a global superhelical variation $\sigma_0 \rightarrow \sigma_0 + \Delta\sigma$ (induced by antibiotics or mutations). The predicted value is given in Supplementary Eq. (S2), where the spacer length-dependent term does not depend on any unknown parameter. After linear expansion in $\Delta\sigma$, the *relative* effect of the superhelical variation on spacer length mutants of the same promoter s is entirely predictable and independent of σ_0 (which is satisfactory since the absolute SC levels are not always known with precision in the analysed *in vivo* data):

$$\log FC(\Delta\sigma, n_0 + \Delta n, s) - \log FC(\Delta\sigma, n_0, s)$$
$$\simeq -k_\theta\, \alpha_0^2\, \Delta n\, \Delta\sigma \simeq -25\, \Delta n\, \Delta\sigma \quad (5)$$

This dependence is shown in Supplementary Figure S2 for a DNA relaxation $\Delta\sigma = 0.03$, and yields an expression ratio of around 2 between spacers differing by one nucleotide. In transcriptomic analyses, promoters are grouped by their spacer length n but differ by their overall sequence s, and the resulting term (R_{sc} in Supp. Eq. S2) remains in the form of strong statistical noise (Figures 4 and 5), imposing to work with the proportion of activated promoters rather than directly with the fold-change values.

Measurement of mutant promoters' responses to opposite supercoiling variations *in vivo*

230 bp sequences upstream of the *pheP* start codon were synthesised with mutations in the spacer (GeneCust), and individually cloned into pUCTer-luc plasmids upstream of a luciferase reporter gene (*luc*) (Supplementary Table S1). *E. coli* strain MG1655 cells were then transformed with these

plasmids using a standard electroporation procedure. The measurement protocol is described elsewhere (7). Briefly, *E. coli* cells carrying the plasmids with the different promoters were grown at 37°C in LB medium in a microplate reader (Tecan Spark). The OD_{600nm} and luminescence were measured every 5 min to follow bacterial growth and promoter expression, respectively. DNA relaxation was induced by injecting 5 μl of novobiocin (50, 100, 150 and 200 μg/ml final concentrations tested), whereas an increase in negative SC was induced by injecting 2 μl of seconeolitsine (25, 50, 75 and 100 μM final concentration tested) (35). The responses to these opposite DNA SC variations were then computed by comparing the luminescence values (in triplicates) of the novobiocin or seconeolitsine-shocked strain compared to the same strain injected with water (novobiocin solvent) or DMSO (seconeolitsine solvent), 60′ or 5′ after shock, respectively. The employed firefly luciferase has a lifetime of around 45′ in *E. coli* (36), and buffers the repressive effect of novobiocin. Confidence intervals and p-values were computed using Student's distributions. Raw datapoints are provided in Supplementary Figure S3. We previously showed (7) that the presence of the employed plasmids does not affect bacterial growth, and that results are not affected by plasmid copy-number variations (we compute relative expression levels between mutant promoters rather than absolute levels, and consistent observations were obtained in plasmid-borne or in chromosomal luciferase fusions), in agreement with other studies (37,38). The employed plasmids are well established as reflecting the average SC level of the chromosome (39), and also specifically in response to novobiocin-induced relaxation (40–42).

Parameter fitting in mutation data

In Figure 1B and 2B, for simplicity, the curves were drawn based on the orientational contribution only (Eq. 1), with the value $\theta_P = 543°$. In Figure 2B, the datapoints of Borowiec *et al*. (43) with a 18-nt spacer fall on the predicted line without adjustment, suggesting $Q_{sp} = 0$ for these promoters (see Eq. 5). The 19-nt spacer data of Aoyama *et al.* (44) were fitted with a value $Q_{sp}(n, s) = 1.1\,k_B T$ (Supplementary Eq. S1), corresponding to a factor 3 in expression. A putative interpretation of this additional cost is the elastic (longitudinal compression) energy required in the spacer to match the distance between RNAP binding sites (contained in ΔG_{sp} in Eq. 2).

Expression fold-changes measured in microplates with *pheP*-derived promoters were reproduced (Figure 3), starting from a level $\sigma = -0.06$, with an overtwisting magnitude $\Delta\sigma = -0.02$ (seconeolitsine), and a relaxation magnitude $\Delta\sigma = 0.005$ (novobiocin). This lower relaxation magnitude probably partly reflects a buffering effect of the reporter system and should thus be considered as an effective value used in the modelling, as further suggested by the lower repressive effect of novobiocin compared to batch cultures (45). The spacer-independent effect of SC, $R_s c$ (Eq. 6), was estimated from the repression (novobiocin) and activation (seconeolitsine) level observed with the 17-nt spacer, with values $R_{sc} = 0.4\,k_B T$ and $R_{sc} = -0.97\,k_B T$ corresponding to activation factors of 0.67 and 2.1, respectively (Figure 3, 17-nt spacers). These presumably reflect the modula-

tion of promoter opening energy by the two opposite superhelical variations (7).

Genome-wide analyses of spacer responses to supercoiling variations

Transcriptomic responses to DNA relaxation or inheritable SC variations were collected from the literature (Supplementary Table S2). The curated map of *E. coli* promoters with associated genes and spacer lengths was retrieved from Ecocyc (46). Only σ70-dependent promoters were retained and classified depending on their spacer length and response to the investigated condition, under standard statistical selection procedures (adjusted *P*-value < 0.05). For transcriptomes of *E. coli* evolved strains, a less stringent *P*-value threshold (0.1) was applied to have enough statistical power for the analysis (47). For other species, TSS maps were retrieved from the literature (48–51), and the positions of promoter elements were predicted using bTSSfinder (52); details are given in Supplementary Information. The relation between promoter activation and spacer length was quantified either by a Student's t-test between activated and repressed promoters (Figures 4A and 5B, C, Supplementary Figure S4) or by linear regression (Figure 4B). All computations were carried using a homemade Python package. All error bars shown are 95% confidence intervals.

RESULTS

Regulatory model of -35/-10 alignment during closed-complex formation

Following previous works (8,26), we hypothesised that, for a simultaneous binding of the −35 and −10 hexamers by the RNAP holoenzyme during closed-complex formation, an intermediate state must be achieved where the spacer DNA is (un)twisted to a favourable orientation (Figure 1A). We translated this notion into a quantitative regulatory model, based on a thermodynamic description of transcription (11) where this contribution can be computed from the torsional energy of spacer DNA, while RNAP is assumed to impose an optimal angle between the two sites. Treating DNA as a homogeneous polymer of known torsional stiffness, this energy depends on the spacer length and its superhelical state before RNAP binding, and the only adjustable parameter is the optimal angle of RNAP (Figure 1A, detailed hypotheses and equations are given in Materials and Methods and Supplementary Information).

Since SC affects transcriptional activity at many other steps of the process (promoter opening, escape, elongation, etc), we assume these other regulatory effects to be independent of the spacer length. Similarly, all other mechanisms by which the spacer length and sequence modulate transcriptional activity (e.g. variable longitudinal distance between −35 and −10 elements, 3D deformations, specific interactions with RNAP within the closed-complex, etc.) are assumed to be independent of the superhelical state of the promoter preceding RNAP binding. Based on these simplifying assumptions, the modulation of the torsional angle between the −35 and −10 hexamers is the only mechanism of *coupled* dependence between spacer length and SC, and
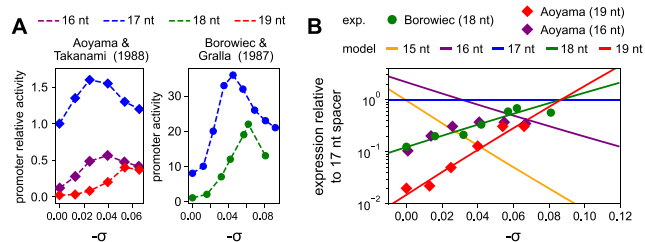
Figure 2. Comparison of *in vitro* transcription assays to model predictions. (**A**) Left panel: relative activities of promoters with 17- and 19-nt spacers (blue and red, respectively) depending on SC level σ, data from (44). Right panel: activity of promoters with 18-nt and 17-nt spacers (green and blue, respectively, data from (43)). (**B**) Transcription model predictions (solid lines) compared to the normalised *in vitro* promoter expression data from (44) (red diamonds) and (43) (green dots). Each datapoint was divided by the corresponding value obtained in the reference mutant promoter with a 17-nt spacer (see text). The slopes of the lines are parameter-free and proportional to $n - 17$ (where *n* is the spacer length), while the intercepts are promoter sequence-specific (see Materials and Methods).

the quantitative contribution of the spacer length to the SC-sensitivity of the promoter can be computed without any adjustable parameter, as developed in the following. The validity of these hypotheses should then be tested by comparing the predictions to available data, and may also depend on the specific promoter sequences.

The first prediction is the quantitative magnitude of this regulatory contribution. Geometrically, the large size of the spacer (around 17 nt) compared to that of many dimeric TFs (such as Fis or CRP, 8−10 nt) implies that physiologically relevant superhelical variations (of around ±0.06) are sufficient to induce a quantitative regulatory effect (same magnitude as an addition/removal of one nucleotide in the spacer, $17 \times 0.06 \simeq 1$). In turn, this regulatory effect is tuned by the mechanical cost of aligning an ill-oriented spacer (e.g. the 18-nt spacer in the left panel of Figure 1 A), which depends on the experimentally known value of the torsional stiffness of DNA (31). Figure 1B shows that, after computation, the addition/removal of one nucleotide in the spacer corresponds to a factor of around 2 in promoter expression (e.g. compare the 18 versus 17 nt spacers at σ = −0.06). While this value is milder than the 'on/off' regulation of many TFs affecting specific promoters, it is sufficient for a biologically relevant effect, especially if this factor affects RNAP activity in a global manner. This magnitude also approximately matches the observed effect of varying the spacer length (Supplementary Figure S1). But importantly, while the deformation energy has a symmetrical effect on the expression of spacers either too long or too short (same value of 16 and 18 nt spacers at σ = −0.06), the orientational effect of SC affects them in opposite directions, with short spacers being activated at relaxed levels (left) while long spacers are rather activated at highly negative levels (right).

In vitro **validation of model predictions on mutant promoters**

In order to rigorously test the model, we developed an analysis of previously obtained *in vitro* transcription data involving spacer length mutants of model promoters. In this experimental protocol, plasmid templates containing *lacP^s*-

derived promoters (43) and *PSC*-derived promoters (44) were prepared at well-defined superhelical densities and incubated with RNAP, minimising the effect of external regulatory factors. The raw datapoints (Figure 2A) exhibit a clear (spacer length-independent) tendency of activation by negative SC, which likely reflects the facilitated melting of the sequence downstream of the −10 element during open-complex formation (7) and possibly other regulatory contributions, here assumed to be independent of spacer length. To eliminate these other factors and allow a direct comparison with the model, we therefore normalised each datapoint by the one obtained at the same σ with the reference 17-nt spacer length (Figure 2B; by construction, the blue curve is flat, see rigorous theoretical analysis in Materials and Methods, Eq. 5). The prediction is a linear dependence of the datapoints (coloured lines), where the *slopes* are set by the DNA torsional stiffness value without any free parameter, and proportional to ($n − 17$), where *n* is the spacer length. In particular, the slope of the 19-nt spacer line is twice that of the 18-nt one (both activated by negative SC), and similarly for 15- and 16-nt spacers with negative values (repression by negative SC). Remarkably, the datapoints from two independent datasets obtained with different promoters (with 18- and 19-nt spacers) align along the predicted slopes without any adjustment. Since these slopes reflect the torsional elasticity of DNA, this observation gives a strong support to our hypothesis that this elasticity is indeed the key regulatory parameter in this process (whereas the intercepts of these lines reflect promoter-specific features, see Materials and Methods).

In the study by Aoyama *et al.* (44), the same data were also collected with 16-nt spacers, but, in direct contrast to our predictions, these are slightly more activated at strong SC levels (Figure 2, purple datapoints). Together with the previous ones, this observation suggests that our hypotheses may be valid for some spacer lengths (18 and 19 nt), but fail for 16-nt spacers. Possible reasons and biological implications are developed in the Discussion, but briefly, the latter behaviour can be expected if their differential SC-sensitivity is dominated not by the orientational deformation of DNA during closed-complex formation as we assume but, for example, by other mechanisms occurring during open-complex formation. The question then arises, if this deviation is a feature of short spacers in general, thus strongly reducing the usefulness of our model, or if it is an exception of 16-nt spacers or of this particular promoter. We therefore ran additional *in vivo* experiments on mutant promoters containing opposite spacer lengths of 15 and 19 nucleotides.

In vivo **validation with opposite superhelical variations**

We constructed spacer length mutants of the *pheP* promoter of *E. coli*, with a 17-nt (native), 15-nt (two deletions) or 19-nt spacer (two insertions, Supplementary Table S1). This promoter is SC-sensitive (45) and not regulated by any identified TF (46), and is thus a relevant candidate for this regulation mechanism based on the basal interaction with RNAP. Promoters were fused on plasmids in front of a luciferase reporter gene (Figure 3 A), and their expression was recorded in *E. coli* cells grown in LB rich medium in a mi-
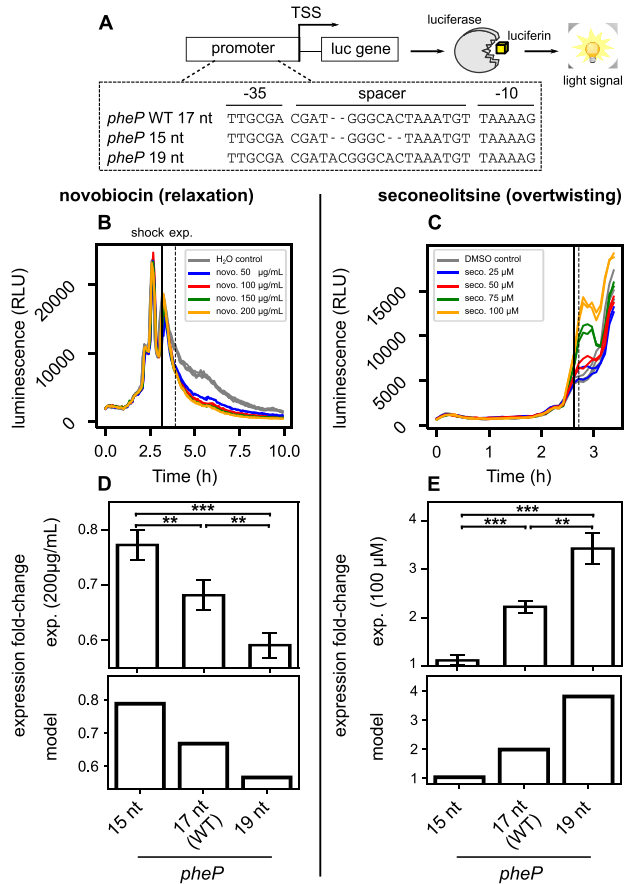
Figure 3. Responses of *pheP*-derived mutant promoters of variable spacer length to opposite SC variations. (**A**) Promoter sequences were synthesised from *pheP* (*E. coli*), with mutated spacers of different lengths. They control the expression of a *luc* gene encoding firefly luciferase, generating luminescence from luciferin substrate. (**B**) Promoter expression monitored in a microplate reader (bacteria carrying plasmids with *pheP* native promoter in LB medium), with a novobiocin shock applied in mid-exponential phase (time point quoted 'shock') at different sublethal concentrations. (**C**) same as (B), with a seconeolitsine shock performed at different sublethal concentrations, which induces a stronger and more transient expression variation. (**D, E**) Expression fold-changes computed 60′ (time point quoted 'exp.') after novobiocin shock (200 μg/mL), or 5′ after seconeolitsine shock (100 μM) from the experiments (upper panels), or predicted by the model (lower panels) assuming SC variations compatible with the observed expression variations levels (see Materials and Methods). All raw datapoints are in Supplementary Figure S3.

croplate reader. Expression levels were computed shortly after treatment by the gyrase inhibitor novobiocin (3) or the topoisomerase I inhibitor seconeolitsine (35,53) applied during exponential phase (Figure 3B, C). These drugs induce opposite SC variations in *E. coli*, DNA relaxation and overtwisting respectively (35), and thus provide complementary and independent tests of our predictions. Again, for a rigorous analysis of the data, we derived several parameter-free predictions (see dedicated paragraph in Materials and Methods, Eq. 7).

The native *pheP* promoter was repressed by novobiocin-induced DNA relaxation (Figure 3B), in agreement with previous data (45), and was activated by seconeolitsine (Figure 3C); both effects increased with the applied dosage. This

result might be explained by SC-assisted promoter opening (7), as in the *in vitro* data above. Note that the horizontal scale differs in the two plots because the shocks were applied at slightly different times, just after (resp. before) the maximal expression peak of *pheP*, in order to obtain reproducible (opposite) fold-change values with the two drugs. As a result, the growth stages of the cells might differ slightly between the two experiments, with only minor effect expected (optical densities of 0.55 versus 0.48 at shock time, see Supplementary Figure S3). The relatively modest (but highly significant) repression levels in response to novobiocin are due to a buffering effect of the reporter system (see Materials and Methods). Next, comparing the regulatory effect of the shocks on the mutants vs native promoter, we observed that novobiocin represses the 19-nt-spacer promoter with much stronger magnitude, whereas the 15-nt promoter was almost insensitive to DNA relaxation (Figure 3D), following model predictions. Conversely, under seconeolitsine-induced DNA overtwisting, the activation fold-change was strongest for the 19-nt promoter, and very low for the 15-nt promoter (Figure 3E). These four independent observations (differential effect of either superhelical variation on a shorter or longer spacer) were quantitatively reproduced by fitting the magnitude of the effective superhelical variations induced by the drugs (Figure 3D, E, one fitted parameter for each shock, see Materials and Methods). These results confirm, after the *in vitro* data above, the prediction that 19-nt spacers are favoured by strong negative SC levels, and 15-nt spacers are rather favoured by DNA relaxation. The opposite behaviour of 16-nt spacers noted above thus seems to be an abnormal case where our hypotheses break down, and these promoters will therefore be disregarded in the upcoming analyses (see Discussion). For all other spacer lengths, since the proposed mechanism relies on the basal interaction between RNAP and promoter elements, we then wished to enlarge the analysis to the entire genome of *E. coli*, in order to test its validity and relevance at the global scale.

Global effect of the spacer in promoter SC-sensitivity

We first looked at the variability of spacer lengths among σ70-dependent promoters of *E. coli*, based on an available curated promoter map (46). We focused on σ70 since it is predominant in the exponential phase where the analysed samples were collected. The most frequent spacer length is 17-nt (27% total promoters) and most others are distributed between 15- and 19-nt (Supplementary Figure S5), and we therefore focused on that range in all analyses. Based on the results above, we hypothesised that promoters with short spacers would be more activated by DNA relaxation than those harbouring long ones (Supplementary Figure S2). However, in contrast to the mutation studies above, genome-wide analyses involve the comparison of promoters differing by many additional factors beyond their spacer, including genomic context, surrounding sequences, bindings of transcriptional regulators, etc. We thus searched for a statistical relation between promoter selectivity during SC variations and spacer length, rather than a prediction valid for each analysed promoter (see theoretical analysis in Materials and Methods).
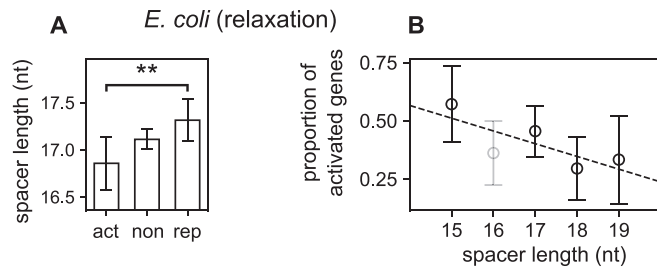
Figure 4. Genome-wide relation between spacer length and promoter selectivity during DNA relaxation in *E. coli*. (**A**) Comparison of the mean spacer lengths of σ70-dependent promoters activated (act), not significantly affected (non) or repressed (rep) by norfloxacin-induced DNA relaxation (LZ54 versus LZ41 strains (45)). As expected from our modelling for a DNA relaxation, activated promoters have significantly shorter spacers compared to repressed ones ($P = 0.007$). (**B**) Proportion of activated promoters among those responsive to DNA relaxation, depending on their spacer length (linear regression $P = 0.07$). As observed in *in vitro* data above (Figure 2), 16-nt spacers (in grey) do not follow the model and were excluded from statistical analyses (see Discussion).

In *E. coli*, the transcriptomic response to DNA relaxation was obtained with DNA microarrays (45), after a norfloxacin shock in two alternate topoisomerase mutant strains (54). The analysis shows that promoters activated by DNA relaxation indeed harbour significantly shorter spacers than repressed ones (Figure 4A, $P = 0.007$). Accordingly, classifying the promoters based on their spacer length (Figure 4B) exhibits a clear decreasing tendency (correlation $P = 0.07$). The relatively high level of noise is due to the heterogeneity of promoter sequences within each group, and also likely to a fraction of inaccurately annotated promoters, since a single-nucleotide resolution in the definition of −10 and −35 hexamers is required for an accurate analysis but not always achieved, especially since the −35 element is not always well-defined (15). As expected from the observations above, the 16-nt spacers again deviate from the model predictions (they were excluded from statistical analyses, see Discussion for a functional analysis of these promoters). For all others, these results suggest that, at the global scale, the variability of spacer length is used by bacterial cells for the selectivity of promoters in response to DNA relaxation.

Analysis of transcriptomic data in various species

Since the investigated mechanism relies on highly conserved molecular actors, RNAP and topoisomerases, it might affect a particularly broad range of bacterial species, and we therefore wished to extend this analysis to other organisms. But while the transcriptomic response to DNA relaxation induced by gyrase inhibitors has been recorded in several species (3), curated and accurate promoter maps are generally lacking. We therefore based our analysis on available maps of transcription start sites (TSS) obtained from specifically designed transcriptomic data (Supplementary Table S2), followed by a scan for promoter motifs (52). We thus obtained a list of putative promoters with associated σ factors and associated spacer lengths for two other enterobacteria, *Salmonella enterica* and the phytopathogen *Dickeya dadantii*, and at a drastically larger evolutionary distance, for the cyanobacterium *Synechococcus elongatus* and the small tenericute *Mycoplasma pneumoniae*. However, it must be noted that promoter prediction programs perform poorly in the detection of -35 elements: using the *E. coli* promoter map as a benchmark dataset, we found that the predicted −35 position deviated from the annotated one in around 50% promoters. In other species, this inaccuracy presumably resulted in a much higher level of statistical noise than in the annotated *E. coli* promoters above. In spite of this difficulty, a difference of spacer length in the same direction as in *E. coli* was observed in all investigated species (Supplementary Figure S4), albeit with weaker magnitudes and levels of statistical significance. Altogether, while improvements in promoter definition are clearly required for a solid conclusion, this systematic observation suggests that the variability of spacer length might indeed underpin a selective activation and repression of promoters by global SC variations throughout the prokaryotic kingdom.

Inheritable selection of promoters based on the spacer length

We finally investigated if the present mechanism could be involved not only in transient DNA relaxation responses induced by antibiotic shocks, but also in inheritable variations of global gene expression in the longest-running evolution experiment (55,56). Indeed, in this experiment involving the growth of *E. coli* cells in a daily refreshed minimal medium, point mutations affecting the SC level were quickly and naturally selected as they provided substantial fitness gains (55). A first mutation (in *topA*) was fixed before 2000 generations, and a second mutation (in *fis*) before 20 000 generations, both leading to an inheritable increase of negative SC (Figure 5A). Based on our modelling and the previous observations, we therefore expected promoters with a long spacer to experience enhanced expression in the evolved strains compared to the ancestor. Such a relation is indeed observed, both after 2000 generations (Figure 5B, $P = 0.04$) and 20 000 generations (Figure 5C, $P = 0.026$). The signal is significant but slightly weaker than that observed with antibiotics (Figure 4); this may be explained by the inheritable (rather than transient) nature of the SC variation, which induces an adaptive response of the cells via other regulatory pathways. Again, these results suggest that promoters of different spacer lengths respond differently to SC variations due to, for example, mutations in topoisomerase genes that are observed even between closely related species (57), according to predictable rules.

DISCUSSION

While the spacer length and sequence are known to modulate transcriptional activity, we wished to quantitatively model and test the long-proposed idea (8,43,44) that SC plays a specific role in this process through a simple orientational effect during closed-complex formation. This effect indeed emerged as a predictable quantitative signal, both in specifically designed mutant promoter assays, and as a statistical tendency in whole-genome data. The model and the latter results altogether suggest that this mechanism has a widespread relevance in bacterial transcription, although more detailed and comprehensive analyses will be
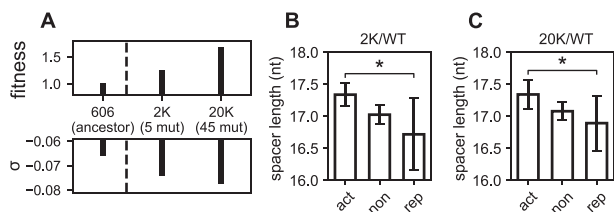
Figure 5. Contribution of the spacer to the global selectivity of promoters by inheritable increase of negative supercoiling in *E. coli*. (**A**) Reproduced from (47). Strains from the longest-running evolution experiment (56): 606 (ancestral genetic background), 2K (clone isolated from population at 2000 generations), 20K (at 20 000 generations). Evolved strains exhibit higher chromosomal SC density σ compared to the ancestor, due to the natural acquisition of two point mutations: one in *topA* before 2000 generations (among the five observed), and one in *fis* before 20,000 generations (among the 45 observed). Those mutations are associated to fitness gains through global expression changes (47). (**B**) Comparison of mean spacer lengths for promoters activated (act), not significantly affected (non) or repressed (rep) in the 2K evolved strain compared to the ancestor. As expected from our modelling for an increase of negative SC, activated promoters have longer spacers compared to repressed ones ($P = 0.04$). (**C**) Same for the 20K evolved strain compared to the ancestor, where the same difference is observed ($P = 0.026$). The wider confidence intervals for repressed promoters result from their lower number in the investigated conditions.

required to confirm it in various species. Interestingly, the underwinding of promoter DNA upstream of the melted region was observed also in the closed-complex at eukaryotic promoters, between +10 and +20 nt after the TATA box (58). Since eukaryotic RNA Polymerases present many structural similarities with bacterial RNAP, and since SC is also ubiquitous in eukaryotic chromosomes, comparable regulatory mechanisms involving the torsional register of DNA might thus also play a regulatory action in eukaryotic transcription.

Limitations of the regulatory model: the case of 16-nt spacers

The model was based on the hypothesis that the investigated mechanism (torsional orientation of −10 and −35 binding sites) could be decoupled from both (i) further regulatory effects of SC (assumed to be independent of the spacer length) and (ii) further modulating effects of the spacer length and sequence (assumed to be independent of SC). These hypotheses, already supported by a collection of qualitative observations on individual promoters (8), are validated here *a posteriori* by the quantitative agreement between model predictions and analysed data of different kinds. A notable exception, however, are promoters involving 16-nt long spacers, for which all observations converge to an opposite behaviour resembling that of long spacers (Figures 2 and 4B). It does not imply that the proposed mechanism does not occur for this family of promoters, but possibly that its effects are overcome by a stronger opposite regulatory effect of SC at a later stage of transcription, in particular during open-complex formation where different elements of the promoter make extensive and complex contacts with RNAP, and the destabilisation of the double helix by SC has a drastic influence on the expression level (7,13,59).

Interestingly, a well-studied class of promoters involving 16-nt spacers are those encoding stable RNAs in *E. coli*, subject to stringent control (both P1 and P2 promoters of all seven ribosomal operons and most promoters of tRNAs, based on the EcoCyc database). These experience a strong repression by ppGpp as well as DNA relaxation occurring in the cell upon transition to stationary phase (60–62). Both effects were attributed to the unusual kinetics of promoter opening and escape due to their G/C-rich 'discriminator' sequence inducing unusual interactions with RNAP in the open-complex (59,63,64). We note however that, based on the observations above, the repressive effect of DNA relaxation is also favoured by the 16-nt long spacer of these promoters (Figures 2 and 4), and both elements (discriminator and spacer) might thus contribute to their behaviour. To further test this hypothesis, we looked if a similar relation exists among promoters of protein-encoding genes. Indeed, we found that those containing 16-nt spacers exhibit significantly G/C-richer discriminators compared to all other promoters (Supplementary Figure S6, $P < 0.001$), suggesting a tight relation between both properties, and defining a subclass of promoters containing 16-nt long spacers and G/C-rich discriminators. It is conceivable that this subclass experiences a specific pathway in transcription initiation, responsible for its peculiar SC-sensitivity beyond the range of our model. Indeed, increasing the spacer from 16- to 17-nt enables the formation of an intermediate state with untwisted -10 element (but no opening of the promoter), which might then provide torsional energy facilitating the melting of the downstream region during open-complex formation (65). In the open-complex, the reactivity of spacer DNA to potassium permanganate or DMS was also shown to depend on its length or sequence (24,66). Altogether, besides its role in the torsional register of the −10 and −35 elements for closed-complex formation, the untwisting of the spacer may also affect subsequent steps of transcription initiation, particularly in the case of promoters containing a 16-nt spacer.

Additional factors influencing the relative orientation of −10/−35 elements

While we focused on the effect of SC on the direct interaction of RNAP with promoter DNA, the (un)twisting of the spacer has been proposed as the mechanistic basis for the regulatory action of several TFs, including MerR, which regulates the (*mer*) operon encoding components of the mercury (Hg) resistance system (28). The *mer* promoter includes a MerR binding site overlapping −35 and −10 elements, separated by a 19-nt spacer, and this unusual length was shown to be essential for normal activation by MerR (67). In the absence of Hg(II), MerR binds to the *mer* promoter in its repressor conformation. The presence of Hg(II) causes a conformational change in MerR and, in turn, the untwisting of DNA and reorientation of -35/-10 sites for effective open-complex formation (28). In agreement with this model, a global increase in negative SC facilitates MerR-mediated activation, impedes MerR-mediated repression, and conversely for DNA relaxation (68). The same mechanism presumably applies to other metal-dependent regulators (28,69), to the activation of the 19-nt spacer *mom* promoter by the C protein of bacteriophage Mu (70), and to the activation of the 19-nt spacer *soxS* promoter by the SoxR regulator of *E. coli* (71). The

general elastic model proposed here may thus be enriched to include these more specific actors.

Effect of spacer sequence on −35/−10 alignment by RNA Polymerase

Apart from the limitations mentioned above, we only considered how SC modulates the relative torsional orientation of −10 and −35 elements depending on the spacer length, and neglected any effect of its base sequence. However, the latter has been shown to affect the transcriptional activity of various promoters (19–23). The torsional orientation between the −35 and −10 hexamers, considered in this study, might contribute to this sequence-dependent regulation, especially since the crystallographic structures of DNA oligomers collected in the Nucleic Acids Database exhibit a strong variability of twist angles (by a factor of around 2) among dinucleotides of different sequences (31). We therefore extended our model to implement these sequence-dependent values (32), and estimate the resulting adjustment of the regulatory effect of SC, for all σ70-dependent promoters of *E. coli* (Supplementary Figure S7, with details of the computation in Supplementary Information). Overall, the maximal span of the sequence contribution remains weaker than the gain/loss of one nucleotide in the spacer (Supplementary Figure S7A), confirming our hypothesis that the effect of the spacer sequence is weaker than that of its length (for the considered orientational angle). More precisely, we asked if this modulation results from the sequence-induced heterogeneity of the average twist angle (i.e. the DNA structure, α_0 in Eq. 1) or stiffness (i.e. elasticity, k_θ), or both. Imposing a sequence-averaged value to either of these two parameters (Supplementary Figures S7B, C) showed that that the sequence-dependence of DNA stiffness alone has almost no regulatory effect (Supplementary Figure S7B). The sequence thus contributes mostly by modulating the total twist angle of the spacer, i.e., its average structure rather than its elasticity. As noted, this modulation yet remains modest compared to that induced by the variability in spacer length.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Travers,A. and Muskhelishvili,G. (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth?. *Nat. Rev. Microbiol.*, **3**, 157–169.
2. Dorman,C.J. and Dorman,M.J. (2016) DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys. Rev.*, **8**, 89–100.
3. Martis,B. S., Forquet,R., Reverchon,S., Nasser,W. and Meyer,S. (2019) DNA supercoiling: an ancestral regulator of gene expression in pathogenic bacteria?. *Comput. Struct. Biotechnol. J*, **17**, 1047–1055.
4. Verma,S.C., Qian,Z. and Adhya,S.L. (2019) Architecture of the *Escherichia coli* nucleoid. *PLoS Genet.*, **15**, e1008456.
5. Jovanovich,S.B. and Lebowitz,J. (1987) Estimation of the effect of coumermycin A1 on *Salmonella typhimurium* promoters by using random operon fusions. *J. Bacteriol.*, **169**, 4431–4435.
6. Pruss,G.J. and Drlica,K. (1989) DNA supercoiling and prokaryotic transcription. *Cell*, **56**, 521–523.
7. Forquet,R., Pineau,M., Nasser,W., Reverchon,S. and Meyer,S. (2021) Role of the discriminator sequence in the supercoiling sensitivity of bacterial promoters. *mSystems*, **6**, e0097821.
8. Wang,J.Y. and Syvanen,M. (1992) DNA twist as a transcriptional sensor for environmental changes. *Mol. Microbi.*, **6**, 1861–1866.
9. Wood,D.C. and Lebowitz,J. (1984) Effect of supercoiling on the abortive initiation kinetics of the RNA-I promoter of ColE1 plasmid DNA. *J. Biol. Chem.*, **259**, 11184–11187.
10. Dorman,C.J. (2019) DNA supercoiling and transcription in bacteria: a two-way street. *BMC Mol. Cell Biol.*, **20**, 26.
11. Bintu,L., Buchler,N.E., Garcia,H.G., Gerland,U., Hwa,T., Kondev,J., Kuhlman,T. and Phillips,R. (2005) Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.*, **15**, 125–135.
12. Klein,C.A., Teufel,M., Weile,C.J. and Sobetzko,P. (2021) The bacterial promoter spacer modulates promoter strength and timing by length, TG-motifs and DNA supercoiling sensitivity. *Sci. Repor.*, **11**, 24399.
13. Ruff,E.F., Record,M.T. and Artsimovitch,I. (2015) Initial events in bacterial transcription initiation. *Biomolecules*, **5**, 1035–1062.
14. Mitchell,J.E., Zheng,D., Busby,S.J.W. and Minchin,S.D. (2003) Identification and analysis of 'extended –10' promoters in *Escherichia coli*. *Nucleic Acids Res.*, **31**, 4689–4695.
15. Kumar,A., Malloch,R.A., Fujita,N., Smillie,D.A., Ishihama,A. and Hayward,R.S. (1993) The minus 35-recognition region of escherichia coli sigma 70 is inessential for initiation of transcription at an 'extended minus 10' promoter. *J. Mol. Biol.*, **232**, 406–418.
16. Soberón-Chávez,G., Alcaraz,L.D., Morales,E., Ponce-Soto,G.Y. and Servín-González,L. (2017) The transcriptional regulators of the CRP family regulate different essential bacterial functions and can be inherited vertically and horizontally. *Front. Microbiol.*, **8**, 959.
17. Aoyama,T., Takanami,M., Ohtsuka,E., Taniyama,Y., Marumoto,R., Sato,H. and Ikehara,M. (1983) Essential structure of E. coli promoter: effect of spacer length between the two consensus sequences on promoter function. *Nucleic Acids Res.*, **11**, 5855–5864.
18. Mulligan,M.E., Brosius,J. and McClure,W.R. (1985) Characterization in vitro of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter. *J. Biol. Chem.*, **260**, 3529–3538.
19. Chan,B. and Busby,S. (1989) Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA Polymerase. *Gene*, **84**, 227–236.
20. Thouvenot,B., Charpentier,B. and Branlant,C. (2004) The strong efficiency of the *Escherichia coli* gapA P1 promoter depends on a complex combination of functional determinants. *Biochem. J.*, **383**, 371–382.
21. Liu,M., Tolstorukov,M., Zhurkin,V., Garges,S. and Adhya,S. (2004) A mutant spacer sequence between -35 and -10 elements makes the Plac promoter hyperactive and cAMP receptor protein-independent. *Proc. Natl Acad. Sci. U.S.A.*, **101**, 6911–6916.
22. Hook-Barnard,I.G. and Hinton,D.M. (2009) The promoter spacer influences transcription initiation via sigma70 region 1.1 of Escherichia coli RNA Polymerase. *Proc. Natl Acad. Sci. U.S.A.*, **106**, 737–742.
23. Hook-Barnard,I.G. and Hinton,D.M. (2007) Transcription initiation by mix and match elements: flexibility for Polymerase binding to bacterial promoters. *Gene Reg. Syst. Biol.*, **1**, 275–293.

24. Auble,D.T. and DeHaseth,P.L. (1988) Promoter recognition by *Escherichia coli* RNA Polymerase: Influence of DNA structure in the spacer separating the 10 and 35 regions. *J. Mol. Biol.*, **202**, 471–482.

25. Warne,S.E. and deHaseth,P.L. (1993) Promoter recognition by *Escherichia coli* RNA Polymerase. Effects of single base pair deletions and insertions in the spacer DNA separating the -10 and -35 regions are dependent on spacer DNA sequence. *Biochemistry*, **32**, 6134–6140.

26. Unniraman,S. and Nagaraja,V. (2001) Axial distortion as a sensor of supercoil changes: a molecular model for the homeostatic regulation of DNA gyrase. *J. Genet.*, **80**, 119–124.

27. Amouyal,M. and Buc,H. (1987) Topological unwinding of strong and weak promoters by RNA polymerase: a comparison between the lac wild-type and the UV5 sites of Escherichia coli. *J. Mol. Biol.*, **195**, 795–808.

28. Brown,N.L., Stoyanov,J.V., Kidd,S.P. and Hobman,J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol. Rev.*, **27**, 145–163.

29. Pemberton,I.K., Muskhelishvili,G., Travers,A.A. and Buckle,M. (2000) The G+C-rich discriminator region of the tyrT promoter antagonises the formation of stable preinitiation complexes11Edited by M. Yaniv. *J. Mol. Biol.*, **299**, 859–864.

30. Sreenivasan,R., Shkel,I.A., Chhabra,M., Drennan,A., Heitkamp,S., Wang,H.-C., Sridevi,M.A., Plaskon,D., McNerney,C., Callies,K. *et al.* (2020) Fluorescence-detected conformational changes in duplex DNA in open complex formation by *Escherichia coli* RNA polymerase: upstream wrapping and downstream bending precede clamp opening and insertion of the downstream duplex. *Biochemistry*, **59**, 1565–1581.

31. Xu,F. and Olson,W.K. (2010) DNA architecture, deformability, and nucleosome positioning. *J. Biomol. Struct. Dyn.*, **27**, 725–739.

32. Cevost,J., Vaillant,C. and Meyer,S. (2018) ThreaDNA: predicting DNA mechanics' contribution to sequence selectivity of proteins along whole genomes. *Bioinformatics*, **34**, 609–616.

33. Harteis,S. and Schneider,S. (2014) Making the Bend: DNA tertiary structure and protein-DNA interactions. *Int. J. Mol. Sci.*, **15**, 12335–12363.

34. Zhabinskaya,D. and Benham,C.J. (2012) Theoretical analysis of competing conformational transitions in superhelical DNA. *PLoS Comput. Biol.*, **8**, e1002484.

35. Pineau,M., Shiny,M.B., Forquet,R., Baude,J., Villard,C.,Grand, L., Popowycz,F., Soulère,L., Hommais,F., Nasser,W., Reverchon,S. *et al.* (2021) What is a supercoiling-sensitive gene? Insights from topoisomerase I inhibition in the Gram-negative bacterium Dickeya dadantii. bioRxiv doi: https://doi.org/10.1101/2021.04.09.439150, 11 April 2021, preprint: not peer reviewed.

36. Thompson,J.F., Hayes,L.S. and Lloyd,D.B. (1991) Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene*, **103**, 171–177.

37. Dorman,C.J., Barr,G.C., Ni Bhriain,N. and Higgins,C.F. (1988) DNA supercoiling and the anaerobic and growth phase regulation of tonB gene expression. *J. Bacteriol.*, **170**, 2816–2826.

38. Schneider,R., Travers,A. and Muskhelishvili,G. (2000) The expression of the *Escherichia coli* fis gene is strongly dependent on the superhelical density of DNA. *Mol. Microbi.*, **38**, 167–175.

39. Hsieh,L.-S., Burger,R.M. and Drlica,K. (1991) Bacterial DNA supercoiling and [ATP]/[ADP]. Changes associated with a transition to anaerobic growth. *J. Mol. Biol.*, **219**, 443–450.

40. Cameron,A. D.S., Stoebel,D.M. and Dorman,C.J. (2011) DNA supercoiling is differentially regulated by environmental factors and FIS in *Escherichia coli* and *Salmonella enterica*. *Mol. Microbi.*, **80**, 85–101.

41. Ouafa,Z.-A., Reverchon,S., Lautier,T., Muskhelishvili,G. and Nasser,W. (2012) The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium *Dickeya dadantii*. *Nucleic Acids Res.*, **40**, 4306–4319.

42. Drlica,K. and Snyder,M. (1978) Superhelical *Escherichia coli* DNA: relaxation by coumermycin. *J. Mol. Biol.*, **120**, 145–154.

43. Borowiec,J.A. and Gralla,J.D. (1987) All three elements of the lac ps promoter mediate its transcriptional response to DNA supercoiling. *J. Mol. Biol.*, **195**, 89–97.

44. Aoyama,T. and Takanami,M. (1988) Supercoiling response of E. coli promoters with different spacer lengths. *Biochim. Biophys. Acta*, **949**, 311–317.

45. Blot,N., Mavathur,R., Geertz,M., Travers,A. and Muskhelishvili,G. (2006) Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep.*, **7**, 710–715.

46. Keseler,I.M., Mackie,A., Santos-Zavaleta,A., Billington,R., Bonavides-Martínez,C., Caspi,R., Fulcher,C., Gama-Castro,S., Kothari,A., Krummenacker,M. *et al.* (2017) The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res.*, **45**, D543–D550.

47. El Houdaigui,B., Forquet,R., Hindré,T., Schneider,D., Nasser,W., Reverchon,S. and Meyer,S. (2019) Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling. *Nucleic Acids Res.*, **47**, 5648–5657.

48. Webber,M.A., Ricci,V., Whitehead,R., Patel,M., Fookes,M., Ivens,A. and Piddock,L.J.V. (2013) Clinically relevant mutant DNA gyrase alters supercoiling, changes the transcriptome, and confers multidrug resistance. *mBio*, **4**, e00273-13.

49. Forquet,R., Jiang,X., Nasser,W., Hommais,F., Reverchon,S. and Meyer,S. (2022) Mapping the complex transcriptional landscape of the phytopathogenic bacterium Dickeya dadantii. *mBio*, e0052422.

50. Junier,I., Unal,E.B., Yus,E., Lloréns-Rico,V. and Serrano,L. (2016) Insights into the mechanisms of basal coordination of transcription using a genome-reduced bacterium. *Cell Syst.*, **2**, 391–401.

51. Vijayan,V., Jain,I.H. and O'Shea,E.K. (2011) A high resolution map of a cyanobacterial transcriptome. *Genome Biol.*, **12**, R47.

52. Shahmuradov,I.A., Mohamad Razali,R., Bougouffa,S., Radovanovic,A. and Bajic,V.B. (2017) bTSSfinder: a novel tool for the prediction of promoters in cyanobacteria and Escherichia coli. *Bioinformatics*, **33**, 334–340.

53. Ferrándiz,M.-J., Martín-Galiano,A.J., Arnanz,C., Camacho-Soguero,I., Tirado-Vélez,J.-M. and de la Campa,A.G. (2016) An increase in negative supercoiling in bacteria reveals topology-reacting gene clusters and a homeostatic response mediated by the DNA topoisomerase I gene. *Nucleic Acids Res.*, **44**, 7292–7303.

54. Zechiedrich,E.L., Khodursky,A.B., Bachellier,S., Schneider,R., Chen,D., Lilley,D.M. and Cozzarelli,N.R. (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli*. *J. Biol. Chem.*, **275**, 8103–8113.

55. Crozat,E., Philippe,N., Lenski,R.E., Geiselmann,J. and Schneider,D. (2005) Long-term experimental evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics*, **169**, 523–532.

56. Tenaillon,O., Barrick,J.E., Ribeck,N., Deatherage,D.E., Blanchard,J.L., Dasgupta,A., Wu,G.C., Wielgoss,S., Cruveiller,S., Médigue,C. *et al.* (2016) Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature*, **536**, 165–170.

57. Rovinskiy,N.S., Agbleke,A.A., Chesnokova,O.N. and Higgins,N.P. (2019) Supercoil levels in E. coli and Salmonella chromosomes are regulated by the C-terminal 35–38 amino acids of GyrA. *Microorganisms*, **7**, 81.

58. Dienemann,C., Schwalb,B., Schilbach,S. and Cramer,P. (2019) Promoter distortion and opening in the RNA Polymerase II Cleft. *Mol. Cell*, **73**, 97–106.

59. Ohlsen,K.L. and Gralla,J.D. (1992) DNA melting within stable closed complexes at the *Escherichia coli* rrnB P1 promoter. *J. Biol. Chem.*, **267**, 19813–19818.

60. Schneider,D.A., Ross,W. and Gourse,R.L. (2003) Control of rRNA expression in *Escherichia coli*. *Curr. Opin. Microbiol.*, **6**, 151–156.

61. Dennis,P.P., Ehrenberg,M. and Bremer,H. (2004) Control of rRNA synthesis in *Escherichia coli*: a systems biology approach. *Microbiol. Mol. Biol. Rev.*, **68**, 639–668.

62. Hori,H., Tomikawa,C., Hirata,A., Toh,Y., Tomita,K., Ueda,T. and Watanabe,K. (2014) Transfer RNA synthesis and regulation. In: *eLS*. American Cancer Society.

63. Lamond,A.I. (1985) Supercoiling response of a bacterial tRNA gene. *EMBO J.*, **4**, 501–507.

64. Winkelman,J.T., Chandrangsu,P., Ross,W. and Gourse,R.L. (2016) Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of E. coli ribosomal RNA promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E1787–E1795.

65. Auner,H., Buckle,M., Deufel,A., Kutateladze,T., Lazarus,L., Mavathur,R., Muskhelishvili,G., Pemberton,I., Schneider,R. and Travers,A. (2003) Mechanism of transcriptional activation by FIS:

role of core promoter structure and DNA topology. *J. Mol. Biol.*, **331**, 331–344.

66. Newlands,J.T., Ross,W., Gosink,K.K. and Gourse,R.L. (1991) Factor-independent activation of Escherichia coli rRNA transcription: II. Characterization of complexes of rrnB P1 promoters containing or lacking the upstream activator region with Escherichia coli RNA polymerase. *J. Mol. Biol.*, **220**, 569–583.

67. Parkhill,J. and Brown,N.L. (1990) Site-specific insertion and deletion mutants in the mer promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. *Nucleic Acids Res.*, **18**, 5157–5162.

68. Condee,C.W. and Summers,A.O. (1992) A mer-lux transcriptional fusion for real-time examination of in vivo gene expression kinetics and promoter response to altered superhelicity. *J. Bacteriol.*, **174**, 8094–8101.

69. Heldwein,E.E. and Brennan,R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature*, **409**, 378–382.

70. Basak,S. and Nagaraja,V. (2001) DNA unwinding mechanism for the transcriptional activation of momP1 promoter by the transactivator protein C of bacteriophage Mu. *J. Biol. Chem.*, **276**, 46941–46945.

71. Hidalgo,E. and Demple,B. (1997) Spacing of promoter elements regulates the basal expression of the soxS gene and converts SoxR from a transcriptional activator into a repressor. *EMBO J.*, **16**, 1056–1065.