

Characterization of Gene Repression by Designed Transcription Activator-like Effector Dimer Proteins

Nicole A. Becker,¹ Justin P. Peters,² Tanya L. Schwab,¹ William J. Phillips,¹ Jordan P. Wallace,¹ Karl J. Clark,¹ and L. James Maher III^{1,*}

¹Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine and Science, Rochester, Minnesota and ²Department of Chemistry and Biochemistry, University of Northern Iowa, Cedar Falls, Iowa

ABSTRACT Gene regulation by control of transcription initiation is a fundamental property of living cells. Much of our understanding of gene repression originated from studies of the *Escherichia coli lac* operon switch, in which DNA looping plays an essential role. To validate and generalize principles from *lac* for practical applications, we previously described artificial DNA looping driven by designed transcription activator-like effector dimer (TALED) proteins. Because TALE monomers bind the idealized symmetrical *lac* operator sequence in two orientations, our prior studies detected repression due to multiple DNA loops. We now quantitatively characterize gene repression in living *E. coli* by a collection of individual TALED loops with systematic loop length variation. Fitting of a thermodynamic model allows unequivocal demonstration of looping and comparison of the engineered TALED repression system with the natural *lac* repressor system.

INTRODUCTION

Gene regulation by control of transcription initiation is fundamental to living cells. Interactions between proteins and DNA and between proteins can drive RNA polymerase recruitment to or exclusion from promoter sequences in DNA. Control is typically through accessory and regulatory proteins, often tuned by post-translational modifications (1–3). Differences exist in the modes of regulation of eukaryotic, archaeal, and prokaryotic transcription initiation (4). However, both activation and repression of transcription initiation are observed in all three kingdoms of life.

Key insights into the control of prokaryotic transcription initiation were originally gained from classic investigations of the *Escherichia coli lac* operon (5) and the left and right promoters of coliphage λ (6). *lac* control illustrates both repression and activation functions of accessory proteins influencing RNA polymerase binding to the *lac* promoter (3,5,7). The λ system similarly illustrates negative and positive control but was also the first system to demonstrate cooperative binding by clusters of λ repressor proteins locally and through DNA looping (6,8,9).

We have been studying repression of transcription initiation in the LacI repressor system (10–16) because of our interest in understanding how the bending and twisting

rigidities of the DNA double helix are managed in living cells (17). The classic studies of Müller-Hill (5,18–21) and Record (22–24) were the first to demonstrate that the *lac* switch features auxiliary (distal) operators in addition to the proximal operator that overlaps the promoter. It was shown that repression of transcription initiation is controlled by the effective concentration of *lac* repressor at the proximal operator. Effective repressor concentration at this operator is increased by simultaneous binding of bidentate *lac* repressor tetramer at distal operators. This repression enhancement occurs by cooperativity at a distance via DNA looping (25–29). We have previously studied gene control by assembling elements of the *lac* control switch that allow us to deduce probabilities of DNA looping as a function of DNA length using expression of the *lacZ* gene as the readout (10,13). This approach allows sensitive measurement of biophysical details of DNA looping energetics in vivo at basepair (bp) resolution using ensemble experiments.

Our past studies have illuminated fundamental aspects of the LacI repressor DNA looping mechanism, including the interplay of intrinsic operator affinity (controlled by both DNA sequence and the binding of inducer), operator position, DNA bending and twisting flexibilities, and architectural DNA-binding proteins that modify the physical properties of DNA (10–15). With this background, we have recently sought to exploit fundamental principles of

Submitted July 14, 2020, and accepted for publication October 13, 2020.

*Correspondence: maher@mayo.edu

Editor: Wilma Olson.

<https://doi.org/10.1016/j.bpj.2020.10.007>

© 2020 Biophysical Society.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



the *lac* operon in designing an artificial DNA looping system for application in controlling transcription initiation at any promoter in *E. coli* or other organisms. Our premise is that adequate understanding of the natural *lac* system should enable construction of an artificial system mimicking some of its features but targeted to regulate arbitrary promoters.

Transcription activator-like effector (TALE) proteins originating in bacterial pathogens of plants (30,31) employ independent base-specific DNA recognition modules to bind the DNA major groove, allowing engineered targeting. We recently described elements of an artificial control system based on fusions between designed TALE proteins and dimerization domains controllable by small molecules (32). This platform allows us to fuse DNA-binding domains with dimerization domains to create artificial DNA looping proteins. Depending on the choice of dimerization domains, dimerization can be constitutive or made to be dependent on either the presence or absence of small molecules. Our initial study introduced the design of these sequence-specific DNA-binding proteins, confirmed their ability to act as repressors by targeting *lac* operator sequences, and presented preliminary evidence of repression by DNA looping. This evidence came from the observation that repression was enhanced by a distal (upstream) operator, and enhancement depended on the spacing between proximal and distal operators (a signature of DNA looping). Further, repression depended on TALE protein dimerization to form transcription activator-like effector dimers (TALEDs).

In the initial study, the sequence symmetry of the idealized *lac* operator meant that TALE monomer binding was able to occur in either of two orientations, such that each operator spacing could support two or four competing DNA loops that formed simultaneously, depending on the specific operators and TALEDs being studied. Although producing clear evidence for looping enhancement of repression, data quantitation and interpretation were complicated by the potential for multiple TALED-mediated DNA loop geometries.

We now extend our previous results to characterize in detail promoter repression by a designed TALED. We measure how gene repression depends on the relative orientation of operators, and document DNA looping by measuring its length dependence at bp resolution for cases in which only a single loop conformation is possible at each operator spacing. Importantly, these new, to our knowledge, coherent data for single loops allow meaningful quantitative thermodynamic modeling of engineered TALED-based gene control elements. In turn, this new analysis permits a first systematic comparison with corresponding biophysical parameters obtained from our prior studies of the natural *lac* system. This analysis sets the stage for implementation of TALED-directed DNA looping for gene control in other prokaryotic and eukaryotic systems and comparison with other designed approaches (33–35).

MATERIALS AND METHODS

DNA looping reporter constructs

Episomal and plasmid DNA looping constructs (Figs. S1 and S2; Table S1) were based on plasmid pJ2280 (32). Episomal spacing constructs were created by modifications of pFW11-null as described in the Supporting Materials and Methods (36,37). Plasmid constructs contain the complete *lacZ* coding sequence downstream of the promoter and operator(s).

TALE-FKBP protein expression

Cloning of genes encoding designed TALEs involved described methods (38). TALE-FKBP protein expression plasmid was created using a modified version of plasmid pJ1035 (promoter of moderate strength) (37). Plasmid pJ1035 contains the bacterial UV5 promoter with complete -10 and -35 box sequences. See Supporting Materials and Methods and Fig. S3 for full details.

E. coli β -galactosidase reporter assay

LacZ expression (E) was measured using a liquid β -galactosidase colorimetric enzyme assay (39) adapted as previously described (32). Assays were performed with a minimum of three colonies repeated on each of 2 days for at least six data points. Normalized reporter expression (E'), with or without TALE-FKBP, allows for comparisons among experiments:

$$E' = \frac{E_{two\ operators}}{E_{single\ operator}}$$

Repression was quantitated in terms of repression ratio (RR):

$$RR = \frac{E_{-TALE-FKBP}}{E_{+TALE-FKBP}}$$

with the contribution to the repression ratio due to free repressor binding at the proximal operator defined as RR_F ,

$$RR_F = \left[\frac{E_{-TALE-FKBP}}{E_{+TALE-FKBP}} \right]_{single\ operator};$$

the overall contributions to the repression ratio due to free repressor and DNA looping defined as RR_T ,

$$RR_T = \left[\frac{E_{-TALE-FKBP}}{E_{+TALE-FKBP}} \right]_{two\ operators};$$

and the contributions to the repression ratio due to DNA looping defined as RR' ,

$$RR' = \frac{RR_T}{RR_F}$$

Data fitting to thermodynamic model

The thermodynamic model of promoter repression used to fit data relating gene expression to the presence and spacing of operator sequences has been previously described (10,13,17). The adaptation of this model to the current analysis is explained in Supporting Materials and Methods. Briefly, the fraction of proximal operator bound by TALED protein as a function of DNA operator-operator length is modeled with five adjustable parameters

evaluating the distribution of possible states of the proximal operator through a partition function for the system. Fit parameters give insight into the physical properties of the nucleoprotein loop.

RESULTS AND DISCUSSION

TALED design

A designed TALED (Fig. 1 A; Fig. S3) was created using a semiautomated method to recognize a 15-bp subsequence within the asymmetric *lac* O₂ operator (38). TALE fusion with a C-terminal FKBP(F36M) mutant domain (Fig. 1 A, “DD”) in place of the natural TALE transcription activation domain allows constitutive homodimerization, with affinity reported to be 30 μM in vitro (40).

lac looping model systems

As in our prior studies, we demonstrated and characterized engineered DNA looping in vivo using promoter-reporter constructs of the form shown in Fig. 1 B. The *lac* UV5 promoter driving *lacZ* is flanked by identical O₂ operators (Fig. S1) derived from the *lac* operon. We intentionally chose to target *lac* operators to allow use of promoter-reporter constructs previously created for analysis of DNA looping by LacI repressor and to facilitate direct comparison of results. The center-to-center operator spacing is systematically varied to monitor the relationship between the energetically unfavorable DNA bending and twisting required for TALED-driven looping and the transcriptional readout. TALED binding to the isolated proximal operator inhibits promoter function (Fig. 1 C), and this repression is enhanced by increasing local TALED concentration and promoter distortion by looping (Fig. 1 D).

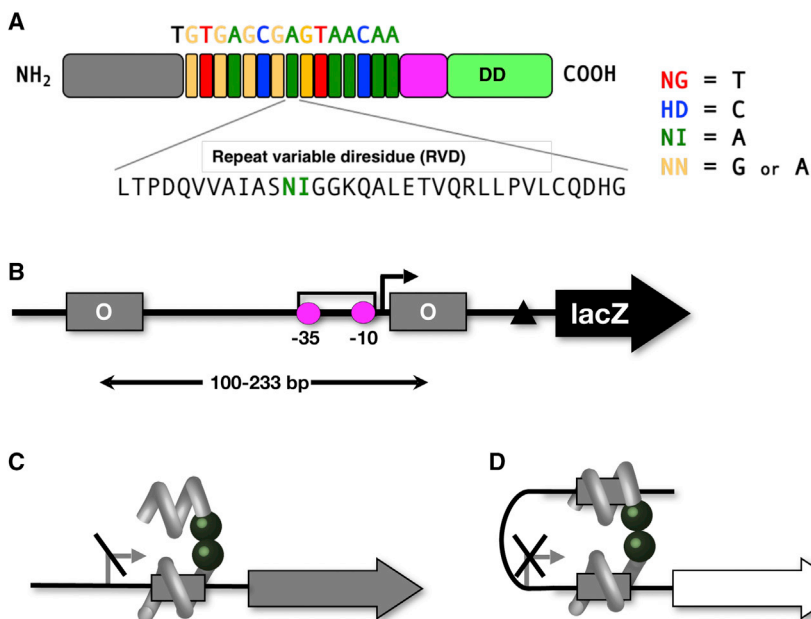


FIGURE 1 Experimental design. (A) TALE protein design and fusion to C-terminal FKBP(F36M) mutant dimerization domain (“DD”) is shown. Tandem 34-amino acid repeats (single letter amino acid codes) with programmed base-specific repeat variable diresidue domains are indicated in colors. The 15-bp DNA sequence recognized by this TALE protein is indicated above. (B) Example of promoter construct design for DNA looping studies is given. *lac* operators flank a *lac* UV5 promoter (broken arrow shows transcription start site with -10 and -35 sequences indicated) such that the proximal operator is just downstream of the promoter and the distal operator is at various distances (measured operator center-to-center) upstream. The *lacZ* gene acts as reporter, and the Shine-Dalgarno sequence is indicated (triangle). (C) A schematic of weak repression by TALED binding only the proximal operator in an unlooped configuration is given. (D) Strong repression by TALED-dependent DNA looping for an example operator configuration is shown. To see this figure in color, go online.

Effect of operator orientation on TALED-dependent gene repression in vivo

Our prior study (32) involved TALED recognition of a symmetrical O_{sym} *lac* operator, complicating interpretation of results because the directional TALE protein can bind the symmetrical operator in either of two orientations. For combinations involving such operators, two or four competing DNA loop configurations are possible. To unequivocally confirm DNA looping and measure single coherent TALED-driven DNA loops, we designed a TALE that recognizes the asymmetrical *lac* O₂ operator, supporting a single defined binding geometry (Fig. 2 A) that can be controlled depending on the orientation of the O₂ operator. With respect to the operator orientation shown in Fig. 2 A, the inverted O₂ orientation is termed invO₂ (Fig. 2 B).

We sought to determine if the stability of a DNA loop driven by the TALED homodimer depends on the relative binding orientations of the two anchoring TALEs (Fig. 2 C, left two columns). We therefore collected plasmid-based reporter expression data for three different operator spacings as operator orientations were altered (Table S2). From these results we highlight the RR, which compares reporter expression with and without TALED (Fig. 2 C). Relative to constructs with a single proximal operator in either orientation, constructs with two operators all showed increased repression (Fig. 2 C), consistent with DNA looping (32). Interestingly, loop stability (indicated by extent of reporter repression) varied somewhat as a function of operator orientation, but effects on RR were generally less than twofold, and there was not a consistent trend that convergent, divergent, or parallel operator orientations were favored. This result suggests that, for the DNA loop sizes studied here,

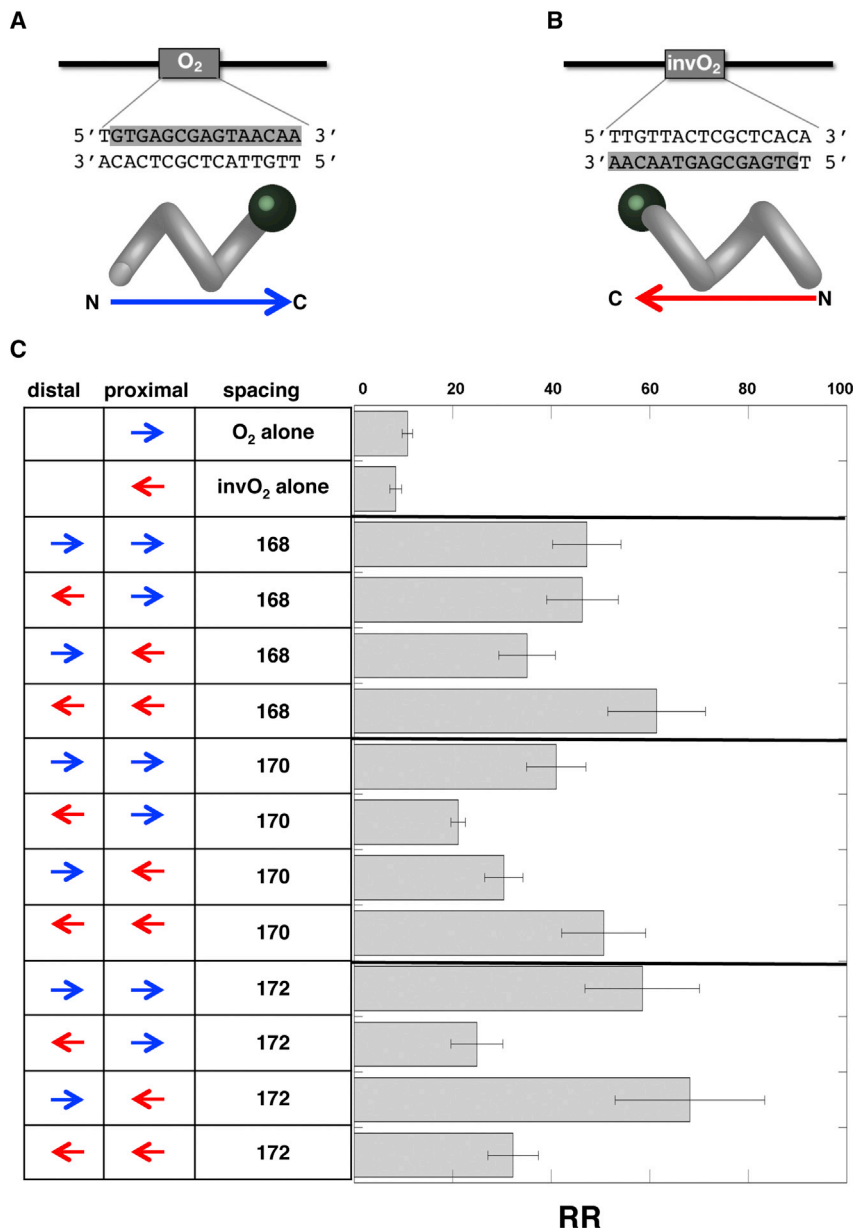


FIGURE 2 Effect of TALE-operator orientation on repression looping. (A) TALE targeting of the purine-rich strand of O₂ when this asymmetric operator is oriented in the forward direction yields the indicated protein binding polarity. (B) The O₂ operator in a flipped orientation (invO₂) recognized by the same TALE yields the opposite protein polarity. (C) Data comparing repression ratios (as defined in Materials and Methods) for the indicated operator configurations and center-to-center spacings are given. To see this figure in color, go online.

TALED protein flexibility appears to accommodate different loop geometries. In subsequent studies, we explicitly distinguish each family of loops (i.e., O₂-O₂, invO₂-O₂, etc., where the first listed operator is promoter distal and the second listed operator is promoter proximal).

DNA looping by TALE homodimers: effect of operator spacing and context for single loops

Given the apparent similarity of loop stabilities for different operator orientations, we chose to collect reporter expression data for one representative series of constructs with O₂-O₂ and invO₂-O₂ operator configurations at different spacings. In the process, we also studied how results were

affected by placement of the promoter-reporter construct on the single-copy F' episome (our conventional choice to mimic the chromosome) versus a low-copy-number plasmid. Our systematic TALED studies occur in a LacI impaired background (LacI Y282D) by placement of the promoter-reporter construct on plasmids versus homologous recombination into the F' episome. It was therefore important to determine whether this more convenient plasmid context gives results comparable with those obtained in the F' episome. Plasmid-based and episome-based data (Tables S3 and S4) measuring repression as a function of operator center-to-center spacing are shown in Fig. 3.

Fig. 3 A shows raw reporter activity from the indicated bacterial strains and operator orientations under conditions

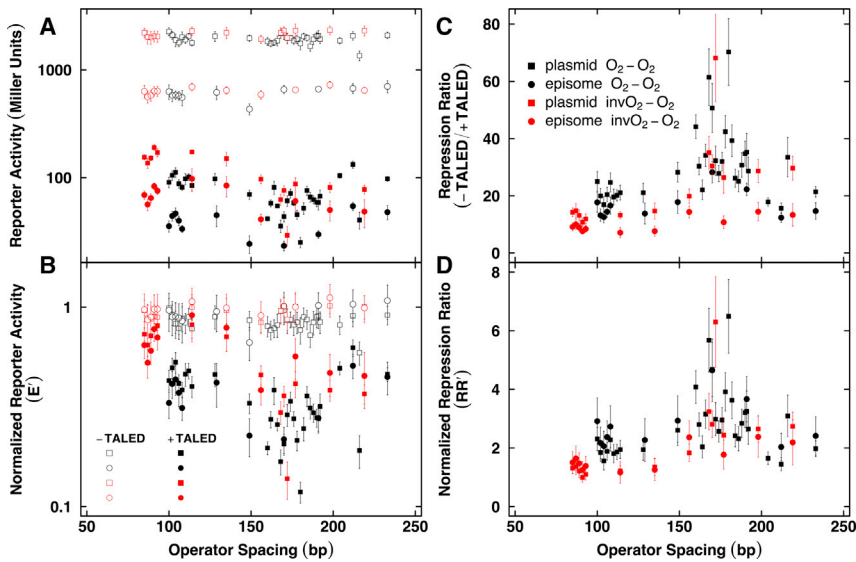


FIGURE 3 Transcriptional activity and repression by single TALE D loop configurations as a function of operator spacing. (A) Absolute reporter activity for constructs of the indicated operator pairs at the indicated center-to-center spacings in the indicated context (low-copy plasmid or single-copy episome) is shown. (B) The data from (A) normalized to control data from corresponding constructs containing only a proximal operator are given. (C) Repression ratio (as defined in [Materials and Methods](#)) for the data in (A) is shown. (D) Normalized repression ratio (as defined in [Materials and Methods](#)) summarizing the improvement in repression specifically due to looping is shown. To see this figure in color, go online.

with or without expression of the homodimer TALE D. It is immediately evident that operator pairs lead to length-dependent promoter repression in the presence of TALE D protein relative to strains lacking TALE Ds (*solid* versus *open symbols* in [Fig. 3 A](#)). Also evident is the higher reporter activity from cells with plasmid-borne reporters versus reporters on the single-copy F' episome (*squares* versus *circles* in [Fig. 3 A](#)). Both results are consistent with expectations. Interestingly, when reporter activity is normalized to the activity of reference constructs carrying only a single proximal operator (E'), the data coalesce into the coherent pattern seen in [Fig. 3 B](#). This result indicates that DNA looping behavior is comparable for episomal and plasmid constructs, suggesting that DNA packaging is similar in both contexts and the titration effect of operator copy number does not substantially influence repression for the intracellular TALE D concentration studied here.

TALE D-dependent effects are best seen by expressing reporter data as the RR ([Fig. 3 C](#)) and specific loop-dependent contributions by expressing reporter data as the normalized RR' ([Fig. 3 D](#)), as described in [Materials and Methods](#). Striking in both [Fig. 3, C and D](#) is the evidence for a local maximum in repression in the vicinity of 175-bp operator separation.

Several different theoretical models have been proposed for interpretation of experimental protein-mediated looping data (41–43). A key challenge is inclusion of sufficient structural detail to accurately capture the complicated dependence of loop stability on loop length when multiple loop conformations may coexist, with multiple topoisomers conceivable for each such loop. Despite these complications, certain features typically exist in all DNA looping models: 1) the looping is oscillatory with the period of oscillations closely matching the helical repeat of DNA, 2) the amplitude of these oscillations decreases with increasing

length of the loop, 3) the envelope of oscillations shows a peak in the range between 50 and 200 bp, and 4) the amplitude of oscillations is not correlated with the probability of loop formation.

To analyze and interpret these data more completely, we performed fitting with an established thermodynamic model of promoter repression by DNA looping (see [Materials and Methods](#) and [Supporting Materials and Methods](#) (10,13,37)). We confined our fitting to spacings ranging from 160 to 192 bp for which the experimental data are most complete. [Fig. 4](#) shows both results for the fit region (*solid lines*) as well as model predictions extrapolated to the other data (*dotted lines*). This approach emphasizes TALE D looping regimes that are and are not well treated by the simple thermodynamic model. The results for the O_2 - O_2 loop series are shown as normalized reporter activity (E') in [Fig. 4 A](#) and as normalized repression ratio (RR') in [Fig. 4 B](#). Results from episome and plasmid contexts are similar enough to be well characterized by a single set of model parameters (*black lines* in [Fig. 4](#)), clearly demonstrating the characteristic oscillation of repression as a function of operator separation, interpreted as the result of the face-of-the-helix dependence of looping energy favoring repression by untwisted loops. This result firmly establishes DNA looping driven by TALE Ds. The comparable, but smaller, $invO_2$ - O_2 data set is similarly analyzed in [Fig. 4, C and D](#), with the O_2 - O_2 model in black for comparison.

Values of thermodynamic fitting parameters for these data are shown in [Table 1](#). Parameter estimates are interpreted tentatively to generate a useful visual guide but must be considered with caution because of the low sampling of length data in some regions, relative simplicity of the theoretical model, and difficulty in fitting the highly oscillatory function. The fit values for the apparent helical repeat (10.42 bp/turn for the O_2 - O_2 loop series and 10.75 bp/turn for the $invO_2$ - O_2 loop series) are comparable with that

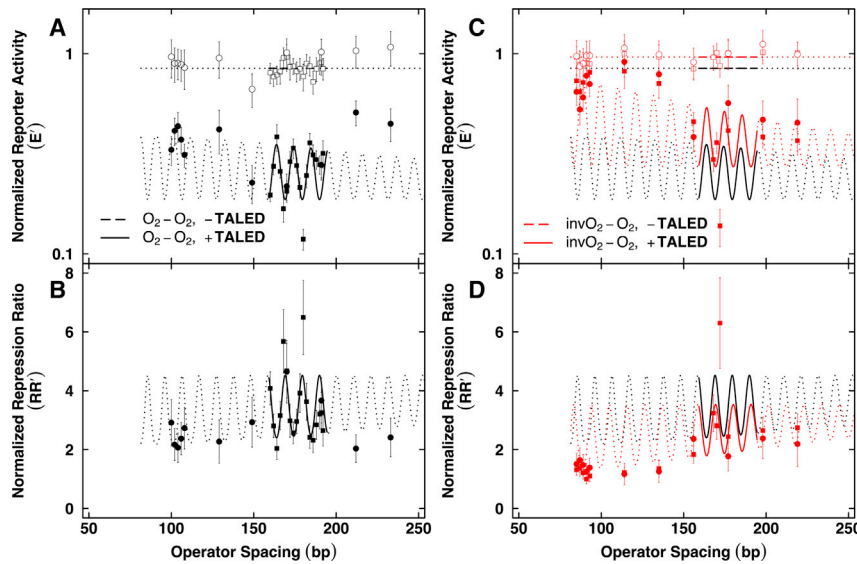


FIGURE 4 Thermodynamic model fits (*solid lines*) and predictions (*dotted lines*) as a function of operator spacing for transcriptional activity and repression data from single TALE loop configurations. (A) Best-fit model for normalized reporter activity data (as defined in [Materials and Methods](#)) for the indicated O_2 - O_2 operator spacings in the absence or presence of TALE protein is given. (B) Best-fit model for normalized repression ratio data of O_2 - O_2 operator spacings is given. (C) shows the same as in (A) except with data for $invO_2$ - O_2 configurations. Fit from (A) is in black for comparison. (D) shows the same as in (B) but for $invO_2$ - O_2 configurations. Fit from (B) is in black for comparison. To see this figure in color, go online.

observed for LacI looping (10.73 bp/turn). However, is clear from the 95% confidence interval ([Table 1](#)) that other fit values are possible for this parameter. The lower apparent helical repeat of the O_2 - O_2 loop series merits further investigation and potentially could suggest different geometrical considerations for the TALE looping system. The fit value of the torsional modulus of the looped DNA (C_{app}) also includes twist flexibility imparted on the system by the flexible TALE linker amino acids in the protein-DNA loops. Estimates of C_{app} (1.36 and 0.96) are lower than the common in vitro value of $2.4 [\times 10^{-19} \text{ erg-cm}]$ ([44–46](#)). This we attribute to the participation of the flexible TALE proteins as part of the loop, a result already well established for LacI looping ([10–15](#)). The normalized parameter K°_{max} is comparable for both data sets (3.09 vs. 2.64), consistent with the results of [Fig. 2](#) showing that operator orientation has minimal effect on the overall extent of repression. For virtually all spacings examined, the repression ratio is greater in the O_2 - O_2 loop series than the $invO_2$ - O_2 loop series ([Fig. 3, C and D](#)), which is captured in the model as a greater value of the normalized parameter K°_{NSL} . Finally, we note that $sp_{optimal}$ represents an optimal value (among others separated by hr) that sets the phasing for the oscillations. The model confirms that there is an optimal spacing for repression in each data set near 180 bp, consistent with visual inspection of the data.

However, one striking shortfall of the simple modeling is illustrated for operator spacings near 100 bp ([Fig. 4](#)). The model poorly captures E' minima but is satisfactory for maxima ([Fig. 4, A and C](#)), whereas the opposite is true for RR' (satisfactory fits at minima but poor fits at maxima; [Fig. 4, B and D](#)). We interpret this discrepancy as evidence for a second physical phenomenon not treated in the simple thermodynamic model: TALE dimerization. Unlike the stable LacI dimer of dimers, TALE dimerization in this de-

signed system is weak and likely insufficient to overcome the DNA bending strain required for small DNA loops. Indeed, there is precedent for dimerization-dependent looping behavior in the *E. coli gal* operon ([47](#)). GalR is a homolog of LacI but forms a weaker dimer of dimers. Thus, DNA looping requires that pairs of GalR dimers separately occupying two operators must overcome DNA stiffness to form this dimer of dimers (tetramerization), analogous to TALE formation in our system. Galactose is the natural inducer and appears to affect both tetramerization and DNA binding ([48](#)), whereas chemical inducers only affecting dimerization were described in our previous work ([32](#)) but are not studied here. Moreover, any repression looping system based on dimeric proteins that simultaneously bind two DNA sites will show a repression optimum that depends on the concentration of the protein dimer. This occurs because at high concentrations, preformed dimers may saturate both DNA sites without looping. Future experiments to estimate and systematically alter in vivo TALE concentrations will allow exploration of this variable.

Global analysis of TALE and LacI repression loops

Because of our past experience analyzing repression loops by LacI repressor ([10–15](#)), mixed competing TALE loop configurations ([32](#)), and the single TALE loop configurations described in this work, we have the opportunity for global comparison of the repression loop properties based on our data for these systems. Examples from each of these data sets are shown together in [Fig. 5, A and B](#).

In comparison with our previous *lac* studies, a direct example is provided by episome-based O_{sym} - O_{sym} constructs. The brown circles in [Fig. 5, A and B](#) show reporter activity for four operator spacings near 86 bp with or

TABLE 1 Thermodynamic Model Fits for Data with a Single Loop Configuration

Parameter ^c	TALED ^a		WT LacI ^b	
	O ₂ -O ₂	invO ₂ -O ₂	O _{sym} -O ₂ , -IPTG	O _{sym} -O ₂ , +IPTG
<i>hr</i> (bp/turn)	10.42 ± 0.20	10.75 ± 0.92	11.44 ± 0.74	10.73 ± 0.49
<i>C</i> _{app} (× 10 ⁻¹⁹ erg-cm)	1.36 ± 0.21	0.96 ± 1.06	0.76 ± 0.42	0.64 ± 1.11
<i>K</i> _{max} ^o	3.09 ± 0.52	2.64 ± 2.47	68.62	17.85
<i>K</i> _{NSL} ^o	1.71 ± 0.39	0.29 ± 2.43	10.45	0
<i>sp</i> _{optimal} (bp)	179.45 ± 0.38	180.10 ± 0.99	78.27 ± 0.53	78.82 ± 0.46
<i>K</i> _O	5.56	5.78	2.45	0.13
<i>K</i> _{max}	17.16	15.28	167.78 ± 60.60	2.39 ± 0.54
<i>K</i> _{NSL}	9.53	1.70	25.55 ± 31.29	0.00 ± 2.06
<i>K</i> _{O2} [-TALED]	0.08	0.12	N/A	N/A
<i>K</i> _{NS} [-TALED]	0.08	0.02	N/A	N/A
<i>K</i> _O	plasmid 9.82	plasmid 10.01	N/A	N/A
<i>K</i> _{max}	30.32	26.47	N/A	N/A
<i>K</i> _{NSL}	16.83	2.95	N/A	N/A
<i>K</i> _{O2} [-TALED]	0	0.02	N/A	N/A
<i>K</i> _{NS} [-TALED]	0.18	0.06	N/A	N/A

N/A, not applicable.

^aAll parameters are +TALED except when otherwise indicated.

^bThe normalized parameters were not used for fitting in Becker et al. (10).

^cParameters directly determined from fitting include a 95% confidence interval.

without TALED protein (*solid* versus *open* symbols). The green circles in Fig. 5, A and B show reporter activity from an *E. coli* strain with wild-type lac repressor (WT LacI, *solid* symbols) or a strain with a totally disabled lac repressor (LacI Y282D, *open* symbols). The latter is equivalent to the absence of TALED so that the maximal promoter activity in the absence of functional protein is the same for both. The extent of decreased report activity caused by protein-mediated looping (WT LacI, *green* versus TALED, *brown*) is very similar (*solid* symbols in Fig. 5 A), despite different anchoring proteins. Even though this particular operator configuration leads to very high levels of total repression (several hundredfold), the contribution from DNA looping is only modest (*RR'*-value of 2–6 in Fig. 5 B) because each protein binds the operator tightly even in the absence of a distal operator (*RR* of 113 vs. 75, respectively, for WT LacI versus TALED).

To better observe the effects of looping on repression, a majority of the previous work with lac repressor explored the combination of a strong distal operator sequence (O_{sym})

with a weak proximal operator sequence (O₂). In an episome, proximal O₂ alone with wild-type LacI accounted for an *RR*-value of ~3 relative to the absence of repressor (i.e., LacI Y282D mutant). A similar value was observed for WT LacI assayed with or without isopropyl β-D-1-thiogalactopyranoside (IPTG) inducer. *RR*-values for proximal O₂ alone with TALED are slightly larger, ~6 (episome) vs. ~11 (plasmid), and vary less than twofold based on reporter type. The blue circles (*open* versus *solid*) in Fig. 5 display lac repressor data (± IPTG inducer) for episomal constructs with operator spacings ranging from 60 to 90 bp. Here, several important observations can be made. From a total repression enhancement up to 100-fold, lac-repressor-mediated looping is the dominant contributor with an *RR'* of 5–33 and a mean of ~15 (Fig. 5 B), in contrast to the maximal TALED *RR'* of ~6. For clarity, Fig. 5, C and D display the thermodynamic model fits to the different data sets with a single loop configuration. By comparing the fit value of the normalized parameter *K*_{max}^o (Table 1), it is evident that the contribution of looping to repression is higher for lac repressor (~69) than for TALED (~3). However, each of the distal operator sequences was weak in the TALED data sets, with *RR* of either ~11 (proximal O₂ alone, plasmid) or ~8 (proximal invO₂ alone, plasmid). Creation of heterodimer TALEDs capable of binding both strong distal and weak proximal operator sequences could further tease out this effect.

Comparing repression by lac repressor and by TALEDs

We set out to apply artificial TALED proteins in the context of a set of promoter-reporter constructs previously assembled to study DNA looping by LacI repressor in vivo. Although this approach created some complications in our initial report (binding of the initial TALEs to the symmetrical lac O_{sym} operator can occur in either of two orientations), we now have defined TALEs that recognize asymmetrical operators so that single defined loops are created using distinct orientations of the O₂ operator sequence. This has allowed detailed measurement of the defined TALED loop system in living bacteria, facilitating comparison with loops driven by LacI repressor. As shown in Fig. 5, C and D, the classic oscillation pattern in our new, to our knowledge, data makes it unequivocal that the TALED system drives DNA looping. Fig. 5 D summarizes the observation that the contribution of looping to repression is higher for lac repressor than for TALEDs. The two systems have similarities and differences.

With respect to similarities, loops driven by LacI or TALEDs have in common that the optimal loop length is far smaller than expected for naked DNA (near four persistence lengths). Over the length scales we previously studied for LacI, DNA bending appears to be a much smaller obstacle than expected for looping, with DNA twist energy playing a more obvious role (10). We have interpreted this

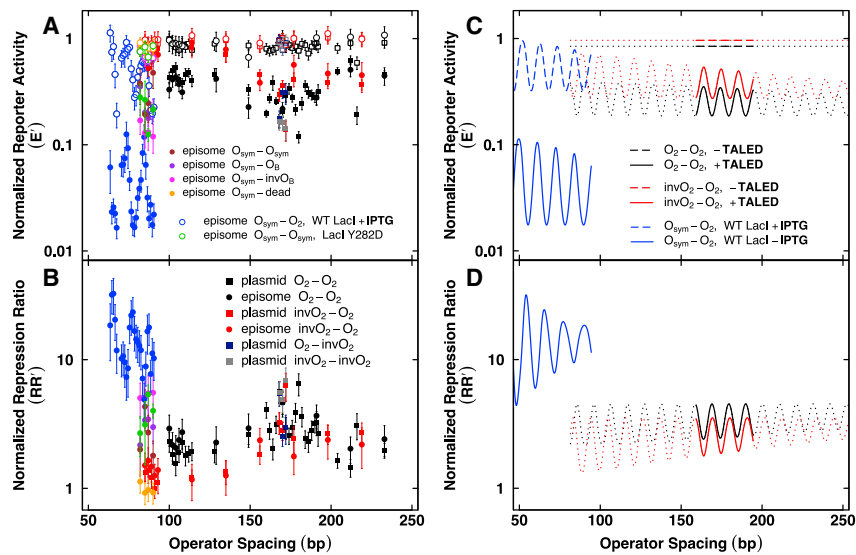


FIGURE 5 Compilation of repression data for DNA loops anchored by TALEDs or LacI repressor and comparison of thermodynamic model fits (solid lines) and predictions (dotted lines) for data from single configuration *lac* and TALED loops. (A) Normalized reporter activity for all indicated constructs from this and prior published reports from our laboratory is shown, including LacI repressor loops and all TALED single loop configurations (involving only O_2 and $invO_2$ operators) and multiple competing TALED loop configurations (involving O_{sym}). (B) Data are plotted as normalized repression ratio. (C) Best-fit model for reporter activity from the indicated cases is given. Black indicates fits for O_2 - O_2 operator spacings with TALED. Red indicates fits for $invO_2$ - O_2 operator spacings with TALED. Blue indicates fits for O_{sym} - O_2 operator spacings with LacI repressor. (D) Model fits for normalized repression ratio are given. To see this figure in color, go online.

as evidence for the effect of architectural DNA-binding proteins in vivo (10–12).

With respect to differences, our data sets show strong LacI repressor looping even for loops smaller than 100 bp in length, in agreement with prior studies (20), whereas minimal TALED looping is observed in the data collected near 100 bp. Several considerations may explain the different behavior of the two systems. First and foremost, whereas the LacI tetramer is a stable dimer of dimers, TALE dimerization via the FKBP(F36M) domain is weak, with an equilibrium dissociation constant in the tens of micromolar. This implies that loops involving expensive DNA deformation may not be supported by the weak TALE dimerization interface. This concept would explain relatively low repression and little torsional dependence for short operator spacings, then a regime of maximal repression near 175 bp that depends on the second operator with gradually decreasing torsional oscillation as operator spacing increases. Second, the concept of a small-molecule chemical inducer is different between the systems. Whereas *lac* induction involves a small metabolite that decreases repressor affinity for its DNA operators, TALED anti-repression by small molecules in our engineered system does not change operator affinity but alters protein dimerization (32). Third, occupancy of the promoter-proximal operator by a TALE protein, even if TALE dimerization has been blocked, results in higher basal repression for TALE monomer binding to this operator than for the LacI weakened by inducer binding. Fourth, although LacI binding undoubtedly changes the physical properties of the occupied operator DNA (49–51), DNA recognition by TALE protein wrapping of the operator major groove is expected to confer additional rigidity (31), constraining operator conformations within repression loops. This consideration, together with weak TALE dimerization, may explain the large optimal loop lengths observed

in the TALED system. Thus, whereas the apparent optimal DNA loop length for *lac* repressor was found to be near 80 bp (10,20), the maximum for TALEDs appears to be near 175 bp (Fig. 5).

The concept of an optimal DNA length for protein-mediated looping is, in itself, interesting. It is likely that looping optima reflect the same principles embodied in predictions of the wormlike chain polymer model for DNA cyclization (17,52). Because of its high relative stiffness, the effective end-end concentration (J) is extremely low for short DNA, rising rapidly to an optimal length for DNA cyclization before gradually falling for longer lengths because of entropic effects. These considerations based on the wormlike chain polymer model (53) allow estimation of apparent values of the DNA persistence length (which contains contributions from both proteins and negatively supercoiled DNA in vivo). These estimated persistence lengths are ~ 11 nm for LacI repressor and ~ 17 nm for the TALE homodimer studied here, contrasting with ~ 50 nm for DNA in vitro. The different optimal looping DNA lengths may reflect the considerations raised above or even the possibility that protein-DNA loops have evolved to recruit architectural DNA-binding proteins that reduce apparent DNA stiffness (10–12,16).

SUPPORTING MATERIAL

Supporting Material can be found online at <https://doi.org/10.1016/j.bpj.2020.10.007>.

AUTHOR CONTRIBUTIONS

Project conception and manuscript preparation, N.A.B., J.P.P., K.J.C., and L.J.M.; experimental work, N.A.B., J.P.P., T.L.S., J.P.W., and W.J.P.

ACKNOWLEDGMENTS

We thank members of the Maher laboratory for assistance.

This work was supported by the Mayo Foundation, the Mayo Clinic Graduate School of Biomedical Sciences Summer Undergraduate Research Fellowship program, and by the National Institutes of Health (grant no. GM75965 to L.J.M.).

REFERENCES

- Travers, A., and G. Muskhelishvili. 2007. A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep.* 8:147–151.
- Rojo, F. 2001. Mechanisms of transcriptional repression. *Curr. Opin. Microbiol.* 4:145–151.
- Dalma-Weiszhausz, D. D., and M. Brenowitz. 1996. Interactions between RNA polymerase and the positive and negative regulators of transcription at the *Escherichia coli* gal operon. *Biochemistry.* 35:3735–3745.
- Struhl, K. 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell.* 98:1–4.
- Müller-Hill, B. 1996. The Lac Operon: A Short History of a Genetic Paradigm. Walter de Gruyter, Berlin, Germany.
- Ptashne, M. 2004. A Genetic Switch: Phage Lambda Revisited, Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Straney, S. B., and D. M. Crothers. 1987. Lac repressor is a transient gene-activating protein. *Cell.* 51:699–707.
- Hochschild, A., and M. Ptashne. 1986. Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell.* 44:681–687.
- Johnson, A. D., B. J. Meyer, and M. Ptashne. 1979. Interactions between DNA-bound repressors govern regulation by the lambda phage repressor. *Proc. Natl. Acad. Sci. USA.* 76:5061–5065.
- Becker, N. A., J. D. Kahn, and L. J. Maher, III. 2005. Bacterial repression loops require enhanced DNA flexibility. *J. Mol. Biol.* 349:716–730.
- Becker, N. A., J. D. Kahn, and L. J. Maher, III. 2007. Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Res.* 35:3988–4000.
- Becker, N. A., J. D. Kahn, and L. J. Maher, III. 2008. Eukaryotic HMGB proteins as replacements for HU in *E. coli* repression loop formation. *Nucleic Acids Res.* 36:4009–4021.
- Bond, L. M., J. P. Peters, ..., L. J. Maher, III. 2010. Gene repression by minimal lac loops in vivo. *Nucleic Acids Res.* 38:8072–8082.
- Becker, N. A., J. P. Peters, ..., T. A. Lionberger. 2013. Mechanism of promoter repression by Lac repressor-DNA loops. *Nucleic Acids Res.* 41:156–166.
- Becker, N. A., A. M. Greiner, ..., L. J. Maher, III. 2014. Bacterial promoter repression by DNA looping without protein-protein binding competition. *Nucleic Acids Res.* 42:5495–5504.
- Becker, N. A., and L. J. Maher, III. 2015. High-resolution mapping of architectural DNA binding protein facilitation of a DNA repression loop in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 112:7177–7182.
- Peters, J. P., III, and L. J. Maher, III. 2010. DNA curvature and flexibility in vitro and in vivo. *Q. Rev. Biophys.* 43:23–63.
- Oehler, S., E. R. Eismann, ..., B. Müller-Hill. 1990. The three operators of the lac operon cooperate in repression. *EMBO J.* 9:973–979.
- Oehler, S., M. Amouyal, ..., B. Müller-Hill. 1994. Quality and position of the three lac operators of *E. coli* define efficiency of repression. *EMBO J.* 13:3348–3355.
- Müller, J., S. Oehler, and B. Müller-Hill. 1996. Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J. Mol. Biol.* 257:21–29.
- Müller-Hill, B. 1998. The function of auxiliary operators. *Mol. Microbiol.* 29:13–18.
- Mossing, M. C., and M. T. Record, Jr. 1986. Upstream operators enhance repression of the lac promoter. *Science.* 233:889–892.
- Bellomy, G. R., M. C. Mossing, and M. T. Record, Jr. 1988. Physical properties of DNA in vivo as probed by the length dependence of the lac operator looping process. *Biochemistry.* 27:3900–3906.
- Law, S. M., G. R. Bellomy, ..., M. T. Record, Jr. 1993. In vivo thermodynamic analysis of repression with and without looping in lac constructs. Estimates of free and local lac repressor concentrations and of physical properties of a region of supercoiled plasmid DNA in vivo. *J. Mol. Biol.* 230:161–173.
- Oehler, S., and B. Müller-Hill. 2010. High local concentration: a fundamental strategy of life. *J. Mol. Biol.* 395:242–253.
- Schleif, R. 1987. Gene regulation: why should DNA loop? *Nature.* 327:369–370.
- Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* 61:199–223.
- Rippe, K. 2001. Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.* 26:733–740.
- Rippe, K., P. H. von Hippel, and J. Langowski. 1995. Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* 20:500–506.
- Bogdanove, A. J., and D. F. Voytas. 2011. TAL effectors: customizable proteins for DNA targeting. *Science.* 333:1843–1846.
- Mak, A. N.-S., P. Bradley, ..., B. L. Stoddard. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science.* 335:716–719.
- Becker, N. A., T. L. Schwab, ..., L. J. Maher, III. 2018. Bacterial gene control by DNA looping using engineered dimeric transcription activator like effector (TALE) proteins. *Nucleic Acids Res.* 46:2690–2696.
- Hao, N., K. E. Shearwin, and I. B. Dodd. 2017. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun.* 8:1628.
- Kim, J. H., M. Rege, ..., J. E. Phillips-Cremens. 2019. LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods.* 16:633–639.
- Morgan, S. L., N. C. Mariano, ..., K. C. Wang. 2017. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* 8:15993.
- Whipple, F. W. 1998. Genetic analysis of prokaryotic and eukaryotic DNA-binding proteins in *Escherichia coli*. *Nucleic Acids Res.* 26:3700–3706.
- Peters, J. P., N. A. Becker, ..., L. J. Maher, III. 2011. Quantitative methods for measuring DNA flexibility in vitro and in vivo. *Methods Enzymol.* 488:287–335.
- Ma, A. C., M. S. McNulty, ..., S. C. Ekker. 2016. FusX: a rapid one-step transcription activator-like effector assembly system for genome science. *Hum. Gene Ther.* 27:451–463.
- Miller, J. H. 1992. A Short Course in Bacterial Genetics. Cold Spring Harbor Laboratory Press, New York.
- Rollins, C. T., V. M. Rivera, ..., T. Clackson. 2000. A ligand-reversible dimerization system for controlling protein-protein interactions. *Proc. Natl. Acad. Sci. USA.* 97:7096–7101.
- Czapla, L., M. A. Grosner, ..., W. K. Olson. 2013. Interplay of protein and DNA structure revealed in simulations of the lac operon. *PLoS One.* 8:e56548.
- Mulligan, P. J., Y. J. Chen, ..., A. J. Spakowitz. 2015. Interplay of protein binding interactions, DNA mechanics, and entropy in DNA looping kinetics. *Biophys. J.* 109:618–629.
- Zhang, Y., A. E. McEwen, ..., S. D. Levene. 2006. Analysis of in-vivo LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS One.* 1:e136.
- Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* 170:957–981.

45. Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting. II. Topoisomer analysis. *J. Mol. Biol.* 170:983–1007.
46. Shore, D., J. Langowski, and R. L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. USA.* 78:4833–4837.
47. Lewis, D. E., and S. Adhya. 2015. Molecular mechanisms of transcription initiation at gal promoters and their multi-level regulation by GalR, CRP and DNA loop. *Biomolecules.* 5:2782–2807.
48. Semsey, S., K. Virnik, and S. Adhya. 2005. A gamut of loops: meandering DNA. *Trends Biochem. Sci.* 30:334–341.
49. Colasanti, A. V., M. A. Grosner, ..., W. K. Olson. 2013. Weak operator binding enhances simulated Lac repressor-mediated DNA looping. *Biopolymers.* 99:1070–1081.
50. Lewis, M., G. Chang, ..., P. Lu. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science.* 271:1247–1254.
51. Swigon, D., B. D. Coleman, and W. K. Olson. 2006. Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. *Proc. Natl. Acad. Sci. USA.* 103:9879–9884.
52. Shimada, J., and H. Yamakawa. 1984. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules.* 17:689–698.
53. Johnson, S., M. Lindén, and R. Phillips. 2012. Sequence dependence of transcription factor-mediated DNA looping. *Nucleic Acids Res.* 40:7728–7738.