

Analysis of *In-Vivo* LacR-Mediated Gene Repression Based on the Mechanics of DNA Looping

Yongli Zhang^{1*}, Abbye E. McEwen², Donald M. Crothers¹, Stephen D. Levene^{2,3*}

1 Departments of Chemistry and Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, United States of America, 2 Institute of Biomedical Sciences and Technology, University of Texas at Dallas, Richardson, Texas, United States of America, 3 Department of Molecular and Cell Biology, University of Texas at Dallas, Richardson, Texas, United States of America

Interactions of *E. coli* lac repressor (LacR) with a pair of operator sites on the same DNA molecule can lead to the formation of looped nucleoprotein complexes both *in vitro* and *in vivo*. As a major paradigm for loop-mediated gene regulation, parameters such as operator affinity and spacing, repressor concentration, and DNA bending induced by specific or non-specific DNA-binding proteins (e.g., HU), have been examined extensively. However, a complete and rigorous model that integrates all of these aspects in a systematic and quantitative treatment of experimental data has not been available. Applying our recent statistical-mechanical theory for DNA looping, we calculated repression as a function of operator spacing (58–156 bp) from first principles and obtained excellent agreement with independent sets of *in-vivo* data. The results suggest that a linear extended, as opposed to a closed v-shaped, LacR conformation is the dominant form of the tetramer *in vivo*. Moreover, loop-mediated repression in wild-type *E. coli* strains is facilitated by decreased DNA rigidity and high levels of flexibility in the LacR tetramer. In contrast, repression data for strains lacking HU gave a near-normal value of the DNA persistence length. These findings underscore the importance of both protein conformation and elasticity in the formation of small DNA loops widely observed *in vivo*, and demonstrate the utility of quantitatively analyzing gene regulation based on the mechanics of nucleoprotein complexes.

Citation: Zhang Y, McEwen AE, Crothers DM, Levene SD (2006) Analysis of *In-Vivo* LacR-Mediated Gene Repression Based on the Mechanics of DNA Looping. PLoS ONE 1(1): e136. doi:10.1371/journal.pone.0000136

INTRODUCTION

The *lac* operon of *E. coli* provides an important paradigm for gene regulation [1], in which DNA looping is a central aspect of transcriptional repression. The gene products of the *lac* operon are three enzymes important for metabolism of lactose, an alternative cellular energy source. In the wild-type *lac* operon there are three *lac* repressor (LacR) binding sites, or operators: one primary operator (O_1) located at the +11 position relative to the start of transcription, and two auxiliary operators located 92 bp upstream (O_3) and 401 bp downstream (O_2) relative to the primary binding site. DNA looping between the primary operator and either of the auxiliary operators enhances occupancy of the primary site by LacR [2,3], thereby blocking transcription by preventing RNA polymerase binding to the promoter.

The Record [4] and Müller-Hill [5] groups reported classic studies of repression as a function of the helical phasing or DNA length between a primary and one auxiliary *lac* operator, providing early evidence for DNA looping as a mode of transcriptional control. These results and those of studies involving other proteins [6] have led to a long-standing question: how can DNA loops shorter than 100 bp form efficiently *in vivo*, given the large energy barrier created by strong DNA bending and/or twisting deformations [7,8]? The prevailing explanation is that a DNA molecule has greater apparent flexibility *in vivo* and thus the actual DNA bending and twisting energy for loop formation is lower than that estimated from *in-vitro* DNA-elasticity parameters. Such enhanced apparent flexibility could be attributed to nonlinear behavior of DNA elasticity accompanying strong DNA distortion [9–13], or result from dynamic and non-specific protein binding and bending [14–19]. Indeed, Becker *et al.* [19] addressed the latter hypothesis directly by investigating effects of deletion of both genes that encode subunits for the non-specific DNA-binding protein HU. They found that loop-mediated repression mediated by LacR was substantially reduced in HU-deletion strains and that this

phenotype could be partially rescued by ectopic expression of the human DNA-bending protein HMG.

Analyses of DNA looping often rest on the assumption that the proteins mediating the loop are rigid and play no active role in looping other than providing end constraints at DNA binding sites. However, recent experimental [20,21] and theoretical studies [22] have questioned this assumption and suggest that both protein geometry and flexibility play important roles in the formation of small DNA loops. Protein conformational flexibility can potentially lower the free energy of DNA bending and twisting required for loop formation; if the protein assembly is sufficiently flexible neither enhanced DNA flexibility nor protein-induced bending promoted by additional factors may be needed to stabilize small loops.

.....
Academic Editor: Peter Fraser, The Babraham Institute, United Kingdom

Received August 30, 2006; **Accepted** November 30, 2006; **Published** December 27, 2006

Copyright: © 2006 Zhang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Joint DMS/NIGMS Mathematical Biology Initiative (NIH GM 67242 to SDL and I.K. Darcy), by NIH grant GM 21966 to DMC, and by the National Foundation for Cancer Research through the Yale-NFCR Center for Protein and Nucleic Acid Chemistry.

Competing Interests: The authors have declared that no competing interests exist.

* **To whom correspondence should be addressed.** E-mail: sdlevene@utdallas.edu (YZ); ylzhang@lbl.gov (SDL)

‡ **Current address:** Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York, United States of America

A full understanding of the role of DNA looping in gene repression requires a complete and rigorous analysis of the plethora of data obtained from *in-vivo* experiments. Previous analyses [2,4,5,19,23,24] have several limitations. In addition to neglecting mechanical contributions from protein flexibility, results are often analyzed by treating DNA looping as being quantitatively equivalent to the related process of DNA cyclization [25–31]. We have shown that important distinctions exist between these two processes and that neglecting these differences can potentially lead to misinterpretation of the helical-phase dependence of looping, for example [22]. The major obstacle to quantitatively analyzing experimental data has thus been lack of an accurate and computationally efficient theory for DNA looping [7].

Here we describe a comprehensive analysis of the thermodynamics of LacR-mediated repression, including a rigorous statistical-mechanical theory for DNA loop closure [22]. Our treatment considers the mechanics of a protein-mediated loop in terms of a rigid-body approximation that applies both to the base pairs of DNA and to the protein domains that constitute the nucleoprotein assembly. DNA conformations in this model are parameterized using three conventional angular parameters: tilt, roll, and twist, corresponding to rotations of a base pair about the x , y , and z axes, respectively, of a conventionally chosen local Cartesian-coordinate frame [28]. The geometric arrangement of protein domains is specified by using a similar local coordinate frame fixed within each rigid-body entity of a protein structure (Figure 1). Interaction potentials between base-pair steps and protein domains are taken as quadratic forms in the angular displacements from mechanical equilibrium in the absence of loop-closure constraints. This model therefore allows for conformational flexibility among protein domains and within protein-DNA contacts. We compute the mechanical minimum-energy conformation of the protein-mediated loop and calculate thermodynamic quantities by including thermal fluctuations about this conformation through a harmonic approximation [29]. The approach has many advantages over previous methods in terms of accuracy, computational efficiency, and versatility. It has been applied successfully to the analysis of DNA cyclization data as a special case and provides a basis for understanding the general principles that govern loop-mediated protein-DNA interactions [22]. Swigon et al. [32] recently considered *in-vitro* LacR-mediated DNA looping using a similar strategy [29], although it is not clear to what extent the entropy of particular LacR conformations was considered. Here we extend our approach to investigate LacR-dependent, DNA-loop-regulated gene repression *in vivo*.

Several crystal structures of LacR and the LacR-operator complex, shown in Figure 1, reveal that the repressor can be considered as a dimer of dimers [33,34]. Each LacR monomer consists of a DNA-binding headpiece, a core domain, and a tetramerization domain. In the crystal structures, a “v”-shaped tetramer is formed from two dimers via a four-helix bundle that comprises the tetramerization domain. This structure has the DNA-binding domains symmetrically placed about a two-fold or dyad rotational axis that lies in the plane of the “v.” The interior angle between the two LacR dimers is about 60° and protein binding induces a local 45° kink in the DNA. However, electron microscopy [35] showed that 44% of LacR in solution is present in an extended conformation (~180° between the two arms, Figure 1C), with the remaining 56% of complexes in the “v” shape. Additional solution studies support the existence of an extended LacR conformation in small loops containing intrinsically bent DNA sequences [36,37]. A reasonable way of reconciling the discrepancies is to assume an inherently bistable

structure for LacR such that v-shaped and extended conformations can exist in equilibrium. Assuming that the binding affinity to operator DNA is independent of LacR conformation [32], the proportion of each repressor structure in LacR-mediated DNA loops depends on the sum of the free energies arising from DNA and protein distortion during loop formation.

RESULTS

Conformational model for the LacR tetramer

Based on the symmetry and modular structure of the LacR tetramer, we model the protein as a dimeric assembly consisting of rigid-body domains connected by semiflexible joints (Figures 1C, D). There are three sets of protein-related rotation angles in addition to those for the DNA dinucleotide steps: two sets for the contacts made by protein domains with the last and first base pairs of the DNA and one set for the contact between protein dimers [22]. These angles describe the kinematics of protein domains joined at the positions shown in Figure 1C. Nearest-neighbor interactions between protein dimers, dinucleotide steps, and between protein domains and DNA are governed by harmonic potentials (see Equation 3 in Materials and Methods) with thermal fluctuations of each DNA base pair expressed in terms of standard deviations of the corresponding angular parameters from their static values. For homogeneous DNA, the standard deviations, σ , for tilt and roll are identical and related to DNA bending persistence length, P , by $\sigma = (1/P)^{1/2}$ where P is given in base pairs and σ is expressed in radians. The deformability of the protein assembly in this model is similarly specified in terms of standard deviations of the protein-DNA and protein-protein tilt/roll/twist rigid-body parameters.

Following previous observations, we focus on the two canonical LacR geometries: the v-shaped structure characterized by an interior angle of 60° (Figures 1A, B) and the extended tetramer structure, with a 180° interior angle between dimers (Figure 1C). Note, however, that because of strain within the loop, the equilibrium value of this angle is not generally identical to that in the absence of constraints (Figure 1D) [22].

The main thermodynamic quantity to be evaluated is the J factor (see Equation 5 in Materials and Methods), defined by Jacobson and Stockmayer as a measure of the circularization propensity of linear polymer chains [38]. The J factor can be understood in several equivalent ways: (i) as a quantity proportional to the equilibrium constant for formation of a closed chain from an open chain. This process requires association of two chain ends with a consequent reduction from six translational degrees of freedom to three [29]; the J factor thus has units of concentration. With this interpretation it is clear that the free energy of DNA looping is given by $\Delta G_{loop} = -k_B T \ln J$. Note, however, that this formulation of the J factor omits the thermodynamic contribution from protein-DNA association. (ii) The effective concentration of one end of a chain in the vicinity of the other. In the particular case of DNA looping that we discuss here, the J factor is the effective concentration of an auxiliary operator-bound LacR molecule at the primary operator. Due to the tethering effect of DNA looping this concentration can be much higher than the bulk free LacR concentration [2,7], leading to increased occupancy of the primary site by the repressor and enhanced gene repression. (iii) As the ratio of statistical-mechanical partition functions for closed and open chains [29].

Multiple DNA loop conformations

For the v-shaped LacR, there are three classes of mechanical-equilibrium looped conformations (Figure 2) [33], denoted “WT”

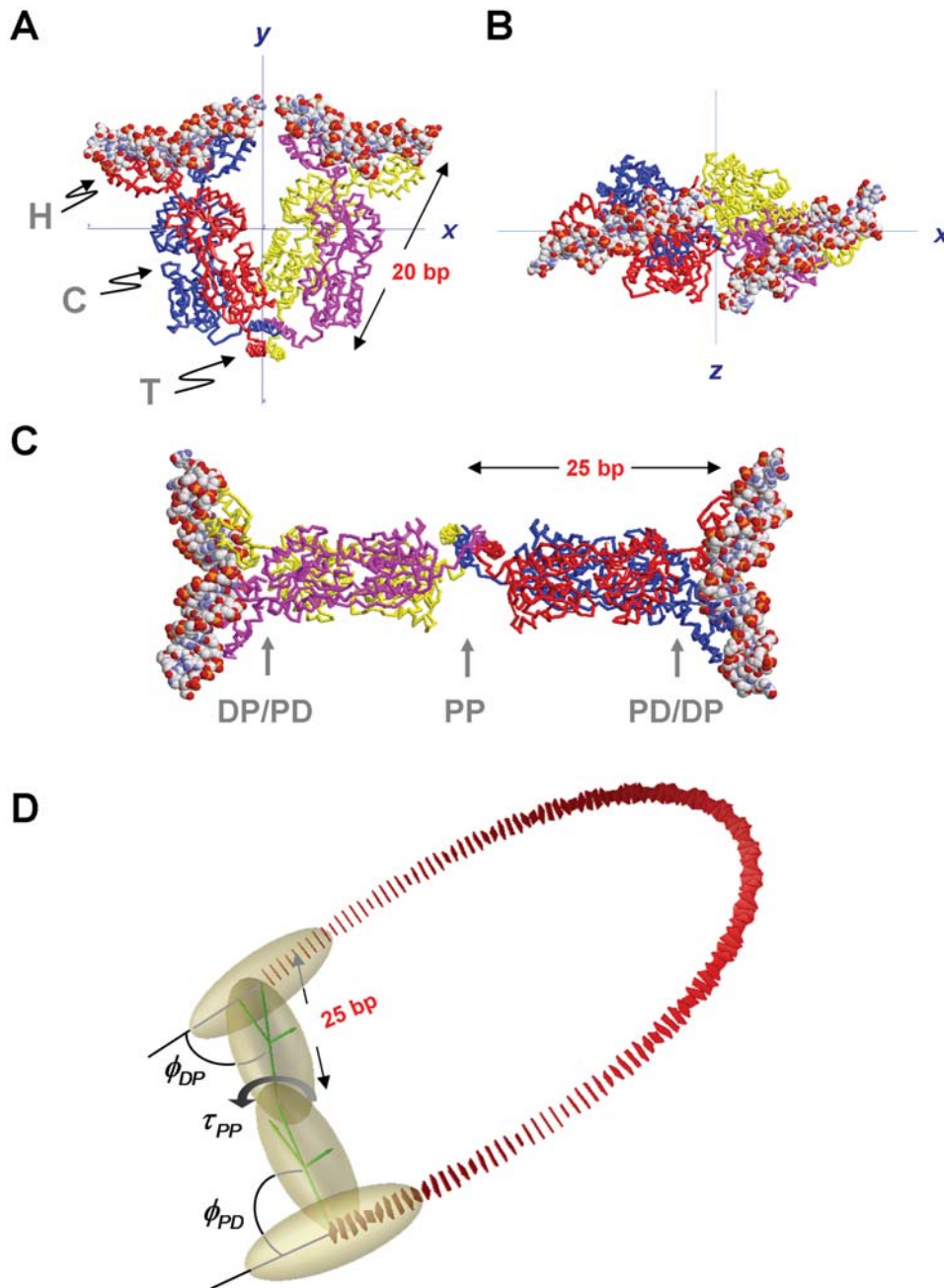


Figure 1. Structures of the *lac*-repressor tetramer complexed with DNA. (A) Crystallographic structure of the “v”-shaped LacR complexed with the symmetric operator sequence GAATTGTGAGCGCTACAATT (PDF accession number 1LBG) [34], shown along the z axis. The α -carbon trace of each lac monomer is rendered in a separate color; DNA segments are shown in a space-filling representation. The three repressor domains are indicated: *H*, headpiece; *C*, core; *T*, tetramerization. The *x* axis is an approximate two-fold or dyad axis in the structure. (B) View of the complex shown in (A) along the *x* axis. Helical axes of the bound DNA segments project slightly (35°) out of the plane of the “v” structure, implying that a small degree of DNA writhe may be induced by LacR-mediated looping. (C) A hypothetical structure of LacR in its extended conformation. This model was generated from the “v”-shaped structure shown in (A) by increasing the interior angle from 60° to $\sim 180^\circ$. The three semi-flexible joints modeling the elastic properties of the tetramer are indicated by vertical arrows. Note that an increase in the length of the LacR major axis from 20 bp to 25 bp occurs when the tetramer isomerizes from the “v-shaped” to the extended structure. (D) Simplified elastic model for LacR and a simulated 137-bp DNA loop mediated by the extended LacR structure. DNA base pairs are represented by rectangular slabs (red). Two sets of coordinate axes (green) represent the local coordinate frames embedded in the protein subunits (gold) that mediate DNA looping. The coupling of protein and DNA geometry is characterized by tilt, roll, and twist values for the DNA-protein, protein-protein, and protein-DNA interfaces. Three of these variables are shown here: the DNA-protein roll angle, ϕ_{DP} ; the protein-protein twist angle, τ_{PP} ; and the protein-DNA roll angle, ϕ_{PD} (see Materials and Methods for details). doi:10.1371/journal.pone.0000136.g001

(“wrapping towards”), “WA” (“wrapping away”), and “LB” (“looping beside”), depending on the DNA trajectory’s approach relative to the inside or outside of the “v.” Because a pseudo-dyad

axis is shared by the LacR DNA-binding headpiece and the operator sequence, operator-binding affinity should be independent of the local orientation of the DNA binding site. This

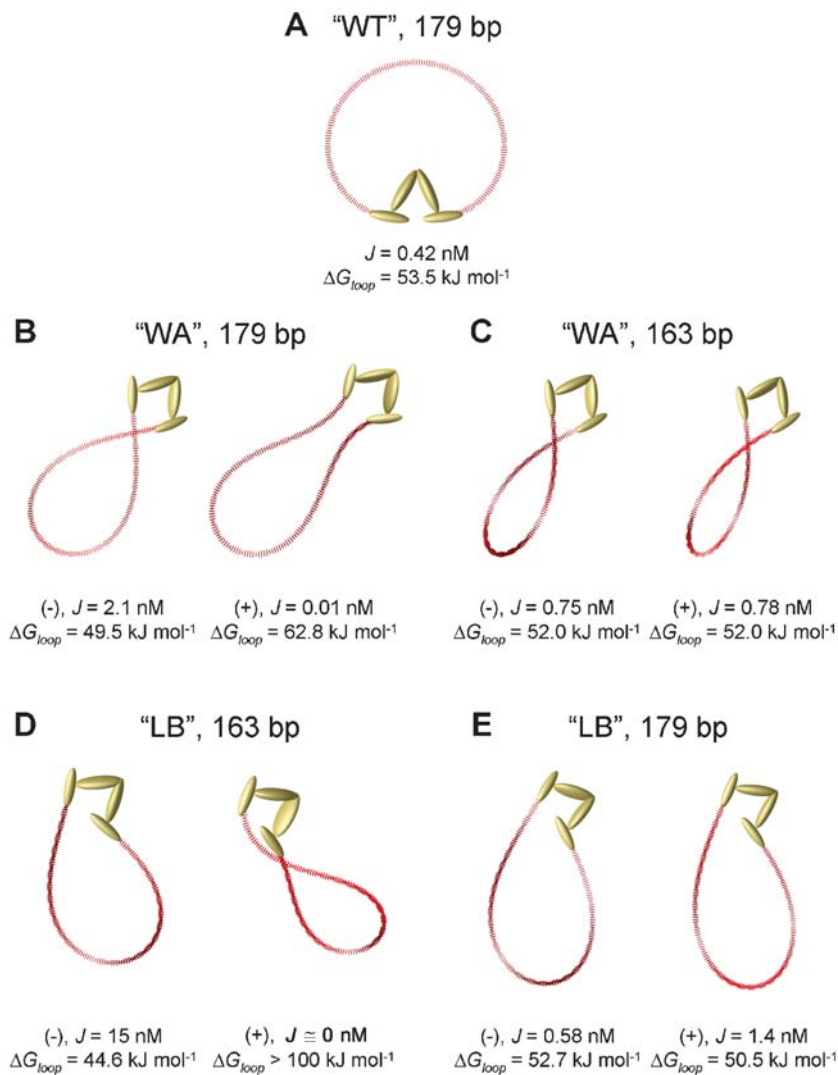


Figure 2. Looped DNA conformations mediated by the v-shaped LacR tetramer structure for 179- and 163-bp DNAs. Three classes of loop conformations: (A) "wrap toward" ("WT"), (B), (C) "wrap away" ("WA"), and (D), (E) "loop beside" ("LB"), are shown along with their respective J factors and ΔG_{loop} values. At most two alternative loop topoisomers are found for small DNA loops such as those considered here; these correspond to underwound (–) and overwound (+) DNA conformations. The DNA-loop sizes 179 and 163 bp correspond to maxima and minima, respectively, in the DNA-length dependence of J for the "WA" or "WT" conformations (see Figure 3A). The 179-bp "WT" conformation shown in (A) is dominated by one planar topoisomer because of its in-phase LacR binding sites at this DNA length; however, there are two alternative 179-bp "WA" conformations, which are shown in (B). The two "WA" topoisomers for 163-bp DNA are essentially isoenergetic and shown in (C); note that the loop crossings differ in topological sign. In contrast, the "LB" conformations (D) and (E) have minimum and maximum J values at 179 and 163 bp, respectively. Only the 179-bp (–) "LB" topoisomer shown in (D) is populated to any appreciable extent at thermal equilibrium, whereas both (–) and (+) forms have similar free energies in the case of 163-bp "LB" loops.

doi:10.1371/journal.pone.0000136.g002

property is expected even though the operator sequence may not itself be palindromic. In general, each class consists of a pair of overtwisted and undertwisted topoisomer solutions [22]. These three classes of loop conformations were also found by Olson and coworkers [32], described using different nomenclature [39].

For our calculations, we used a planar v-shaped structure to represent the repressor. In the crystal structure of the LacR-operator complex, the helical axes of the operator sites do not lie in the mean plane of the repressor structure and are instead displaced by about 20 degrees (Figure 1B). However, we found that J factors were relatively insensitive to this angle (see Figure S1). In contrast to the v-shaped LacR structure, there is only one class of "simple loops" ("SL") formed by the extended LacR tetramer (Figure 1D) [22,32].

Helical dependence of DNA looping for different LacR conformations

The computed J factors for the three classes of v-shaped loop conformations are shown as functions of loop size (or DNA length) in Figure 3A. Values of \bar{J} for particular conformations and corresponding values of the looping free energy are also given in Figure 2. Remarkably, the LB conformation has the largest \bar{J} value among the three classes of v-shaped protein structures and dominates the distribution (Figure 3A). There is a one-half-turn difference in the helical-phase dependence of \bar{J} for LB conformations relative to those for the WA and WT conformations. The difference in phasing arises because the LB conformation involves a 180-degree rotation of one operator element about the sequence

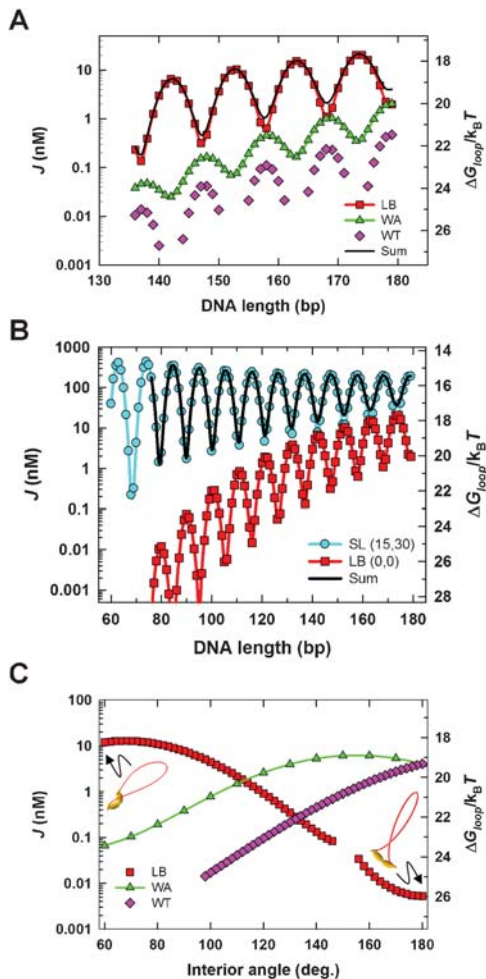


Figure 3. J factor and ΔG_{loop} values versus DNA length or LacR interior angle for four classes of loop conformations. (A) Comparison of J factors and ΔG_{loop} for the three classes of loop conformations mediated by the v-shaped *lac* repressor. Protein assemblies are taken to be rigid (i.e., DNA-protein, protein-protein, and protein-DNA flexibility parameters were all set to 0). (B) Length dependence of J factors and ΔG_{loop} for the extended (SL) and LB conformations. Protein-flexibility parameters are given in parentheses as $(\sigma_0^{PP} = \sigma_\phi^{PP} = \sigma_\tau^{PP}, \sigma_0^{DP} = \sigma_\phi^{DP} = \sigma_\tau^{DP} = \sigma_\phi^{PD} = \sigma_\tau^{PD} = \sigma_\tau^{DP})$. Taken together with (A), the J-factor length dependence shows that the extended LacR conformation dominates all of the v-shaped forms for loops smaller than 180 bp. (C) The dependence of J and ΔG_{loop} values on the interior angle between LacR domains is shown for three classes of 153-bp loops as the repressor structure opens from the near-crystallographic value of 60° to an extended form (180°). Protein assemblies were taken to be rigid, as in (A). WA and WT loops become degenerate at large angles, which can be seen from the identical J factors attained with the extended form of LacR. A small difference ($\sim 1.5k_B T$) between the asymptotic ΔG_{loop} value for WA and WT conformations in (C) and the corresponding value on the SL curve in (B) is due to differences in protein flexibility and tetramer dimensions (dimer major-axis length of 25 bp in (B) versus 20 bp in (C)). Because of broken symmetry, LB loops adopt a highly strained conformation as the interior angle approaches 180°. For comparison, projections of 3-d conformations for LB loops with interior-angle values of 60° and 180° are shown as insets. Gaps in the curves indicate that no stable mechanical-equilibrium conformations were found for “LB” loops when the interior angle was between 146° and 156°, nor for “WT” loops having interior-angle values less than 98°. This behavior is characteristic of abrupt transitions between mechanical minima, usually when a loop bifurcates to either an over-twisted or under-twisted conformation. Without a stable mechanical-equilibrium conformation, the perturbation method employed in our statistical-mechanical theory cannot be applied.

doi:10.1371/journal.pone.0000136.g003

dyad with respect to the WA and WT conformations. Moreover, the amplitude of the helical-phase dependence for the LB conformation is significantly larger than that for the WA conformation. This dependence of J-factor amplitude on loop conformation militates against the general use of empirical formulas based on DNA-cyclization theory to estimate the DNA torsional rigidity [19,24,28] and underscores the need to explicitly consider protein geometry and mechanics in models of DNA looping [22].

The length dependence of \bar{J} for SL loops mediated by the extended LacR structure is shown in Figure 3B. J factors for the SL conformation greatly exceed those for even the most thermodynamically favorable v-shaped conformation, LB. This difference between SL and LB loops is particularly pronounced for DNA loop sizes less than 100 bp, which is the range used in many studies of *in-vivo* gene repression regulated by DNA looping. The comparison in Figure 3B involves different protein-flexibility parameters for the two tetramer structures. Because the v-shaped tetramer is locked in place by interactions between the central domains of the two LacR dimers, it is likely that conformational flexibility in this compact conformation is substantially less than that of the extended conformation in which these interactions have been broken [34,40,41].

Figure 3B also shows that differences in protein structure and conformational flexibility dramatically alter the balance between elastic energy and chain entropy in loop formation as a function of DNA length [22,29]. There is a small, but significant, decrease in chain entropy with increasing loop size for the formation of SL-class loops, indicated by the decay of J-factor peaks with increasing DNA length. This increase in looping free energy stands in contrast to other results [32]. In the case of loops mediated by the v-shaped LacR structure, however, J factors increase with DNA length, demonstrating that these structures are determined by loop elastic energy. The phase dependence of the SL conformation is the same as that for the WT and WA structures and is one-half-turn out of phase relative to the LB conformation. This ~ 5 -bp difference in phasing between SL and LB loops implies that loop sizes that are J-factor minima in the SL length dependence closely coincide with J-factor maxima for the LB conformation.

Solutions for WT and WA conformations, but not those for LB, are expected to approach those for the extended repressor conformation in the limit where the LacR interior angle approaches 180 degrees. We examined J factors for a 153-bp loop formed by each of the three v-shaped structures as a function of the interior angle (Figure 3C). The results show that as the angle opens up from the near-crystallographic value of 60 degrees to the fully extended state (180 degrees, see Figure 1C), the LB conformation becomes increasingly unfavorable whereas WA and WT structures become increasingly favorable. Fully extended, the WA and WT structures are degenerate; as expected, they have identical \bar{J} values and looping free energies. Unlike the WA and WT loops, increasing the interior angle drives the “LB” structure toward the conformation of a loop with approximately parallel ends (in contrast to the approximately antiparallel ends in Figure 1D). Such strained conformations have dramatically diminished J factors.

The J factor is a direct measure of the relative proportions of particular looped conformations at thermodynamic equilibrium. In principle, J factors for all classes of loop conformations should be taken into account in calculating the free energy of LacR-mediated loop formation. However, based on the comparative magnitude of J factors for the SL and v-shaped repressor structures (Figure 3B), we chose to simplify our analysis of *in-vivo* repression data by using \bar{J} values for the SL loop class exclusively. In doing so

we neglected the possible free-energy difference for LacR tetramers in the two conformations, which has not been accurately measured, but may be relatively small [32,35].

Thermodynamic model for *in-vivo* gene repression regulated by LacR-mediated DNA looping

The energetics of loop formation depend not only on the geometry and mechanical properties of protein and DNA expressed in terms of the *J* factor, but also on the binding equilibria relating different protein-DNA association states (Figure 4). In experiments of Müller *et al.* [5], the observable quantity is expression of a reporter gene (e.g., β -galactosidase) as a function of variables such as operator spacing or operator affinity for LacR. To quantitatively analyze gene repression based on a model for DNA looping, we assume that the rate of reporter-gene expression is under thermodynamic control, namely, proportional to the probability that the primary operator is unbound. This assumption has been used in previous analyses of LacR-mediated gene repression [4,42,43].

Based on the equilibria shown in Figure 4 and derivations given in Materials and Methods, the enhancement of gene repression by DNA looping, *R*, is calculated according to the formula

$$R = \frac{E_{no\ loop}}{E_{loop}} = 1 + \frac{\lambda J P_1}{(K_1 + P_1)(K_2 + P_1)} \equiv 1 + \Gamma J \quad (1)$$

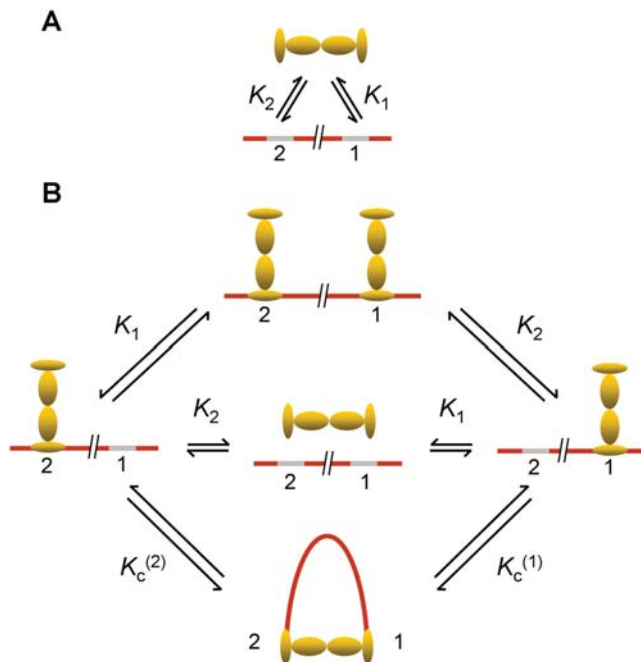


Figure 4. Chemical equilibria among LacR-operator association states assuming a single looped conformation for the LacR-DNA complex. (A) A LacR tetramer can bind to two cognate sites with different affinities, for which K_1 and K_2 are apparent dissociation constants. Assuming site “1” is the primary site near the promoter of the *lac* operon, its occupancy by LacR prevents RNA polymerase from binding to the promoter, blocking transcription of genes under its control. (B) Coupled equilibria involving different LacR-DNA complexes. Here $K_c^{(1)}$ and $K_c^{(2)}$ denote the unimolecular association constants associated with DNA looping and are related to K_1 , K_2 , λ , and J by Equation 9 (see Materials and Methods). doi:10.1371/journal.pone.0000136.g004

where $E_{no\ loop}$ and E_{loop} denote rates of gene expression in the absence of DNA looping (i.e., deletion of the auxiliary site indicated by site 2 in Figure 4), and that in its presence, respectively. In Equation 1 P_1 is the LacR-tetramer concentration in the cell and K_1 and K_2 are equilibrium dissociation constants of LacR for the primary and auxiliary operator sites, respectively. The dimensionless parameter λ mainly accounts for possible allosteric effects when one LacR tetramer associates with two DNA sites, with $\lambda > 1$ for cooperative binding and $\lambda < 1$ for anti-cooperative binding (see Equation 9 in Materials and Methods). We chose $\lambda = 1$ throughout because formation of the bidentate LacR-operator complex is non-cooperative [44]. The factor Γ contains all information concerning protein-DNA association exclusive of the looping contribution. In the special case of strong operator sites, i.e., $K_1, K_2 \ll P_1$, the enhanced repression in Equation 1 can be simplified to $R = 1 + \Gamma/P_1$. This expression confirms the notion that the role of DNA looping is to increase the local protein concentration, thereby enhancing gene repression. It also shows that the enhancement increases with decreasing protein concentration, a conclusion discussed in greater detail below.

In some experiments [5], repression was determined from the ratio of β -galactosidase activity measured for *E. coli* strains lacking a plasmid-borne LacR expression system to that for strains carrying the expression plasmid. Throughout our data analysis we adopt the definition of enhanced repression given in Equation (1), which is more appropriate for characterizing effects of DNA looping. This enhanced repression is the ratio of measured reporter activities for a construct in which the auxiliary operator has been deleted to that for a construct containing both primary and auxiliary operators. Therefore, to calculate *R*, the repression values of Müller *et al.* [5] were normalized relative to the measured value for a primary operator-only construct (120, see Figure 2 in [5]) under identical conditions. The resulting *R* values were then subjected to a multi-parameter curve-fitting analysis described below.

Analysis of *in-vivo* LacR-mediated gene repression based on DNA looping

Müller *et al.* [5] demonstrated a dramatic dependence of repression on helical phasing in systematic measurements of LacR-dependent gene repression at incremental operator spacings between 57.5 bp and 155.5 bp (the non-integral value is due to a 1-bp length difference between operator sequences). To fit these data using a non-linear least-squares method, we chose four adjustable parameters in our model: DNA helical repeat, DNA persistence length (or bending flexibility), DNA torsional rigidity (or twisting flexibility), and protein flexibility. All four parameters implicitly determine the value of the *J* factor in Equation 1. Protein-flexibility parameters (standard deviation values corresponding to angle fluctuations) for both protein-DNA and protein-protein contacts (Figure 1) assumed identical values and Γ (Table 1) was computed using reported values of P_1 , K_1 and K_2 [45,46] according to Equation 1. Note that a single value of the factor Γ applies to all values of the operator spacing in this model. Optimization over the four adjustable parameters was carried out using a simplex algorithm minimizing the following target function

$$\chi^2 = \sum_{i=1}^{N_d} \left[\frac{\log R_{comp,i} - \log R_{exp,i}}{\delta \log R_{exp,i} / R_{exp,i}} \right]^2 \quad (2)$$

where $R_{comp,i}$, $R_{exp,i}$ denote the computed and experimental enhanced-repression values, respectively. To avoid overfitting to

Table 1. Best-fit Values of Adjustable Parameters for LacR-loop-mediated Repression Data.

Data Set	No. of data points, N_d	Γ^c , $\times 10^2$	Fitting error ^d	Best-fit values			
				Persistence length, bp	Torsional rigidity, 10^{-19} erg cm	Helical repeat, bp/turn ⁻¹	Protein flexibility, deg.
Müller <i>et al.</i> [5]	51	1.17 ^a	1.0	95 (± 1)	1.1 (± 0.1)	11.60 (± 0.01)	20.7 (± 0.5)
Becker <i>et al.</i> [18] WT	26	4.66 ^b	1.1	95 (± 3)	0.7 (± 0.1)	11.08 (± 0.04)	19 (± 1)
Becker <i>et al.</i> [18] Δ HU	25	4.66 ^b	1.0	128 (± 2)	0.8 (± 0.1)	10.95 (± 0.03)	16 (± 1)

^{a,b}An identical auxiliary-operator sequence, O_{id} , was used in the experiments of references [5] and [18] with LacR-dissociation constant $K_2 = 0.036$ nM. However, the two sets of experiments differ significantly with respect to primary operators and *in-vivo* LacR concentration. In [5]: $K_1 = 0.36$ nM (O_1) and $P_1 = 85$ nM, whereas in [18]^P, $K_1 = 4.4$ nM (O_2) and $P_1 = 17$ nM. The five-fold difference in LacR levels between Müller *et al.*'s experiments and those of Becker *et al.* are due to the former's use of a LacR expression system that has a stronger promoter than that of the wild-type LacR gene. Additional information about operator-dependent dissociation constants can be found in Table 2.

^c $\Gamma = \frac{\lambda P_1}{(K_1 + P_1)(K_2 + P_1)}$, with $\lambda = 1$ for LacR.

^dThe fitting error is calculated according to $1/N_d \left\{ \sum_{i=1}^{N_d} [(\log R_{comp,i} - \log R_{exp,i}) / \delta \log R_{exp,i}]^2 \right\}^{1/2}$ where $R_{comp,i}$, $R_{exp,i}$ denote the computed and experimental enhanced-repression values, respectively, and $\delta \log R_{exp,i} = \delta_i / R_{exp,i}$ with the reported experimental error for the i -th data point.

doi:10.1371/journal.pone.0000136.t001

low experimental repression values (Figure 5), we chose the weight $\delta \log R_{exp,i} / R_{exp,i} = \delta_i / (R_{exp,i})^2$ in the least-squares fit, with δ_i the reported experimental error for the i -th data point. The total

number of experimental data points, N_d , was equal to 51. We obtained the fit to the experimental data shown in Figure 5A with the corresponding best-fit adjustable parameters given in Table 1; experimental and fitted enhanced-repression values as well as computed J factors for all of the analyses described can be found in Tables S1 and S2.

As shown in Table 1, values of the persistence length and torsional rigidity are respectively reduced by about 37% and more than 50% relative to their corresponding canonical values *in vitro* [47,48]. The fitted value of the DNA helical repeat, 11.60 (± 0.01) bp/turn, is consistent with previously reported *in-vivo* values [6] and is larger than that for topologically unconstrained DNA free in solution (≈ 10.5 bp turn⁻¹) because of DNA unwinding that accompanies negative supercoiling *in vivo* [49]. Our quantitative analysis of these gene-repression data first shows that the high degree of protein flexibility (20.7 $\pm 0.5^\circ$) cannot completely compensate the requirement for increased DNA flexibility *in vivo*. The high overall flexibility of the nucleoprotein assembly is reflected in the decay of repression peaks with operator spacings above 70 bp. This entropy-dominated effect in DNA looping is a unique feature of statistical-mechanical models [22], which take full account of DNA and protein flexibilities.

The strong agreement between calculated repression values and experimental data quantitatively verifies the role of DNA looping in *lac* gene regulation. In particular, our model explains the optimal O_1 – O_3 separation of 92 bp observed in the wild-type operon, which coincides with a strong peak in the enhanced-repression curve (Figure 5A). Once all of the adjustable parameters in the model were determined from the fitting procedure, we were able to compare predicted repression values with additional experimental measurements. Müller *et al.* also investigated the effects of operator quality at fixed operator spacing (~ 92 bp) [5]. In this experiment, the high-affinity auxiliary operator (O_{id}) used in the previous analysis was replaced by three different operator sites with weaker affinities for LacR. Table 2 shows that measured repression values and those calculated according to Equation 1 also give very good agreement.

Recent studies suggest that additional factors responsible for enhanced DNA flexibility, such as HU protein, may play an important role in facilitating loop-mediated gene regulation [19]. We investigated the role of HU in modulating DNA looping by analyzing the comparative LacR-dependent repression data of Becker *et al.* for *E. coli* strains expressing wild-type levels of HU protein and a mutant lacking HU [19]. Our analysis of these data

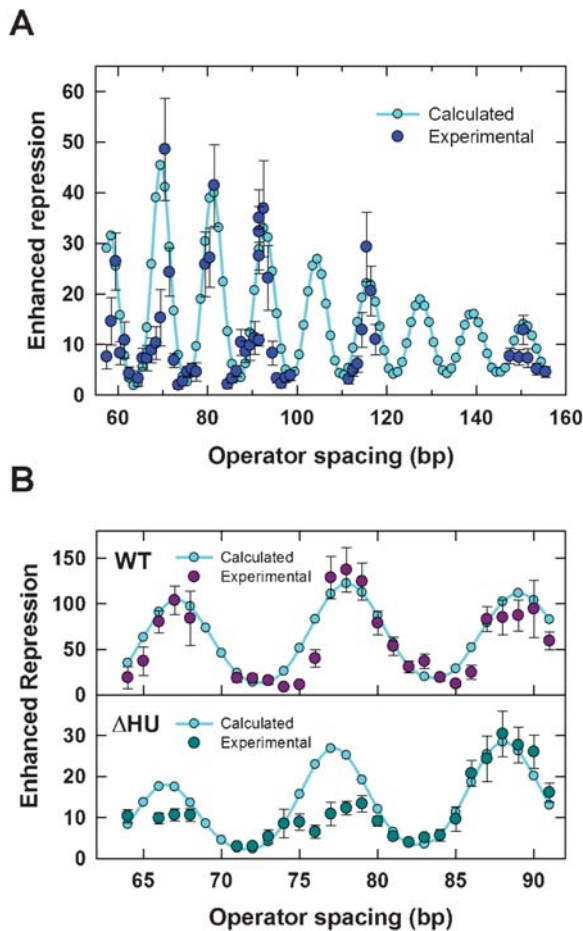


Figure 5. Best fit of the thermodynamic model to *in-vivo* repression data. Best-fit values of the adjustable parameters for all data sets are given in Table 1. (A) *In-vivo* repression data from Figure 3A of Müller *et al.* [5] and optimized fit to Equation 1. (B) *In-vivo* repression data of Becker *et al.* [19] for wild-type (“WT”) and HU-deletion (Δ HU) *E. coli* strains along with their respective optimized fits. doi:10.1371/journal.pone.0000136.g005

Table 2. Effect of operator quality on DNA looping and gene repression at approximately constant helical phasing.

Auxiliary operator	LacR dissociation constant, K_2 (nM) ^a	Measured repression	Calculated repression
O _{id}	0.036	32.3	28.8
O ₁	0.36	35.0	31.7
O ₂	4.4	27.5	30.4
O ₃	167	10.8	11.4

Operator nucleotide sequences and measured values of enhanced repression can be found in [5]. All measurements and calculations were performed using O₁ as the primary operator. Spacing between the primary and auxiliary operators was fixed at 92 bp, except for O_{id}. Because the O_{id} operator is 1 bp shorter (20 bp) than the other operator sequences, spacing between the primary and auxiliary operators, measured center-to-center, is 91.5 bp instead of 92 bp. The J factor is 2.37 μ M for the loop involving O_{id}, and 2.64 μ M for all other operators.

^aAuxiliary-operator dissociation constants, K_2 , were derived from binding free energies, ΔG , given by Vilar and Leibler [45] through the relationship $K_2 = \exp(\Delta G)/(N_A V)$ where N_A is Avogadro's number, and $V = 1$ fL ($1 \cdot 10^{-15}$ L) is the volume of an *E. coli* cell.

doi:10.1371/journal.pone.0000136.t002

sets used a value of Γ that was adjusted to reflect differences in operator affinities and LacR concentration relative to the Müller *et al.* experiments (Table 1). The resulting fit to the wild-type data in Figure 5B (upper panel) shows excellent agreement between the theory and experimental data and gives a low value for the persistence length, identical to that obtained in the fit to the Müller *et al.* data (Table 1). Compared with the Müller *et al.* results, Becker *et al.*'s enhanced-repression values are significantly greater (about three-fold as judged by the repression peaks), whereas the amplitude of the helical dependence is about two-fold smaller. The larger value of Γ in Becker *et al.*'s experiments, due to five-fold lower LacR expression (see Table 1, also Equation 1), is mainly responsible for the increased enhancement, whereas the diminished amplitude results from lower DNA torsional rigidity, which is decreased even further relative to its *in-vitro* value compared with the Müller *et al.* results. The DNA helical-repeat value (11.1 bp turn⁻¹) is consistent with lower levels of negative supercoiling than that in the earlier study [50]. Differences in helical-repeat values between the two data sets are not surprising, given the complex dependence of supercoiling on cellular physiology [51,52] and differences in the *E. coli* strains and DNA constructs used [5,19]. Specifically, in Müller *et al.*'s experiments DNA-looping assays involved operator sequences that were located on the *E. coli* chromosome, whereas in Becker *et al.*'s experiments, operator sequences resided on single-copy F' episomes. The DNA sequences between the operator sites were substantially similar in the two studies; none of the intervening DNA sequences contained any known intrinsic bends or other unusual features.

Using the same value of Γ estimated for the wild-type case, we then fitted Becker *et al.*'s data for repression in an HU deletion strain (Figure 5B, bottom panel). There was a marked increase in the best-fit DNA bending rigidity (Table 1), bringing this value into a range compatible with DNA molecules at moderate ionic strength *in vitro* [47,48]. Although the Δ HU persistence-length value (128 \pm 2 bp) is somewhat smaller than that normally given for mixed-sequence DNA in solution, it is equal within experimental uncertainty to values measured by rotational diffusion experiments at high salt (129 bp in 110 mM Na⁺/10 mM Mg²⁺) [53]. The abundance of multivalent cations and polyamines *in vivo* is expected to have significant effects on DNA elasticity [54];

however, it is also possible that non-specific binding of other architectural DNA-bending proteins present in the cell or sequence-dependent variations in bending flexibility in the region between operator sites may contribute to the slightly reduced persistence length.

Our torsional-rigidity values in the presence and absence of HU compare favorably with those estimated by Becker *et al.* using an empirical formula that contained torsional elasticity only. The model described here takes LacR structure and both bending and torsional flexibility of the entire nucleoprotein assembly into account and thus provides rigorous and quantitative evidence for a direct functional role of HU protein on DNA elasticity and loop-dependent interactions *in vivo*.

Biological consequences of DNA looping

One of the most frequently cited biological roles for DNA looping is to raise the local concentration of a regulatory protein in the vicinity of a promoter element [2,7]. Our rigorous analysis confirms this picture for LacR-mediated looping. As shown in Figure 6, DNA looping in HU-containing wild-type cells boosts the LacR concentration (J factor) at the primary operator (O₁) from its bulk value of 0.017 μ M to between 0.28 and 2.6 μ M. This effect raises the occupancy of the primary operator, the fraction of primary operator sequences bound by LacR, from 0.79 to between 0.985 and 0.998, a value essentially insensitive to helical phasing (Figure 7, upper panel). Such pronounced enhancement of operator occupancy has the consequence of decreasing the expression rate of β -galactosidase (molecules per hour per cell) [5,55] from 1,300 to a range of 12 to 90. In wild-type *E. coli* strains with an O₁-O₃ operator pair separated by 92 bp (Table 2), the predicted O₁ occupancy is 0.9986 (equivalent to about 8 β -galactosidase molecules per hour per cell), in excellent agreement with direct *in-vivo* measurements [3,56].

For a two-operator system, occupancy of the primary operator (Figure 7, upper panel) involves a looped state and two unlooped states (Figure 4). To relate the enhanced operator occupancy and

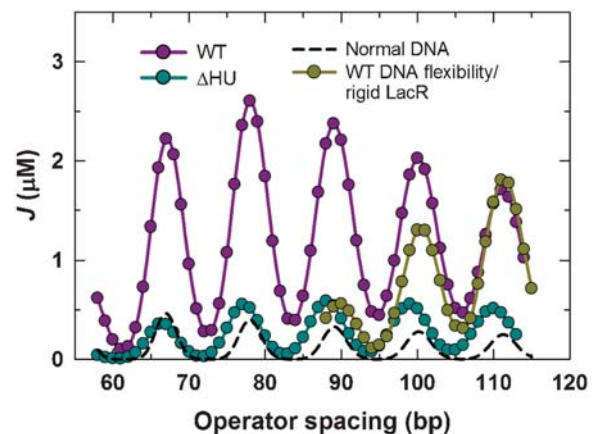


Figure 6. J factor versus operator spacing under different intracellular conditions. The J values corresponding to wild-type, WT, or HU-deletion, Δ HU, *E. coli* strains were calculated using the respective best-fit parameters given in Table 1. J factors corresponding to LacR-mediated loops having normal DNA elasticity were computed using canonical parameters for DNA persistence length (150 bp) and torsional rigidity ($2.4 \cdot 10^{-19}$ erg cm), and a LacR flexibility parameter identical to that for the wild-type strain (19°). Calculations of the J factor for the case of a rigid extended LacR conformation used the same parameters as for the wild-type strain except for a protein-flexibility parameter set equal to 2.0°.

doi:10.1371/journal.pone.0000136.g006

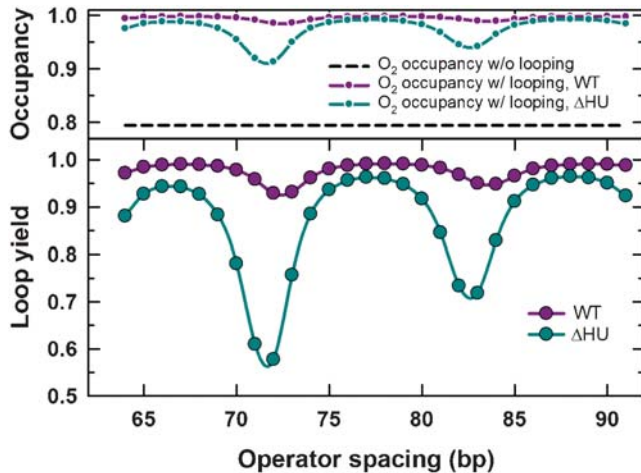
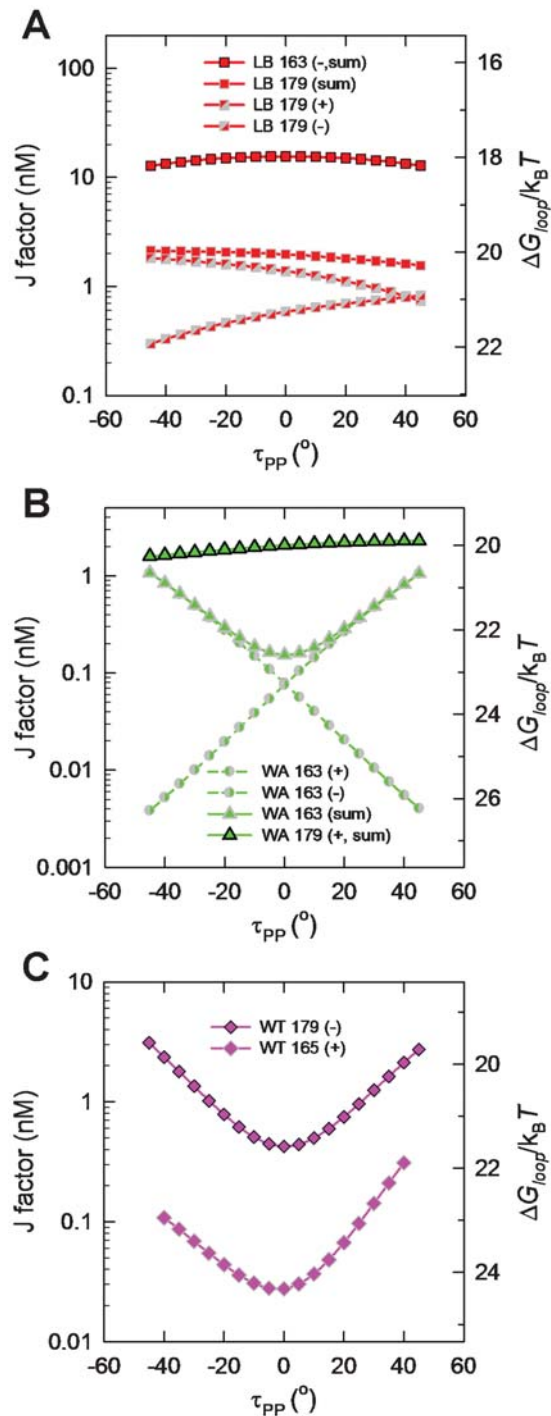


Figure 7. Predicted repressor occupancy of the primary operator and loop yield as a function of operator spacing corresponding to the data of Becker *et al.* [19]. Upper panel: the proportion of the primary operator site (O₂) bound by LacR for wild-type and HU-deletion strains computed from their respective best-fit parameters using formula $1 - E_{loop}$ (see Equation 13). The occupancy of the primary operator in the absence of the looping contribution is shown for reference (dashed line). Lower panel: the loop yield computed using Equation 10 for both wild-type and HU-deletion strains. *J* values used in these calculations are those given in Figure 6.
doi:10.1371/journal.pone.0000136.g007

gene repression to DNA looping, we calculated the loop yield (Figure 7, bottom panel), which is the proportion of looped states relative to all possible states. The loop yield directly correlates with the *J* factor, operator occupancy, and enhanced gene repression as demonstrated by their identical dependence on DNA helical phase (Figures 5B, 6, and 7). Furthermore, the high loop yield (0.929–0.992) confirms that enhanced gene repression is almost exclusively attributable to DNA looping.

In the absence of HU, increases in effective DNA bending and protein rigidities reduce measured gene repression by up to twelve-fold, depending on operator spacing, with an average reduction of 5.6-fold (Figure 5B). The effect of HU is also apparent from decreases in *J* factor (Figure 6), operator occupancy, and loop yield (Figure 7) for the HU-deficient *E. coli* strain compared with wild-type. To put this finding in perspective, we calculated *J*-factor values using a canonical DNA persistence length of 150 bp (Figure 6). As expected, *J* values are much smaller than those obtained *in vivo* in the presence of HU [5,19]. These comparisons quantitatively confirm HU's putative role in facilitating the formation of small DNA loops *in vivo*. DNA torsional rigidity *in vivo* is substantially reduced relative to the *in-vitro* value and is not significantly affected by the presence of HU protein (Table 1). The basis of HU's differential effect on bending and torsional rigidities is not clear. We speculate that DNA supercoiling *in vivo* may enhance nonlinearities in DNA torsional elasticity [11,57] and that this effect is largely independent of HU expression.

Dynamic DNA bending induced by HU protein is not the only factor that reduces the thermodynamic cost of small-loop formation, however. Significant helical-phase-dependent enhanced repression remains even when HU is deleted (Figure 5B). Consistent with this observation, predicted values of operator occupancy (Figure 7, upper panel) are significantly greater than that in the absence of DNA looping and the major proportion of DNA is expected to be in the looped state (Figure 7, bottom panel).



doi:10.1371/journal.pone.0000136.g008

While investigating other contributions to small-loop formation, we noticed that the effective conformational flexibility of LacR is only marginally reduced (from 19° to 16°) in the absence of HU (Table 1). This slight change in the protein-flexibility parameter is probably caused by differing extents of LacR deformation in forming DNA loops and accompanying nonlinearity of protein elasticity.

We assessed the specific contribution of protein flexibility to loop formation by comparing the *J* factors calculated for a flexible

protein assembly with those of a rigid extended LacR structure (Figure 6), assuming the same DNA rigidities as those in the HU-containing wild-type strain. Interestingly, we could obtain solutions only in cases where the operator spacing was greater than 88 bp, indicating that DNA loops smaller than this size have highly unfavorable looping free energies. In the range from 88 to 100 bp, J factors with rigid LacR tetramers are lower than those with flexible tetramer assemblies. This comparison demonstrates the crucial role of protein flexibility in forming small DNA loops. In contrast, above ~ 108 bp, high protein flexibility makes DNA looping unfavorable due to increased entropy loss [22]. Taken together, we conclude from these results that protein structure, protein conformational flexibility, and DNA flexibility induced by non-specific protein-DNA interactions such as those with HU, all contribute significantly to the formation of small DNA loops widely observed *in vivo*.

DISCUSSION

Gene repression in the *lac* system has become a textbook example of how DNA looping modulates the local concentration of a regulatory protein in the vicinity of a promoter. Intensive study over the last three decades has led to a wealth of information about the thermodynamics of *lac* repressor's interaction with single wild-type and mutant operator sequences and the dependence of gene repression on operator spacing. However, the quantitative effect of DNA looping on LacR-mediated gene regulation has been highly controversial. Here we provide a novel analysis of *in-vivo* gene repression from first principles based on a rigorous statistical-mechanical model of DNA looping.

We analyzed two independent data sets that characterize the dependence of LacR-mediated repression on inter-operator spacing: those of Müller *et al.* [5], and the data of Becker *et al.* [19]. Both sets of experiments systematically cover overlapping ranges of operator-site spacing that span at least two full helical turns. Moreover, both studies were carried out using constructs in which the CAP-binding site located near the promoter was abolished, eliminating the need to take possible CAP-dependent DNA bending or CAP-LacR interactions into account in the calculation of loop free energy [58,59]. We chose on this basis not to analyze the classic results of Law *et al.* [4] because those experiments were done with regulatory regions that included an intact CAP-binding site.

The excellent agreement between experiments and this analysis validates the dominance of the extended LacR structure in DNA looping *in vivo* over the v-shaped LacR structure widely observed *in vitro*. This conclusion is largely consistent with analyses of both the Müller *et al.* and Becker *et al.* results by Saiz *et al.* [24,60], in which a single LacR conformation was found to populate $\sim 80\%$ of the DNA loops formed. The remaining proportion was proposed to be in an alternative conformation, although the details of these two distinct conformations could not be determined from their analysis. The small discrepancy with the present work regarding the existence of this alternative conformation *in vivo* may be due to the simplified semi-quantitative model used by Saiz *et al.*, in which DNA bending flexibility and LacR structure and conformational flexibility were not included. Although a small contribution to looping from the "v-shaped" LacR conformation cannot be ruled out, the extended conformation alone is sufficient to account for all of the *in-vivo* repression data. In the case of the Müller *et al.* data set, our computed repression peaks are somewhat broader than the corresponding experimental ones (see Figure 5A). For unknown reasons, the difference in theoretical and experimental peak widths is a particular feature of the Müller *et al.* data set; there

is no obvious broadening of the computed repression peaks relative to those in the Becker *et al.* data (Figure 5B).

Several studies have investigated the specific role of DNA architectural proteins such as HU and HMG in enhancing apparent DNA bending and twisting flexibilities *in vitro* and *in vivo* [14–17,19,61,62]. HU is an abundant protein in *E. coli*, present at levels of 60,000 copies per wild-type cell [63], or equivalently, about one HU dimer per 100 bp DNA. *In vitro*, DNA fragments as short as 100 bp can be readily circularized by DNA ligase in the presence of HU [15], albeit at HU:DNA ratios that significantly exceed the *in-vivo* value. Nevertheless, abolishing HU protein in a deletion strain dramatically reduces gene repression mediated by LacR (Figure 5B) and directly correlates with decreased apparent DNA bending flexibility. Our results therefore suggest that HU plays an important role in regulating DNA bendability *in vivo*.

Because DNA looping is an essential mechanistic feature of many biological processes including transcription, DNA replication, recombination, and repair [64], the principles that govern DNA looping in the *lac* system generally apply to a very large class of problems in biological regulation and function. The quantitative approach described here should open the way to rigorous *in-vitro* and *in-vivo* characterization of this biologically important class of regulatory mechanisms. An even more comprehensive picture of DNA looping will emerge once a more complete understanding of the kinetics of loop formation is attained [20,55,65].

MATERIALS AND METHODS

Statistical-Mechanical Theory of DNA Looping and Computational Methods

Details of the theory have been published elsewhere [22,29]; thus, only a summary of salient features is presented here. We simplify the structure of a protein-mediated DNA loop by treating the nucleoprotein assembly as a connected chain of rigid bodies. The Hamiltonian for a free chain in the absence of constraints is

$$\beta H = \sum_{i=1}^{N-1} \sum_{j=1}^3 \frac{(x_{ij} - \bar{x}_{ij})^2}{\sigma_{ij}^2} \quad (3)$$

where $X_{ij}(j = 1, \dots, 3)$ denotes the instantaneous rotation angle (tilt, roll, or twist) of the i -th rigid body relative to the $(i-1)$ -st one in the presence of thermal fluctuations characterized by σ_{ij} , and \bar{x}_{ij} is the corresponding equilibrium angle. Here N is the total number of rigid bodies in the chain and $\beta = 1/k_B T$. The Hamiltonian for a closed loop is also described by Equation (3), but subject to six constraints due to chain closure [29], i.e.,

$$f^{(k)}(\{x_{ij} : i = 1, \dots, N-1; j = 1, \dots, 3\}) = 0; \quad k = 1, \dots, 6 \quad (4)$$

which are nonlinear functions of the angular parameters. After finding the mechanical equilibrium conformation of the closed loop with minimum elastic energy, the J factor is calculated using the formula [29]

$$J = \frac{8\pi^2 e^{-E_s}}{\sqrt{\pi^m \det(A) \det(F)}} \quad (5)$$

where E_s is the mechanical elastic energy of the loop and $m = 6$. Here A and F are matrices with dimensions $(3N-3) \times (3N-3)$ and 6×6 , respectively, whose elements are functions of the thermal fluctuations σ_{ij} and the first and second derivatives of the constraint functions (left side of Equation 4) with respect to angular parameters at the mechanical-equilibrium conformation.

Unless noted otherwise, all calculations used canonical parameters for duplex DNA: helical twist $\tau_0 = 34.45^\circ$, a sequence-independent twist-angle standard deviation, or twisting flexibility, $\sigma_\tau = 4.388^\circ$, and standard deviations, or bending flexibilities, for all tilt (θ) and roll (ϕ) angles, σ_θ and σ_ϕ , respectively, of 4.678° (equivalent to an isotropically flexible DNA chain with a persistence length of 150 bp). Average values of tilt and roll for DNA were taken to be zero. To model the DNA loops mediated by the v-shaped protein conformation, we used the following angular parameters for the mechanical-equilibrium conformations of the LacR tetramer [22]: $\phi_{DP} = \phi_{PD} = 67.5^\circ$, $\phi_{PP} = 120^\circ$ for “WA;” $\phi_{DP} = \phi_{PD} = 67.5^\circ$, $\phi_{PP} = -120^\circ$ for “WT;” and $\phi_{DP} = -67.5^\circ$, $\phi_{PD} = 67.5^\circ$, $\phi_{PP} = 120^\circ$, $\tau_{DP} = 180^\circ + 34.45^\circ = 214.45^\circ$ for “LB” [34,35]. For the extended LacR conformation, $\phi_{DP} = \phi_{PD} = 67.5^\circ$, $\phi_{PP} = \tau_{PP} = 0$ [35]. The subscripts specify angular-parameter values for contacts between the protein and the last (DP) and first (PD) base pairs of the DNA loop or between the two protein domains (PP). Note that these parameters take protein-induced DNA bending ($\approx 45^\circ$) at the operator sites into account. Slightly different values for the length of the major LacR-dimer axis were used for the v-shaped and extended LacR conformations: 20 bp and 25 bp, respectively [35]. J-factor computations were carried out on a 2.8-GHz Pentium-4 CPU with 2 GBytes RAM. Geometries of optimized LacR-DNA looped conformations were visualized using the OpenDX data visualization package (<http://www.opendx.org/>). Fortran-90 and C-language source code is available upon request.

Analysis of *in-vivo* LacR-mediated repression

We assume that protein-protein association is maintained under all conditions and that all specific protein-DNA interactions are accounted for in the equilibria shown in Figure 4. Intermolecular DNA associations have also been observed in some *in-vitro* experiments [36,66] because of high DNA concentrations that were used. *In vivo*, these bimolecular associations are unlikely and are therefore not shown in Figure 4.

Take p_0 to be the concentration of free LacR tetramer, and let d_0 , d_1 , d_2 , d_{12} , and d_c designate the concentrations of free DNA, DNA with site “1” bound, DNA with site “2” bound, doubly-bound (unlooped) DNA, and the closed loop, respectively. Then the following equations hold

$$K_1 = \frac{p_0 d_0}{d_1} = \frac{p_0 d_2}{d_{12}}, K_2 = \frac{p_0 d_0}{d_2} = \frac{p_0 d_1}{d_{12}}, K_c^{(1)} = \frac{d_c}{d_1}, K_c^{(2)} = \frac{d_c}{d_2} \quad (6)$$

with

$$p_0 + d_1 + d_2 + 2d_{12} + d_c = P_t \quad (7)$$

and

$$d_0 + d_1 + d_2 + d_{12} + d_c = D_t \quad (8)$$

where D_t and P_t are the total DNA and protein concentrations, respectively. Note that K_1 and K_2 are dissociation constants, whereas the loop-closure constants K_c describe an association reaction. By analogy with its definition in DNA cyclization [25,27], the J factor is the ratio of unimolecular association constant to that of bimolecular association, i.e.,

$$J = \frac{K_c^{(1)}}{\lambda/K_2} = \frac{K_c^{(2)}}{\lambda/K_1} \quad (9)$$

Here we include λ as a dimensionless factor that accounts for possible changes in affinity that could accompany three effects: (1)

allosteric binding when one LacR tetramer associates with two DNA sites, (2) nonspecific DNA binding of LacR, and (3) minor decreases in association constant for a DNA-bound LacR molecule compared to free LacR that arise from greater translational and rotational entropy loss of the former in protein-DNA complexes [67].

Although an exact solution to the above system of equations is available and equivalent to solving a cubic equation, here we consider only the special case where $D_t \ll P_t$, which corresponds to most *in vivo* conditions. In this case Equation (7) gives $p_0 \approx P_t$. The other variables can be obtained by replacing p_0 with P_t , expressing d_0, d_1, d_{12} , and d_c in terms of d_0 , and solving for d_0 with Equation (8). Specifically,

$$\frac{d_c}{D_t} = \frac{\lambda J P_t}{K_1 K_2 + (K_1 + K_2 + \lambda J + P_t) P_t} \quad (10)$$

In the case where $\lambda J \gg K_1$, K_2 and $\lambda J \gg P_t$, the looped configuration dominates and protein-DNA association can be approximated in terms of a two-state equilibrium, yielding

$$\frac{d_c}{D_t} \approx \frac{K_a P_t}{1 + K_a P_t} \quad (11)$$

with an apparent association constant

$$K_a = \frac{\lambda J}{K_1 K_2} \quad (12)$$

Equation (12) relates the total binding strength of a protein modeled as two DNA-binding domains connected by a flexible or semiflexible linker to the binding strengths of individual domains and the geometrical and mechanical properties of the linker. Although not formulated with J factors, a similar model was used by Crothers and Metzger to investigate the thermodynamic linkage between monomeric antibody binding strengths and the overall association constants of multivalent antibodies [68]. This model has been revisited in a study of protein-DNA interactions involving proteins with two domains connected by flexible linkers [69]. Here we have derived the general case from the standpoint of DNA looping and extended this approach to cases where the linkers are semi-flexible. The formula has recently been applied to quantitate the role of sequence-dependent DNA bending and flexibility in E2-DNA interactions using a worm-like chain model of DNA [31].

For the *lac* operon, we designate as “1” the primary site near the promoter, and “2” the auxiliary site located upstream of the promoter [5]. Based on the assumption that gene transcription is under thermodynamic control [4], the transcription rate of a reporter gene (the gene for β -galactosidase in this case) under control of the promoter is proportional to the probability that site “1” is free. Consequently, the gene transcription rate when DNA looping takes place is directly proportional to

$$E_{loop} = \frac{d_0 + d_2}{D_t} = \frac{K_1 (K_2 + P_t)}{K_1 K_2 + (K_1 + K_2 + \lambda J + P_t) P_t} \quad (13)$$

When $P_t \gg K_2$, Equation (13) can be simplified to

$$E_{loop} \approx \frac{1}{1 + (\lambda J + P_t)/K_1} \quad (14)$$

which was previously obtained by Law et al [4], assuming 100% occupancy of site “2”. However, the relationship $P_t \gg K_2$ does not hold in general and thus Equation (13) must be used. Similarly, the

rate without DNA looping, which is determined in the absence of site “2”, is proportional to

$$E_{noloop} = \frac{K_1}{K_1 + P_t} \quad (15)$$

The enhanced gene repression due to DNA looping, R , can be expressed as the ratio of the specific enzymatic activity of β -galactosidase in the absence of the site “2” to that in its presence. Then the calculated enhanced gene repression, the ratio of transcription rates in the absence and presence of the loop, can be compared with the experimental R values through the relation shown in Equation 1.

SUPPORTING INFORMATION

Supporting Text Supporting information for “Analysis of *In-Vivo* LacR-Mediated Gene Repression Based on the Mechanics of DNA Looping” by Yongli Zhang, Abbye E. McEwen, Donald M. Crothers, and Stephen D. Levene.
Found at: doi:10.1371/journal.pone.0000136.s001 (0.31 MB DOC)

REFERENCES

- Berg JM, Tymoczko JL, Stryer L (2002) Biochemistry. New York: W. H. Freeman and Company.
- Mossing MC, Record MT (1986) Upstream operators enhance repression of the *lac* promoter. *Science* 233: 889–892.
- Sasse-Dwight S, Gralla JD (1988) Probing co-operative DNA-binding *in vivo*. The *lac* O1:O3 interaction. *J Mol Biol* 202: 107–119.
- Law SM, Bellomy GR, Schlach PJ, Record MT (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.
- Müller J, Oehler S, Müller-Hill B (1996) Repression of *lac* promoter as a function of distance, phase and quality of an auxiliary *lac* operator. *J Mol Biol* 257: 21–29.
- Lee DH, Schleif RF (1989) *In-vivo* DNA loops in AraCBAD - size limits and helical repeat. *Proc Natl Acad Sci U S A* 86: 476–480.
- Schleif R (1992) DNA looping. *Annu Rev Biochem* 61: 199–223.
- Matthews KS (1992) DNA looping. *Microbiol Rev* 56: 123–136.
- Cloutier TE, Widom J (2004) Spontaneous sharp bending of double-stranded DNA. *Mol Cell* 14: 355–362.
- Yan J, Marko JF (2004) Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys Rev Lett* 93.
- Cloutier TE, Widom J (2005) DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc Natl Acad Sci U S A* 102: 3645–3650.
- Wiggins PA, Phillips R, Nelson PC (2005) Exact theory of kinkable elastic polymers. *Phys Rev E Stat Nonlin Soft Matter Phys* 71: 021909.
- Wiggins PA, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, et al. (2006) High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat Nano* 1: 137.
- Flashner Y, Gralla JD (1988) DNA dynamic flexibility and protein recognition - differential stimulation by bacterial histone-like protein HU. *Cell* 54: 713–721.
- Hodges-Garcia Y, Hagerman PJ, Pettijohn DE (1989) DNA ring closure mediated by protein HU. *J Biol Chem* 264: 14621–14623.
- Ross ED, Hardwidge PR, Maher LJ (2001) HMG proteins and DNA flexibility in transcription activation. *Mol Cell Biol* 21: 6598–6605.
- Skoko D, Wong B, Johnson RC, Marko JF (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry* 43: 13867–13874.
- Yan J, Kawamura R, Marko JF (2005) Statistics of loop formation along double helix DNAs. *Physical Review E* 71.
- Becker NA, Kahn JD, Maher LJ (2005) Bacterial repression loops require enhanced DNA flexibility. *J Mol Biol* 349: 716–730.
- van den Broek B, Vanzi F, Normanno D, Pavone FS, Wuite GJL (2006) Real-time observation of DNA looping dynamics of type III restriction enzymes *NarI* and *NarI*. *Nucleic Acids Res* 34: 167–174.
- Gemmen GJ, Millin R, Smith DE (2006) DNA looping by two-site restriction endonucleases: heterogeneous probability distributions for loop size and unbinding force. *Nucleic Acids Res* 34: 2864–2877.
- Zhang Y, McEwen AE, Crothers DM, Levene SD (2006) Statistical-mechanical theory of DNA looping. *Biophys J* 90: 1903–1912.
- Rippe K (2001) Making contacts on a nucleic acid polymer. *Trends Biochem Sci* 26: 733–740.
- Saiz L, Vilar JMG (preprint) *In-vivo* evidence of alternative loop geometries in DNA-protein complexes. <http://arxiv.org/abs/q-bio.BM/0602012>.
- Shore D, Langowski J, Baldwin RL (1981) DNA flexibility studied by covalent closure of short fragments into circles. *Proc Natl Acad Sci U S A* 78: 4833–4837.
- Shore D, Baldwin RL (1983) Energetics of DNA twisting. I. relation between twist and cyclization probability. *J Mol Biol* 170: 957–981.
- Crothers DM, Drak J, Kahn JD, Levene SD (1992) DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol* 212: 3–29.
- Bloomfield VA, Crothers DM, Tinoco IJ (2000) *Nucleic Acids: Structures, Properties and Functions*. Herndon, VA: University Science Books. 800 p.
- Zhang YL, Crothers DM (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys J* 84: 136–153.
- Zhang YL, Crothers DM (2003) High-throughput approach for detection of DNA bending and flexibility based on cyclization. *Proc Natl Acad Sci U S A* 100: 3161–3166.
- Zhang YL, Xi ZQ, Hegde RS, Shakked Z, Crothers DM (2004) Predicting indirect readout effects in protein - DNA interactions. *Proc Natl Acad Sci U S A* 101: 8337–8341.
- Swigon D, Coleman BD, Olson WK (2006) Modeling the *lac* repressor-operator assembly: The influence of DNA looping on *lac* repressor conformation. *Proc Natl Acad Sci U S A* 103: 9879–9884.
- Friedman AM, Fischmann TO, Steitz TA (1995) Crystal structure of *lac* repressor core tetramer and its implications for DNA looping. *Science* 268: 1721–1727.
- Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, et al. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271: 1247–1254.
- Ruben GC, Roos TB (1997) Conformation of *lac* repressor tetramer in solution bound and unbound to operator DNA. *Microsc Res Tech* 36: 400–416.
- Mehta RA, Kahn JD (1999) Designed hyperstable *lac* repressor-DNA loop topologies suggest alternative loop geometries. *J Mol Biol* 294: 67–77.
- Edelman LM, Cheong R, Kahn JD (2003) Fluorescence resonance energy transfer over ~130 basepairs in hyperstable *lac* repressor-DNA loops. *Biophys J* 84: 1131–1145.
- Jacobson H, Stockmayer WH (1950) Intramolecular reaction in polycondensations. 1. The theory of linear systems. *J Chem Phys* 18: 1600–1606.
- Geanakopoulos M, Vasmatazis G, Zhurkin VB, Adhya S (2001) *Gal* repressosome contains an antiparallel DNA loop. *Nat Struct Biol* 8: 432–436.
- Lewis M (2005) The *lac* repressor. *Comptes Rendus Biologies* 328: 521–548.
- Villa E, Balaeff A, Schulten K (2005) Structural dynamics of the *lac* repressor-DNA complex revealed by a multiscale simulation. *Proc Natl Acad Sci U S A* 102: 6783–6788.
- Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, et al. (2005) Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev* 15: 116–124.
- Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, et al. (2005) Transcriptional regulation by the numbers: applications. *Curr Opin Genet Dev* 15: 125–135.

44. von Hippel PH (1979) On the molecular bases of the specificity of interaction of transcriptional proteins with genome DNA. In: Goldberger RF, ed. *Biological Regulation and Development*. New York: Plenum: 279–347.
45. Oehler S, Amouyal M, Kolkhof P, von Wilcken-Bergmann B, Müller-Hill B (1994) Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *EMBO J* 13: 3348–3355.
46. Vilar JMG, Leibler S (2003) DNA looping and physical constraints on transcription regulation. *J Mol Biol* 331: 981–989.
47. Hagerman PJ (1988) Flexibility of DNA. *Annu Rev Biophys Biophys Chem* 17: 265–286.
48. Levene SD (2001) DNA/RNA flexibility. *Nature Encyclopedia of Life Sciences* (Online): Wiley Interscience. <http://www.els.net>.
49. Bliska JB, Cozzarelli NR (1987) Use of site-specific recombination as a probe of DNA structure and metabolism *in vivo*. *J Mol Biol* 194: 205–218.
50. Wasserman SA, White JH, Cozzarelli NR (1988) The helical repeat of double-stranded DNA varies as a function of catenation and supercoiling. *Nature* 334: 448–450.
51. Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A* 84: 7024–7027.
52. Espeli O, Mariani KJ (2004) Untangling intracellular DNA topology. *Mol Microbiol* 52: 925–931.
53. Pörschke D (1991) Persistence length and bending dynamics of DNA from electrooptical measurements at high salt concentrations. *Biophys Chem* 40: 169–179.
54. Pastré D, Piétrement O, Landousy F, Hamon L, Sorel I, et al. (2006) A new approach to DNA bending by polyamines and its implication in DNA condensation. *Eur Biophys J* 35: 214–223.
55. Vilar JMG, Saiz L (2005) DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Curr Opin Genet Dev* 15: 136–144.
56. Yu J, Xiao J, Ren X, Lao K, Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* 311: 1600–1603.
57. Bryant Z, Stone MD, Gore J, Smith SB, Cozzarelli NR, et al. (2003) Structural transitions and elasticity from torque measurements on DNA. *Nature* 424: 338–341.
58. Vossen KM, Stickle DF, Fried MG (1996) The mechanism of CAP-*lac* repressor binding cooperativity at the *E. coli* lactose promoter. *J Mol Biol* 255: 44–54.
59. Fried MG, Daugherty MA (2001) *In vitro* interaction of the *Escherichia coli* cyclic AMP receptor protein with the lactose repressor. *J Biol Chem* 276: 11226–11229.
60. Saiz L, Rubi JM, Vilar JMG (2005) Inferring the *in vivo* looping properties of DNA. *Proc Natl Acad Sci U S A* 102: 17642–17645.
61. Haykinson MJ, Johnson RC (1993) DNA looping and the helical repeat *in vitro* and *in vivo*: effect of HU protein and enhancer location on Hin invertasome assembly. *EMBO J* 12: 2503–2512.
62. van Noort J, Verbrugge S, Goosen N, Dekker C, Dame RT (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci U S A* 101: 6969–6974.
63. Azam TA, Hiraga S, Ishihama A (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5: 613–626.
64. Echols H (1986) Multiple DNA-protein interactions governing high-precision DNA transactions. *Science* 233: 1050–1056.
65. Finzi L, Gelles J (1995) Measurement of lactose repressor-mediated loop formation and breakdown in single DNA-molecules. *Science* 267: 378–380.
66. Watson MA, Gowers DM, Halford SE (2000) Alternative geometries of DNA looping: an analysis using the *SfiI* endonuclease. *J Mol Biol* 298: 461–475.
67. Hill TL (1987) *Linear Aggregation Theory in Cell Biology*. New York: Springer-Verlag.
68. Crothers DM, Metzger H (1972) Influence of polyvalency on binding properties of antibodies. *Immunochemistry* 9: 341–357.
69. Zhou HX (2001) The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins. *Biochemistry* 40: 15069–15073.