

University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

Physics and Astronomy Faculty Publications
and Presentations

College of Sciences

4-2013

The thermodynamics of DNA loop formation, from J to Z

Stephen D. Levene

Stefan M. Giovan

Andreas Hanke

Massa J. Shoura

Follow this and additional works at: https://scholarworks.utrgv.edu/pa_fac



Part of the [Astrophysics and Astronomy Commons](#), and the [Physics Commons](#)

The Thermodynamics of DNA-loop Formation, from *J* to *Z*

**Stephen D. Levene,^{1,2,3,*} Stefan M. Giovan,² Andreas Hanke,⁴ and
Massa J. Shoura^{1,2}**

Departments of ¹Bioengineering, ²Molecular and Cell Biology, and ³Physics, University of Texas at Dallas, Richardson, TX 75083 USA; ⁴Department of Physics and Astronomy, University of Texas at Brownsville, Brownsville, TX 78520 USA

*Corresponding author, E-mail: sdlevene@utdallas.edu

Tel.: 972 883 2503, Fax: 972 883 2409

Address reprint requests to S.D. Levene.

The formation of DNA loops is a ubiquitous theme in biological processes, including DNA replication, recombination, repair, and gene regulation. These loops are mediated by proteins bound at specific sites along the contour of a single DNA molecule, in some cases many thousands of base pairs apart. Loop formation incurs a thermodynamic cost that is a sensitive function of the length of looped DNA as well as the geometry and elastic properties of the DNA-bound protein. The free energy of DNA looping is logarithmically related to a generalization of the Jacobson-Stockmayer factor for DNA cyclization, termed the J factor. Here we review the thermodynamic origins of this quantity, discuss how it is measured experimentally, and connect the macroscopic interpretation of the J factor with a statistical-mechanical description of DNA looping and cyclization.

Keywords: DNA elasticity, J factor, cyclization kinetics, site-specific recombination, partition function

INTRODUCTION

Simultaneous binding of multiple protein factors to two or more DNA sequences is a common aspect of gene expression, genetic recombination, DNA replication, and DNA repair. The target sites involved in these interactions are frequently located at considerable distances along the same DNA molecule, which requires that the intervening DNA form a looped structure. Looped nucleoprotein assemblies, for example some gene-regulatory complexes in mammalian systems [1-3], can involve many proteins and also segments of DNA that are thousands of base pairs long. Such systems are challenging enough to characterize *in vitro*; having an accurate *in-vivo* picture of these processes requires also knowing how DNA flexibility and folding are affected by DNA-bound architectural proteins in prokaryotes [4,5] or chromatin organization [6] and the binding of non-histone proteins [7,8] in eukaryotes. Advances in our quantitative understanding of DNA-loop formation therefore depend on detailed consideration of the mechanics of loop formation that takes into account DNA elasticity, local and global aspects of DNA organization, and the plasticity of protein-DNA and protein-protein interactions.

Here we address the thermodynamic cost, in terms of free energy, of forming specific DNA-loop conformations. The overall free energy of assembling a particular protein-mediated loop amounts to a sum of free energies for protein-protein interactions, protein-DNA binding, and DNA distortion through wrapping, bending, and looping. The protein-protein and protein-DNA terms are typically regarded as independent and separable from that of DNA distortion, though these assumptions can clearly be questioned in the context of a chromatin environment. We focus in this article on the free energy cost of DNA looping in terms of the statistical distribution of polymer conformations. Mechanical constraints applied to DNA segments at the loop ends strongly perturb the polymer statistics and finding a general solution is a challenging problem in statistical mechanics. However, treating the looping problem as a generalization of the more extensively analyzed case of DNA cyclization has helped to illuminate general principles [9]. In particular, the DNA-size and helical-phase dependence of loop formation can be understood in terms of the size and conformation of the protein complex that mediates looping.

The dependence of the free energy of loop formation on DNA size contains non-trivial enthalpic and entropic contributions [9,10]. The behavior of looped domains involving DNA contour lengths smaller than the persistence length, P , (≈ 150 bp or 50 nm at moderate ionic strength) is dominated by the bending and twisting rigidity of the double helix and thus the free energy cost of loop formation is largely enthalpic. In contrast, loops that are much larger than P , incur minimal enthalpic cost, but are entropically unfavorable. Effects of DNA tertiary structure and topology on looping such as supercoiling or knotting of DNA domains are not always considered, but remain nonetheless important.

THE THERMODYNAMICS OF POLYMER CYCLIZATION

The problem of polymer cyclization was first considered by Jacobson and Stockmayer [11], who formulated the free-energy cost of cyclization in terms of a ratio of respective equilibrium constants for formation of a circular polymer chain from a linear monomer, K_c , and conversion of two linear monomers to a linear dimer, K_d . This ratio is commonly known as the Jacobson-Stockmayer factor, J , also called the J factor,

$$J = K_c / K_d = \exp \left[-(\Delta G_c^0 - \Delta G_b^0) / RT \right] \quad (1)$$

where ΔG_c^0 and ΔG_b^0 are the respective Gibbs free-energy changes in the standard state for the cyclization and dimerization reactions. Equivalently, J is the equilibrium constant for forming a

circular and a linear monomer from a linear dimer [12]. By including the free-energy change for the dimerization reaction in equation (1) the contribution of bond formation between terminal residues is subtracted from the overall cyclization free energy. Thus, J specifically quantifies the effect on the global free energy of the polymer of constraining the chain to a circular conformation.

For a random-flights polymer the orientations of successive chain segments are independent and the effect of the end constraints involves only the confinement of both ends of the chain to a common volume element δV . Then,

$$\begin{aligned}\Delta G_c^0 &= \Delta G^* - RT \ln[W(0)\delta V] \\ \Delta G_b^0 &= \Delta G^* - RT \ln[N_{Av}\delta V]\end{aligned}\tag{2}$$

where $W(0)$ is the probability density for the end-to-end distance $h = |\mathbf{h}|$ (see Figure 1) evaluated at $h = 0$, N_{Av} is Avogadro's number, and ΔG^* is the free-energy of bond formation between the first and last residues of the linear chain [13]. It then follows from equation (1) that

$$J = W(0)/N_{Av},\tag{3}$$

which connects the J factor to the statistical fraction of chain conformations having both chain ends within an infinitesimal distance of one another. J can therefore also be considered as the effective concentration of one chain end in the vicinity of the other.

In the case of a random-flights polymer, $W(h)$ has a closed-form solution for arbitrary h (including $h = 0$), which leads to a simple relation for J ,

$$J = \frac{[3 / (2\pi n\ell^2)]^{3/2}}{N_{Av}}\tag{4}$$

where n is the number of chain segments of length ℓ . This formula is equivalent to Jacobson and Stockmayer's original expression. Note that the volume element δV does not appear, having been eliminated in equation (2).

In the early 1980s DNA-cyclization measurements on restriction fragments about 200 to 1500 bp in size [14-16] motivated more sophisticated treatments of the ring-closure problem (Figure 1). Advanced methods were necessary in order to account for the effects of bending and twisting rigidity in DNA molecules on this length scale. Twisting rigidity in particular leads to a strong periodic dependence of J on the fractional number of helical turns. A perturbation method for homogeneous and isotropically flexible DNA circles was introduced by Shimada and Yamakawa [17] that successfully accounts for helical phasing, bending, and twisting rigidities in small circles. Subsequently, Monte Carlo methods became available to compute J from ensembles of simulated helical wormlike chains [12,18,19]. An advantage of simulation-based methods over that of Shimada and Yamakawa is that applications are not limited to homogeneous chains or those with uniform bending and twisting rigidities [12].

The statistical-mechanical interpretation of the J factor can be further generalized to include effects of limited bending and twisting flexibility. Flory *et al.* [13] introduced an extension of the Jacobson-Stockmayer theory for rotational-isomeric-state (RIS) models of semi-flexible polymers. Their formula for J is a product of $W(0)$ with two conditional probability densities evaluated at specific parameter values that define the geometry of chain ends in the cyclized conformation, specifically

$$J = \frac{4\pi W(\mathbf{0})\Gamma_0(1)\Phi_{0,1}(\tau_0)}{N_{Av}} \quad (5)$$

where the spatial probability density W is now a function of the end-to-end vector \mathbf{h} and evaluated at $\mathbf{h} = \mathbf{0}$, $\Gamma_{\mathbf{h}}(\gamma)$ is the conditional probability density for the scalar product, $\gamma = \mathbf{l}_1 \cdot \mathbf{l}_n$, of tangent vectors at the ends of the chain, given \mathbf{h} , and $\Phi_{\mathbf{h},\gamma}(\tau)$ is the conditional probability density for the twist angle τ given \mathbf{h} and γ (Figure 1). The last quantity is evaluated at the "natural" twist angle expected for a pair of adjacent segments modulo 2π ; for sufficiently long chains the last factor may need to be replaced by a sum of conditional-probability terms to account for the fact that chain closure can sample topological states with different linking numbers or topoisomers [20]. Given the generality of the RIS model, it is clear that this approach is well justified for helical wormlike chains such as DNA. Moreover, because the problem reduces to evaluating a set of probability density functions, J can readily be computed from Monte Carlo simulations of cyclized chains [12,18,19,21,22].

DISTINCTIONS BETWEEN DNA LOOPING AND CYCLIZATION

Although the boundary conditions imposed on the ends of the chain by cyclization are well specified by the structure and flexibility of DNA molecules, the constraints involved in DNA-loop formation are wholly dependent on the geometry of the protein or protein complex that mediates the loop. A search for possible boundary conditions that fit a given set of experimental J values is highly inefficient, which makes Monte Carlo simulation impractical for analyzing experimental data. The issue of protein flexibility is also usually neglected in considering the thermodynamics of loop formation, typically because little information is available concerning intramolecular motions in proteins and multiprotein complexes.

We developed a method for computing the J factor for looped protein-DNA complexes, generalizing an approach used to compute J for the cyclization of sequence-dependent DNA circles [23]. The method combines computation of the equilibrium conformation of the DNA circle with subsequent evaluation of statistical-thermodynamic quantities using a harmonic approximation [23]. In this model the DNA conformation is described by parameters defined at dinucleotide steps, i.e., tilt, roll, and twist [24], which allows straightforward incorporation of intrinsic or protein-induced DNA curvature at the base-pair level. The method is similar in principle to the approach taken by Shimada and Yamakawa [17] in that the theory takes advantage of small fluctuations around one stable mechanical configuration in small DNA circles (*e.g.*, less than ~ 1000 bp). We treat the protein subunits mediating the loop as virtual base pairs in the cyclized molecule, forming a connected set of rigid bodies with a limited number of degrees of freedom between the subunits. Once the mechanical equilibrium conformation of the circle is found, fluctuations around the equilibrium conformation are taken into account with the harmonic approximation. The new method is about four orders of magnitude more efficient than Monte Carlo simulation and has comparable accuracy [23], making this algorithm suitable for fitting experimental J -factor data using nonlinear least-squares methods [10].

Examining the behavior of J as a function of protein geometry revealed that there are significant quantitative differences between DNA cyclization and looping [9]. These differences are manifested in the amplitude and phase of J on loop size. Protein-specific geometry and flexibility can couple to the DNA twist in a loop to give unexpected deviations in the periodicity of J compared with that expected according to cyclization theory. Unlike cyclization, multiple looped conformations involving the same protein structure, but different loop geometries, can

coexist. These details should be considered in analyzing DNA-loop formation both *in vitro* and *in vivo*.

MEASURING DNA LOOPING

Two particular aspects of looped DNA structures have been exploited in *in-vitro* and *in-vivo* experiments. One is the cooperative binding of a protein to its two cognate sites, which can be demonstrated by footprinting methods [25]. DNA looping can increase the occupancies of both binding sites; in particular, it can significantly enhance protein association to the lower-affinity site because of the tethering effect of DNA looping. This is believed to be a general mechanism by which many transcription factors recruit RNA polymerases in gene regulation [26,27] (we note, however, that recent data suggest a more complex paradigm may operate in a chromatin environment; see [28]). Another hallmark is the helical dependence of loop formation for sufficiently small loops [29,30], which arises because of DNA's limited torsional flexibility and the requirement for correct torsional alignment of the two protein-binding sites.

Many methods have been used to directly observe DNA looping *in vitro*, such as scanning-probe [31] and electron microscopy [24], and single-molecule techniques [32]. *In-vivo* assays based on helical dependence, in which the DNA length between two protein binding sites is varied and excess repression or activation of a reporter gene is measured [29,33], have been a powerful tool in bacterial systems. Several techniques have been developed in the last decade to investigate loop formation in the cells of higher organisms. Chromosome-conformation-capture (3-C) technology and variants thereof [34] make use of non-specific protein-protein crosslinking combined with digestion and religation of protein-bound DNA fragments to identify long-range interactions across complex genomes.

Most of the available techniques give only relative values of J because it is not normally possible to determine the value of K_d that appears in equation (1) under the exact conditions of the *in-vitro* or *in-vivo* experiment. The most rigorous and reliable approach for measuring absolute values of J *in vitro* is to determine K_c and K_d from measurements of pairs of forward and reverse rate constants for loop formation and dimerization, respectively. In enzymatically catalyzed reactions such as ligase-dependent cyclization [35-37] the intramolecular and intermolecular reactions are typically monitored in separate experiments. This is because the intermolecular pathway generally requires significantly higher concentrations of DNA substrate and enzyme for the reaction to occur at a measurable rate. After correcting for effects of substrate and enzyme concentration by extrapolation, the equilibrium and rate constants obtained in these experiments are regarded as apparent values.

Recently we showed that the kinetics of site-specific recombination mediated by the Cre recombinase of bacteriophage P1 gives quantitative measurements of the absolute value of J for DNA loops in the size range 870 to 3050 bp [38]. Recombinase-based measurements of loop formation have an important advantage over ligase-catalyzed cyclization in that the reaction does not require free DNA ends. Thus, DNA-looping assays can be carried out on covalently closed or supercoiled DNA, a strict requirement for direct J-factor measurements *in vivo*.

The complexity of the recombinase mechanism presents a significant challenge in these measurements. Unlike the ligase reaction, simple Michaelis-Menten enzyme-kinetics models do not apply; instead, the recombination-kinetics data must be analyzed by solving systems of ordinary differential equations (ODEs). The ODEs contain a set of rate constants for elementary recombinase binding and dissociation steps, k_1, k_{-1}, k_2, k_{-2} in addition to those for site synapsis and recombination k_3, k_{-3}, k_4, k_{-4} [38]. Unlike the recombinase binding/dissociation steps, the

latter four rate constants differ for the inter- and intramolecular mechanisms. The value of J is obtained from the quotient $k_3^{(c)}k_{-3}^{(b)}/(k_{-3}^{(c)}k_3^{(b)})$ for the synapsis/recombination steps shown in Figure 2. Here $k_3^{(c)}$ and $k_{-3}^{(c)}$ are the apparent forward and reverse rate constants, respectively, for intramolecular recombination-site synapsis whereas $k_3^{(b)}$ and $k_{-3}^{(b)}$ are the corresponding values for the intermolecular reaction. In practice all four apparent rate constants k_3, k_{-3}, k_4, k_{-4} need to be determined for both the intra- and intermolecular reactions. Thus, reliable fits to the data require significant numbers of data points for each kinetic curve. In our method we monitor site synapsis and recombination in real time using the FRET signal obtained from fluorophore-labeled substrate DNAs. A typical analysis fits the four free parameters to between 500 and 1000 data points; overall uncertainties in J are conservatively estimated to be in the range of 20 - 50%, which is comparable to those for ligase-catalyzed cyclization.

STATISTICAL-MECHANICAL PARTITION FUNCTIONS AND DNA LOOPING

For a system confined to a fixed volume and at constant temperature there is a direct relationship between J and the canonical-ensemble partition functions, Z_{loop} and Z_{lin} , for a loop and the corresponding linear chain, respectively [23]

$$J = \exp\left[-(G_{loop} - G_{lin})/k_B T\right] = 8\pi^2 \frac{Z_{loop}}{Z_{lin}} \quad (6)$$

Although we use the Gibbs free energies of the two forms as the arguments in the exponential, the correct free energy function is formally the Helmholtz free energy. We consider this difference to be negligible, as is normally the case with systems in condensed phases.

Z takes the form

$$Z = \int \left(\prod_{k=1}^N \prod_{a=1}^3 \frac{1}{h^2} dx_a^k dp_a^k d\theta_a^k dl_a^k \right) \Delta\{\underline{\theta}\} \exp\left(-\beta H\{\underline{x}, \underline{p}, \underline{\theta}, \underline{l}\}\right) \quad (7)$$

where h is Planck's constant and $\beta = 1/(k_B T)$. Here $H\{\underline{x}, \underline{p}, \underline{\theta}, \underline{l}\} = \sum_{k=1}^N T^k(\underline{p}^k, \underline{l}^k) + U_{tot}\{\underline{x}, \underline{\theta}\}$ is the sum of kinetic energy terms plus the total potential energy U_{tot} expressed as functions of noncanonical momenta $\{\underline{p}, \underline{l}\}$ and coordinates $\{\underline{x}, \underline{\theta}\}$, with Δ the Jacobian for the transformation from canonical to noncanonical variables [39]. The integration is performed over all segments of the chain k and the three degrees of freedom for each segment, a . Transformation from canonical to non-canonical coordinates is necessary to facilitate factorization of the kinetic-energy contribution to the integral, namely

$$Z = \left(\prod_{k=1}^N Z_{kin}^k \right) \times \int \left(\prod_{k=1}^N \prod_{a=1}^3 dx_a^k d\theta_a^k \right) \Delta\{\underline{\theta}\} \exp\left(-\beta U_{tot}\{\underline{x}, \underline{\theta}\}\right) \quad (8)$$

This expression for the partition function is trivial to evaluate for the linear chain, but a major challenge to compute for the closed loop because of the presence of multiple non-linear constraints.

A closed-form expression is available for J in the case of the harmonic approximation. The derivation in terms of the partition function ratio given in equation (6) is too lengthy to present here; instead we give the result (for details see [23])

$$J = \frac{8\pi^2 \exp(-E_s/k_B T)}{\sqrt{\pi^6 \det(\mathbf{A}) \det(\mathbf{F})}} \quad (9)$$

where E_s is the elastic energy of the loop in its minimum-energy conformation. Equation (9) contains two matrices \mathbf{A} and \mathbf{F} whose elements are functions of the elastic constants of the chain and also the first and second derivatives of the non-linear constraint functions with respect to angular parameters evaluated at the mechanical-equilibrium conformation.

The range of applicability of the harmonic approximation has not been rigorously determined, but it is clear that its use becomes problematic for loops that contain significant levels of excess supercoiling. This is because the harmonic approximation exclusively consists of local (*i.e.*, nearest-neighbor) interactions between rigid-body base pairs and protein subunits. The lack of long-range contributions to the chain's potential energy is therefore expected to become a serious limitation for DNA loops much larger than 500 bp, a size regime where multiple topoisomer species become populated during random closure of DNA circles [40].

SUMMARY

DNA cyclization and loop formation are related processes that can be described by closely similar thermodynamic and statistical-mechanical formalisms. The thermodynamic quantity relevant to both processes is a ratio of equilibrium constants known as the Jacobson-Stockmayer factor or J factor. The J factor is an extremely useful quantity because it gives the conformational free energy cost of forming a DNA loop independent of the free-energy changes associated with protein-DNA and protein-protein interactions accompanying loop formation.

Most experimental approaches for measuring looping probabilities report on relative, rather than absolute, values of the J factor; therefore, rigorous measurements of this quantity remain a significant challenge. We summarize a recent approach based on the kinetics of the Cre site-specific recombination reaction that enabled the J factor to be measured without the ligation of free DNA ends. This method can be readily implemented with covalently closed DNA molecules, which is a critical requirement for DNA-looping measurements *in vivo*. Such advancements in experimental methods for measuring J will likely motivate future theoretical and computational approaches for evaluating looping free energies in complex nucleoprotein systems.

FUNDING

This work was supported by a grant from the NIH/NSF Joint Program in Mathematical Biology (DMS-0800929 from the National Science Foundation) to SDL and from the National Institutes of Health (5SC3GM083779-03) to A.H.

REFERENCES

1. Ferraiuolo, M.A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X.Q., Nadler, M., Blanchette, M. and Dostie, J. (2010) The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res*, **38**, 7472-7484.
2. Liu, Z., Scannell, D.R., Eisen, M.B. and Tjian, R. (2011) Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell*, **146**, 720-731.
3. Sanyal, A., Lajoie, B.R., Jain, G. and Dekker, J. (2012) The long-range interaction landscape of gene promoters. *Nature*, **489**, 109-113.
4. Becker, N.A., Kahn, J.D. and Maher, L.J., 3rd. (2005) Bacterial repression loops require enhanced DNA flexibility. *J Mol Biol*, **349**, 716-730.
5. Becker, N.A., Kahn, J.D. and Maher, L.J., 3rd. (2007) Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Res*, **35**, 3988-4000.
6. Dekker, J. (2008) Mapping *in-vivo* chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J Biol Chem*, **283**, 34532-34540.
7. Reeves, R. (2010) Nuclear functions of the HMG proteins. *Biochim Biophys Acta*, **1799**, 3-14.
8. Travers, A.A. (2003) Priming the nucleosome: a role for HMGB proteins? *EMBO Rep*, **4**, 131-136.
9. Zhang, Y., McEwen, A.E., Crothers, D.M. and Levene, S.D. (2006) Statistical-mechanical theory of DNA looping. *Biophys J*, **90**, 1903-1912.
10. Zhang, Y., McEwen, A.E., Crothers, D.M. and Levene, S.D. (2006) Analysis of *in-vivo* LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS ONE*, **1**, e136.
11. Jacobson, H. and Stockmayer, W.H. (1950) Intramolecular reaction in polycondensations. I. The theory of linear systems. *J Chem Phys*, **18**, 1600-1606.
12. Levene, S.D. and Crothers, D.M. (1986) Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J Mol Biol*, **189**, 61-72.
13. Flory, P.J., Suter, U.W. and Mutter, M. (1976) Macrocyclization equilibria. I. Theory. *J Am Chem Soc*, **98**, 5733-5739.
14. Shore, D. and Baldwin, R.L. (1983) Energetics of DNA twisting. II. Topoisomer analysis. *J Mol Biol*, **170**, 983-1007.
15. Shore, D. and Baldwin, R.L. (1983) Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J Mol Biol*, **170**, 957-981.
16. Shore, D., Langowski, J. and Baldwin, R.L. (1981) DNA flexibility studied by covalent closure of short fragments into circles. *Proc Natl Acad Sci USA*, **78**, 4833-4837.
17. Shimada, J. and Yamakawa, H. (1984) Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules*, **17**, 689-698.
18. Hagerman, P.J. (1985) Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers*, **24**, 1881-1897.
19. Levene, S.D. and Crothers, D.M. (1986) Topological distributions and the torsional rigidity of DNA. A Monte Carlo study of DNA circles. *J Mol Biol*, **189**, 73-83.
20. Shimada, J. and Yamakawa, H. (1985) Statistical mechanics of DNA topoisomers. The helical worm-like chain. *J Mol Biol*, **184**, 319-329.

21. Czapla, L., Peters, J.P., Rueter, E.M., Olson, W.K. and Maher, L.J., 3rd. (2011) Understanding apparent DNA flexibility enhancement by HU and HMGB architectural proteins. *J Mol Biol*, **409**, 278-289.
22. Czapla, L., Swigon, D. and Olson, W.K. (2008) Effects of the nucleoid protein HU on the structure, flexibility, and ring-closure properties of DNA deduced from Monte Carlo simulations. *J Mol Biol*, **382**, 353-370.
23. Zhang, Y. and Crothers, D.M. (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys J*, **84**, 136-153.
24. Bloomfield, V.A., Crothers, D.M. and Tinoco, I.J. (2000) *Nucleic Acids: Structures, Properties and Functions*. University Science Books, Herndon, VA.
25. Hochschild, A. (1991) Detecting cooperative protein-DNA interactions and DNA loop formation by footprinting. *Methods Enzymol*, **208**, 343-361.
26. Ptashne, M. (1986) Gene regulation by proteins acting nearby and at a distance. *Nature*, **322**, 697-701.
27. Wang, Q., Carroll, J.S. and Brown, M. (2005) Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol Cell*, **19**, 631-642.
28. Kulaeva, O.I., Nizovtseva, E.V., Polikanov, Y.S., Ulianov, S.V. and Studitsky, V.M. (2012) Distant activation of transcription: mechanisms of enhancer action. *Mol Cell Biol*, **32**, 4892-4897.
29. Dunn, T.M., Hahn, S., Ogden, S. and Schleif, R.F. (1984) An operator at -280 base-pairs that is required for repression of *araBAD* operon promoter - addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc Natl Acad Sci USA*, **81**, 5017-5020.
30. Schleif, R. (1992) DNA looping. *Annu. Rev. Biochem.*, **61**, 199-223.
31. Rippe, K., Guthold, M., von Hippel, P.H. and Bustamante, C. (1997) Transcriptional activation via DNA-looping: Visualization of intermediates in the activation pathway of *E coli* RNA polymerase- $\sigma(54)$ holoenzyme by scanning force microscopy. *J. Mol. Biol.*, **270**, 125-138.
32. Finzi, L. and Gelles, J. (1995) Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules. *Science*, **267**, 378-380.
33. Müller, J., Oehler, S. and 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.
34. de Wit, E. and de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization. *Genes Dev*, **26**, 11-24.
35. Crothers, D.M., Drak, J., Kahn, J.D. and Levene, S.D. (1992) DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol*, **212**, 3-29.
36. Taylor, W.H. and Hagerman, P.J. (1990) Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J Mol Biol*, **212**, 363-376.
37. Koo, H.S., Drak, J., Rice, J.A. and Crothers, D.M. (1990) Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry*, **29**, 4227-4234.
38. Shoura, M.J., Vetcher, A.A., Giovan, S.M., Bardai, F., Bharadwaj, A., Kesinger, M.R. and Levene, S.D. (2012) Measurements of DNA-loop formation via Cre-mediated recombination. *Nucleic Acids Res*, **40**, 7452-7464.

39. Hanke, A., Giovan, S.M. and Levene, S.D. (2011) Evaluating partition functions for multiscale models of DNA. *Progress in Theoretical Physics Supplement No. 191*, 109-129.
40. Horowitz, D.S. and Wang, J.C. (1984) Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J Mol Biol*, **173**, 75-91.

Figure Captions

Figure 1. (A.) Experimental data and theoretical dependence for the cyclization probability, J , as a function of DNA size. Three sets of experimental data covering a range of DNA-fragment sizes between 230 and 1100 bp are shown [14-16,36]. Curves show the oscillatory helical-phase dependent (gray) and phase-independent (black) behavior of J in Monte Carlo simulation results [12,19] that were fitted to the data in [16]. **(B.)** Variables describing the boundary conditions at the ends of a cyclizing DNA molecule. $\bar{\mathbf{h}}$ is the end-to-end vector of the chain; $\bar{\mathbf{l}}_1$, and $\bar{\mathbf{l}}_N$, are the initial and terminal vectors tangent to the central axis of the chain, respectively, and have scalar product γ . The vector $\bar{\mathbf{a}}$ is normal to the helix axis and serves as a reference for rotations about this axis; the broken line is its projection in the plane normal to the distal end of the helix axis when $|\bar{\mathbf{h}}|$ and γ are equal to zero and unity, respectively. The torsional free energy of the chain depends on the angle τ evaluated from the scalar product of the projection of $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ and a vector normal to both $\bar{\mathbf{l}}_N$ and $\bar{\mathbf{b}}$. The arctangent of the quotient gives τ . Figure adapted from [12].

Figure 2. J-factor measurements obtained via intra- and intermolecular Cre site-specific recombination of loxP target sequences. **(A.)** Intramolecular site-synapsis and resolution steps in the Cre mechanism. A linear DNA bearing directly repeated Cre₂loxP sequences is converted to an intermediate synaptic complex with respective forward and reverse rate constants, $k_3^{(c)}$ and $k_{-3}^{(c)}$ followed by resolution of the intermediate to products with rate constants $k_4^{(c)}$ and $k_{-4}^{(c)}$. The loxP sites are labeled with donor or acceptor fluorophores, which permits monitoring of the synapsis reaction by time-dependent FRET. **(B.)** Fluorescence signal, $F(t)$, which monitors donor quenching via FRET during intramolecular site synapsis and recombination. Fluorescence decays are for molecules having 3-kbp and 870-bp DNA loops. Non-linear least-squares fits to the data to a system of ordinary differential equations (ODEs, solid curves) gave values of the rate constants $k_3^{(c)}$, $k_{-3}^{(c)}$, $k_4^{(c)}$, and $k_{-4}^{(c)}$. **(C.)** Intermolecular site-synapsis and resolution steps in the Cre-loxP reaction scheme. Equimolar ratios of donor-labeled, T^DB, and acceptor-labeled, TB^A, duplexes were used in reactions having total loxP concentrations equal to 12.5 or 25 nM. **(D.)** Intermolecular rate constants were obtained by non-linear least-squares fits to a set of ODEs similar to those in **(B.)**, which describe the time-dependent concentrations of reactants, intermediates and products for the intermolecular recombination pathway.



