5074–5085 *Nucleic Acids Research, 2017, Vol. 45, No. 9* *Published online 4 February 2017*
doi: 10.1093/nar/gkx069

Efficient chromosomal-scale DNA looping in *Escherichia coli* using multiple DNA-looping elements

Nan Hao[1], Kim Sneppen[2], Keith E. Shearwin[1] and Ian B. Dodd[1,*]

[1]Department of Molecular and Cellular Biology, University of Adelaide, North Terrace, Adelaide SA 5005, Australia and [2]Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, Copenhagen Ø 2100, Denmark

Received October 20, 2016; Revised December 14, 2016; Editorial Decision January 17, 2017; Accepted January 25, 2017

ABSTRACT

Genes are frequently regulated by interactions between proteins that bind to the DNA near the gene and proteins that bind to DNA sites located far away, with the intervening DNA looped out. But it is not understood how efficient looping can occur when the sites are very far apart. We develop a simple theoretical framework that relates looping efficiency to the energetic cost and benefit of looping, allowing prediction of the efficiency of single or multiple nested loops at different distances. Measurements of absolute loop efficiencies for Lac repressor and λ CI using gene expression reporters in *Escherichia coli* cells show that, as predicted by the model, long-range DNA looping between a pair of sites can be strongly enhanced by the use of nested DNA loops or by the use of additional protein-binding sequences. A combination of these approaches was able to generate efficient DNA looping at a 200 kb distance.

INTRODUCTION

Cooperative binding of proteins to separate sites on the same DNA molecule mediated by DNA looping is a core mechanism of transcriptional regulation and other DNA-based processes (1–5). DNA acts as a tether that links the proteins bound to it and can foster their interaction to form specific complexes that promote or repress transcription (6,7). The length of the DNA link or loop between interacting protein-bound sites can vary from a few base pairs (bp), as for adjacently bound proteins, to megabasepair distances, as for the interaction between some eukaryotic enhancers and their promoters (8,9). It remains a puzzle how efficient and specific interactions of DNA-bound proteins can occur when the length of the DNA tether between them is very large.

Genome-wide ligation-based DNA proximity screens, such as Hi-C (10), have revealed a multitude of DNA loops in a variety of organisms (e.g. (11–14)). However, the technique does not provide absolute looping efficiencies, and the functionality of the vast majority of observed loops is unclear. Nevertheless, many of the interactions appear to be mediated by specific DNA-binding proteins and have functional significance in gene regulation and genome organization (11,15,16). The ubiquity and complexity of these DNA loops suggests that a generalizable theory for protein-mediated DNA looping will be needed to properly understand and manipulate their formation.

The formation of a DNA loop constrains the DNA structurally and spatially, and thus entails an energetic cost. The resistance of DNA to twisting and bending is expected to dominate the cost of looping at short DNA separations (7,17). However, the loss of entropy due to the restriction of the spatial freedom of the interacting DNA sites increases with the DNA distance d (bp) between them, and is expected to be the major cost for formation of long DNA loops. The cost of DNA looping is often quantitated by the factor J (or J_{loop}). J is the effective concentration of one DNA site relative to the other DNA site. The lower the value of J, the higher the cost of looping. Thus in the entropy-dominated range, J decreases with increasing DNA distance (7).

Efficient DNA looping can occur because the looping cost can be counterbalanced by the energetic benefit provided by the protein–DNA and protein–protein interactions involved in protein-mediated bridging between the DNA sites. However, this benefit due to the interaction between the sites should be the same at different distances, while the cost of DNA looping increases with distance, at least in the entropy-dominated range. Thus, the efficiency of looping will decrease as the separation between the sites increases. Measurements of the relative efficiency of the *same* protein-mediated DNA–DNA interaction *in vivo*, show a steady weakening of the interactions with DNA distance (17–21). We have analyzed looping by the Lac repressor (LacI) and bacteriophage λ CI over a range of distances in *Escherichia coli* (22,23). In these studies, estimates for absolute looping efficiencies were obtained by *in vivo* measurement of enhancement of repression by DNA looping, combined with models incorporating the extensive physicochemical information about these proteins. We found that while LacI could give ∼30% looping between sites separated

*To whom correspondence should be addressed. Tel: +61 8 8313 5362; Fax: +61 8 8313 4362; Email: ian.dodd@adelaide.edu.au

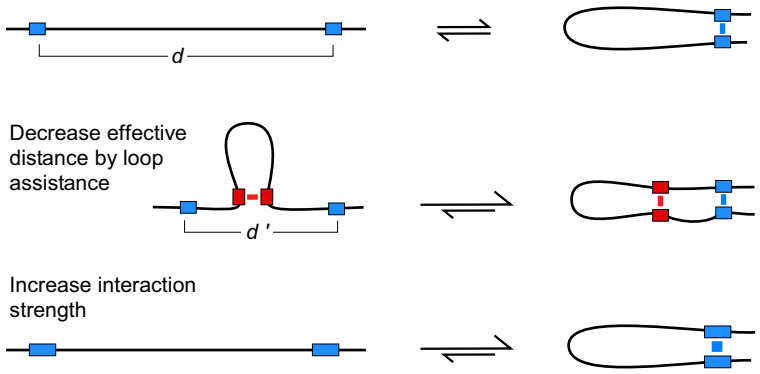

Figure 1. Alternative mechanisms to increase the efficiency of DNA looping.

by 5.6 kb, and CI could give ∼50% DNA looping with a 10 kb separation between sites, looping by these proteins was undetectable at 500 kb separations (23).

The efficiency of long-range DNA looping may be enhanced by decreasing the DNA looping cost or by increasing the interaction benefit (Figure 1). A way to reduce the cost of looping is to use other DNA loops that bring the sites nearer to each other, reducing their effective distance, effectively increasing *J*. This loop assistance effect has been shown using synthetic constructs in eukaryotic cells (24,25). Loop assistance is also assumed to underlie natural and often much longer range enhancer–promoter targeting (26–29). However, in these cases the magnitude of the change in looping efficiency is not clear. Our measurements of the loop assistance effect in *E. coli* for nested arrangements of LacI- and CI-mediated loops showed only modest improvements in efficiency, with the formation of a 2 kb external loop increased by only 1.1- or 1.3-fold when the other protein was able to form an internal 1.4 kb loop (30). Thus, it is not clear whether large improvements in looping efficiencies can be achieved with this mechanism. Increasing the interaction benefit of DNA looping, by increasing the strength of the protein–DNA and protein–protein interactions that hold the two DNA sites together (Figure 1), could also be used to enhance the efficiency of DNA looping (Figure 1), but may also be limited by the biochemistry of the system. For example, for LacI, where the protein–protein and protein–DNA interactions are very strong, looping *in vivo* becomes limited by the difficulty of maintaining LacI at the very low concentrations needed to avoid LacI tetramers binding to both sites and breaking the loop (23,31).

Here we develop a simple theory for single and multiple DNA loops that allows prediction of the fractions of the different looped states from the relevant *J* values and a factor *I* that defines the interaction benefit provided by each of the DNA looping elements involved. The theory predicts that loop assistance should be capable of strong enhancement of long DNA loops. Using LacI and λ CI, we show experimentally that nested loops can give large increases in looping efficiency, driving efficient looping at a 50 kb distance. We show also that the interaction benefit provided by CI can be substantially increased by addition of CI operators, enabling strong loop assistance-dependent LacI looping at a distance of at least 200 kb.

MATERIALS AND METHODS

Strain constructions

All constructs were integrated in single copy in the chromosome of E4643 = MG1655 *rph*$^{+}$ Δ*lacIZYA* (22).The *lacZO2*$^{-}$ gene and the proximal operator/promoter sites (*PlacUV5.lacO2*, λ*OR.PRM* and *4xCI* element) were integrated into λ*attB*, as described (23) (Supplementary Figures S5 and 6). The distal operator sequences (*lacOid*, λ*OL* and *4xCI* element) were either adjacent to the proximal sites (for the 2 kb loop reporters) or present on a separate kanamycin resistance module inserted by recombineering into other chromosomal sites (for the 20, 50, 100 and 200 kb reporters; Supplementary Figure S5). Chromosomal LacI and CI expression modules are shown in Supplementary Figure S7. LacI was expressed from a *PlacI.lacI*$^{+}$ fragment integrated at φHK022 *attB* (23). Lambda CI was expressed from an *OR.PRM.cI.OL1*$^{+}$*2*$^{+}$*3*$^{-}$ module integrated at 186 *attB*, producing 3.3 ± 0.33 wild-type lysogenic units (WLU) of CI (22). Strains not expressing LacI or CI contained integrated empty vectors. Additional details are provided in Supplementary Material.

LacZ assays

A microtitre plate-based kinetic assay was used, with strains grown at 37°C to late log phase in minimal medium + glycerol (1 × M9 salts [10 × M9 salts = 67.8g of NaH_2PO_4, 30.0g of KH_2PO_4, 10g NH_4Cl and 5g NaCl/L H_2O], 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.01 mM $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$, 0.4% glycerol), as described (23).

Extracting estimates of F, J and I from LacI and CI looping reporter data

The reporter data was fitted to extract the key DNA-looping parameters *F* and *J,* by a Monte-Carlo fitting procedure that uses statistical-mechanical models of LacI- and CI-mediated DNA looping regulation of *PlacUV5* and λ*PRM*, and incorporates uncertainty in fixed parameter estimates and data, as previously described (30) (see also Supplementary Material) but with minor variations. For LacI looping, we introduced a parameter for the free [LacI] when one tetramer is removed, $L' = 16.4 \pm 3.3$ nM. For CI looping, we introduced some errors into the parameters for the

free energy change for two CI tetramers forming an octamer $\Delta G_{PTN} = -9.1 \pm 0.5$ kcal/mol (32), and the concentration of CI monomers [CI] = 3.3 ± 0.33 WLU (22). *J* was obtained from ΔG_{oct} (the free energy for *OL-OR* looping by CI octamerization) by the relation $\Delta G_{oct} = \Delta G_{PTN} - RT\ln J$ (M), where $RT = 0.616$ kcal/mol at 37°C. The complexity of CI looping results in large uncertainty in the *J* estimates (23). To better fix these estimates and align them to the estimates from the LacI results, the fitting was done to a combined dataset comprising the CI looping data of Figure 4C, and our previous CI looping data obtained for *OL-OR* separations of 1200, 1500, 1800 and 2000 bp in the absence of LacI looping (30). We made the assumption that the *J* values obtained for these previous data should be the same as the *J* values obtained by analysis of LacI looping for the same separations. Each fitting thus scored not only the match to the data but also the match of each of the *J* values obtained for these data to those expected from the power law of Supplementary Figure S3 (*J* = 296, 234, 193, 173 nM for 1200, 1500, 1800 and 2000 bp spacings). *J* values for the very long range loops (Figure 4C) were not constrained. *I* values were calculated using Equation 1 (Figure 2A).

Two loop model fitting

An iterative Monte Carlo fitting procedure was used to optimize the match between the parameter values for the one-loop and two-loop equations (Equation 1, Figure 2A; Equation 2, Figure 3A) and each set of observed loop fraction *F* values. Each of the three datasets (2–50 kb LacI looping, Figure 5A; 2–50 kb CI looping Figure 5C; 50–200 kb looping Figure 6C) provides six observed loop fraction *F* values (F_1 and $F_{1(2)}$ over three spacings). The fitting runs were repeated with the fixed parameter values and the observed *F* values randomly varied for each run according to their standard deviations and the normal distribution (23). For each spacing, the J_s value was calculated from the $J_{asa'}$ value using the power law exponent of Supplementary Figure S1, $J_s = J_{asa'}(d_s/d_{asa'})^{-1.054}$, where d_s and $d_{asa'}$ are the loop lengths. In each fitting, the variable parameter values were iteratively varied by successive random steps (usually 10^5 iterations) to minimize $\Sigma((F_{observed}-F_{expected})^2/F_{sd})$, where F_{sd} is the standard deviation of the observed *F* estimate. Note that the assumption is made that $J_{asa'}$ in absence of internal protein binding (used for F_1 fit) is the same as $J_{asa'}$ when the internal protein binds but does not form its loop (used for $F_{1(2)}$ fit).

For the 2–50 kb data fitting (Figure 5A and B), individual $J_{asa'}$ values for each spacing and I_{LacI} and I_{CI} values for all spacings were fixed. Individual $J_{aa'}$ values for each spacing were allowed to vary. For the 50–200 kb data fitting (Figure 6B), the $J_{aa'}$ value and I_{LacI} were fixed and applied to all spacings. The $J_{asa'}$ value for the 50 kb spacing was fixed, while $J_{asa'}$ for the 100 and 200 kb spacings and the I_{CI} value for all spacings were allowed to vary.

The means and standard deviations for the fixed and fitted parameters, as well as for the expected *F* values (listed in Figures 5AB and 6B) were obtained from 100 good fits.

Modeling of more than two nested loops

Figure 7A uses a program to calculate *F* for up to 10 nested loops. The length of each additional pair of loop arms (a + a′, b + b′ …) can be specified, as well as individual *I* values for each looping element (Figure 7A shows the case where all loop arms and *I* values are equal). The program calculates the weight of each possible looping combination (each looping element looped or unlooped) by multiplying the *J* values for each component loop, based on the loop length and the power law of Figure 7B and dividing by the *I* values for each of the looped elements, as in Figure 3A and Supplementary Figure S4. The fraction of loop 1 is then calculated as the sum of weights for species where it is looped divided by the sum of weights of all species.

RESULTS

The looping interaction factor *I*

Our aim was to develop a generalizable framework for DNA looping that would allow predictability of DNA looping for combinations of different DNA segments and different DNA looping elements. To do this, we first developed a simple model for single DNA loops that relates the efficiency of looping to the balance between its cost, defined by the factor *J*, and its benefit, defined by a factor *I*.

The purpose of *I* is to combine all of the looping parameters that are *not* dependent on the DNA between the interacting sites. The derivation of *I* for our LacI looping assay system is shown in Figure 2. In this system a distal *lac* operator cooperates by DNA looping with a promoter-proximal *lac* operator overlapping a promoter (23,30); Figure 2A). Our statistical-mechanical model for this system contains seven species, with weights for each species given by the concentration of LacI tetramers (*L* and *L′*), dissociation constants for the two operators (K_P and K_D), a scalar *R* representing the strength of RNAP occupation of the promoter and *J* (23). The efficiency of looping, or the fraction of time that the loop is present, *F*, is given by the sum of weights of the looped species (in this model a single species) divided by the partition sum, which can be expressed simply as $F = J/(J+I)$, where *I* is a function of the dissociation constants, *R* and LacI concentration (Equation 1; Figure 2A). For our LacI looping assay system, where $L = 18$ nM and looping is between a distal *Oid* operator and a proximal *O2* operator at the P_{lacUV5} promoter, $I \sim 80$ nM can be calculated from our *in vivo* parameter estimates (23). Equation 1 should be applicable to any DNA looping protein, including those with alternative DNA-looped species, as long as *J* is similar for each looped conformation, which we expect to be the case in the entropy-dominated regime. Situations where Equation 1 is not applicable are explored in Supplementary Figure S2.

The model for any single DNA loop can thus be simplified to just two species, unlooped and looped, each species representing the combined unlooped species and the combined looped species ((30); Figure 2B). Assigning the unlooped species a statistical weight of 1, it can be seen that the weight of the looped species must be *J*/*I* (Figure 2B). This term shows most clearly how the balance of *J* and *I* determines the favorability of the looped species. Looping is favored by a high *J* (that is, a low energetic cost of loop-

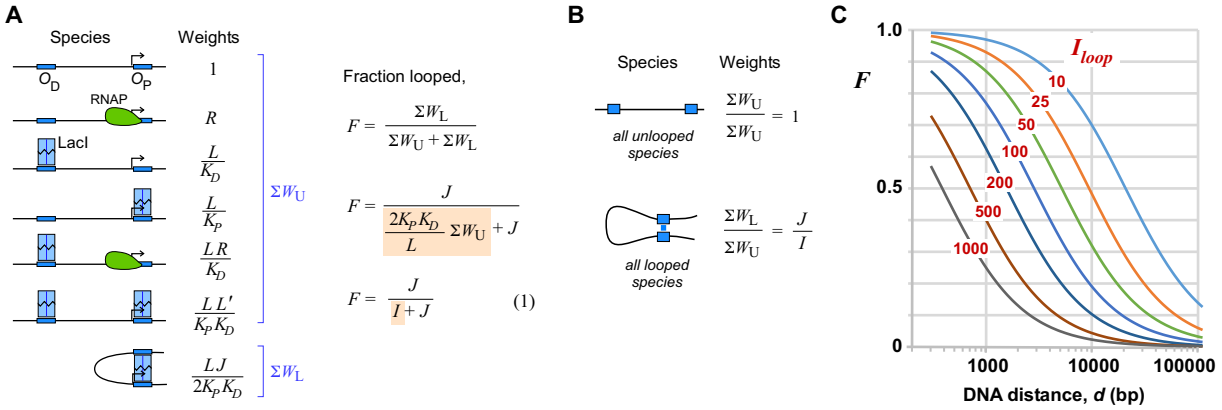

Figure 2. The interaction factor I and DNA looping. (**A**). Derivation of I from the statistical mechanical model for a promoter controlled by binding of LacI tetramers to proximal and distal lac operators (23). L is the free concentration of LacI tetramers ($L' = L$–1.6 nM is the free concentration of LacI tetramers when one is removed), K_D and K_P are dissociation constants, R is a scalar representing RNAP binding. J is the effective concentration of the DNA binding domain of the DNA-tethered LacI tetramer at the unoccupied lac operator. Note that the 1/2 term in the looped species weight is needed to reflect that the DNA-tethered LacI has only one available DNA-binding domain for loop closure while a LacI in solution has two available DNA-binding domains (31); note that species 7 subsumes the parallel and antiparallel loop orientations. The fractional looping F is given by the sum of weights of looped species divided by the sum of weights of all species. I is a function of all the system parameters except J and is inversely related to the interaction strength provided by the LacI-DNA bridge. (**B**) A simplified looping model which lumps together all the unlooped species and all the looped species. Setting the weight for the unlooped species to 1 gives the weight for the looped species as J/I. Thus, looping is favored by high J and low I. (**C**) Predicted effect of DNA separation on the looped fraction F, for sites with different interaction strengths, I. Equation 1 was used, with the J values for each d calculated from previous measurements of LacI looping ((23); Supplementary Figure S1). Note that for $d > 5600$ bp, this involves extrapolation of the J versus d relationship beyond experimental estimates.

ing the DNA) and a low I (that is, a high energetic benefit of the protein-mediated bridge). For efficient DNA looping, I should be small relative to J. Figure 2C shows how the efficiency of looping by a given DNA looping element is predicted to decrease with increasing DNA separation, d, based on extrapolation of our empirically derived power law relationship for J versus d in *E. coli* for LacI looping ((23); Supplementary Figure S1), and how lower values of I increase the efficiency of looping.

This treatment partitions the factors controlling DNA looping into two independent parameters. J should be independent of the protein–DNA looping element that creates the loop—the same J should apply whether a particular segment of DNA is looped by LacI, CI or any other protein. Similarly, I should be independent of the DNA in the loop—the same I should apply for a particular protein at a particular concentration and with particular binding sites no matter which DNA is being looped. This separation of I and J allows for a 'modular' analysis of DNA looping in more complex situations where multiple loops and multiple DNA-looping elements are involved.

Theory for loop assistance by two nested loops

We extended this analysis to nested two-loop arrangements (Figure 3A). Here, sites looped by a different protein (protein 2) are located distances a and a' bp internal to the sites looped by protein 1, with a spacer of length s bp between the two internal sites. The protein 1 and protein 2 looping elements each have a distinct I value, I_1 and I_2. The two different single loop weights are given by simple J/I terms, J_s/I_2 for the protein 2 loop and $J_{asa'}/I_1$ for the protein 1 loop, as in Figure 2B. The weight for the double-looped species is a product of the J_s/I_2 for the internal loop, and a special $J_{aa'}/I_1$ term for the closure of the now smaller a + a' external loop by protein 1. Loop assistance occurs if $J_{aa'} > J_{asa'}$, that is, if the closure of the a + a' loop by protein 1 is more favorable than its closure of the a+s+a' loop. Since the a + a' loop is smaller than the a+s+a' loop, this will always be the case as long as the loops are large enough so that J is dominated by entropic factors. Thus, loop assistance works by shortening the effective distance between looping elements and increasing the effective J.

Figure 3B shows the expected effect of loop assistance on DNA looping efficiency at a range of DNA separations due to a pair of sites nested 300 bp internal to the external loop sites. We used $I = 50$ nM for both proteins (a rough average of $I \sim 80$ nM for LacI and $I \sim 30$ nM for our λ CI looping element—see below), and our measured relationship between DNA distance and J ((23); Supplementary Figure S1). We also assumed that $J_{aa'}$ for the 300 + 300 bp loop is the same as for a 600 bp DNA-only loop ($J = 616$ nM) and that $J_{aa'}$ is constant for the different s spacings.

The model predicts that loop assistance can provide large increases in the efficiency of formation of the external loop at large DNA distances, where the loop forms poorly by itself. The formation of the internal loop is also stimulated, indicating that loop assistance could be strongly synergistic.

Enhancing looping by increasing J: loop assistance can drive efficient formation of 20 and 50 kb DNA loops *invivo*

We previously demonstrated moderate loop assistance *in vivo* between LacI loops and λ CI repressor loops at a 2 kb spacing of the external sites (30). However, the above modeling indicates that loop assistance should be capable of promoting efficient looping over substantially longer DNA distances. To test this we examined nested loops formed by

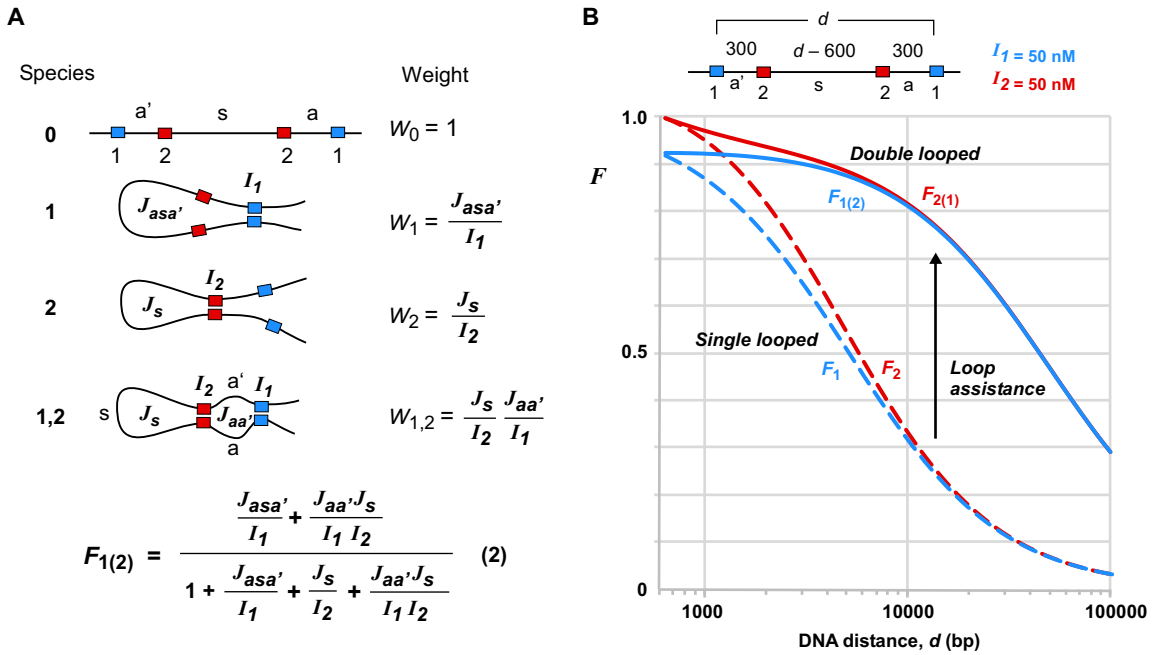

Figure 3. Model for two nested loops and loop assistance. (**A**). The simple single-loop model of Figure 2B can be extended for two nested protein-mediated DNA loops. Weights for the single and double-looped species are given by I terms for each set of the looping elements (I_1 and I_2), and J terms for the three loops, $J_{asa'}$, J_s and $J_{aa'}$. These weights allow calculation of $F_{1(2)}$, the fraction of the protein 1 loop formed in the presence of looping element 2 (shown) and $F_{2(1)}$, the fraction of the protein 2 loop formed in the presence of looping element 1. (**B**) Model-predicted looping fractions versus DNA distance for individual protein loops in the absence (F_1 and F_2) or presence of the other looping element ($F_{1(2)}$ and $F_{2(1)}$), showing loop assistance. As in Figure 2C, the J values for each d were calculated from previous measurements of LacI looping ((23); Supplementary Figure S1). $I = 50$ nM was used for both looping elements.

LacI and CI over 20 and 50 kb separations. We tested arrangements in which a CI loop was located internal to a measured LacI loop, and in which a LacI loop was located internal to a measured CI loop.

Looping by LacI or CI was assayed by measuring how much the presence of a distal LacI or CI binding site improved the repression of a LacI- or CI-controlled promoter that expresses *lacZ* ((23); Figure 4A). For LacI looping we use reporter constructs with *lacZ* expressed from a *lacUV5* promoter controlled by a proximal weak *lacO2* operator, giving ∼50% repression at a fixed low LacI concentration (18 nM). In the presence of a strong distal *lacOid* operator site, *lacZ* expression is decreased from this level roughly proportionally to the fraction of looping (Figure 4A). For measuring CI looping, the *lacZ* gene is expressed from the *PRM* promoter under the control of the proximal *OR* operator site and the distal *OL* site. *OL* and *OR* each contain three CI operators, each able to bind a CI dimer. At the high fixed CI concentration we use (3.3 WLU (22)), *PRM* is activated by a CI tetramer bound to *OR12* but in the absence of *OL*, the repressive *OR3* site is almost completely unoccupied. Looping between *OL* and *OR* due to interactions between this tetramer and a tetramer at *OL*, allows an additional cooperative interaction between a third CI dimer at *OL* and a dimer binding to *OR3*, giving repression of *PRM* (Figure 4A; (22)). Model-based analysis of these reporter data, using available estimates of the underlying biochemical parameters for LacI and CI, allows extraction of the key looping parameters F and J (Materials and Methods; (23)).

For the Lac-external CI-internal nested looping experiments, we began with our existing reporter in which *Oid* and *O2* are separated by 2090 bp, with *OR.PRM* located 390 bp upstream of O2 and with *OL* located 300 bp internal to *Oid* ((30); Figure 4B). To construct the 20 and 50 kb reporters, the DNA portion containing the two promoter-proximal binding sites and the *lacZ* gene was kept the same as in the 2 kb reporters, but the DNA portion containing the two distal operators (or O^- controls) was removed and placed further away on the bacterial chromosome by recombineering, increasing the spacer DNA between the two internal sites from 1400 to 19 700 or 49 600 bp. Steady-state *lacZ* expression was measured in the presence or absence of LacI or CI.

When the distal *Oid* site was located 2090 bp upstream, repression was strongly enhanced due to LacI looping with efficiency $F = 0.69 \pm 0.01$ (Figure 4B). In the presence of CI and its internal binding sites, loop assistance increased looping efficiency 1.3-fold to $F = 0.89 \pm 0.004$. Looping by LacI alone was weak at the 20 kb spacing, $F = 0.12 \pm 0.02$, and almost undetectable at the 50 kb spacing, $F = 0.025 \pm 0.02$. However, in the presence of CI, looping increased 3.7-fold to $F = 0.44 \pm 0.01$ for the 20 kb spacing and 8.8-fold to $F = 0.22 \pm 0.01$ for the 50 kb spacing (Figure 4B).

Very similar results were obtained with the CI-external LacI-internal reporters. In our existing reporter, the *OL* and *OR* sites are separated by 2000 bp, with the LacI *Oid* and *O2* sites each located 300 bp internal to the CI sites ((30); Figure 4C). At this distance a high CI looping efficiency, $F = 0.84 \pm 0.03$, occurs even in the absence of loop assistance.

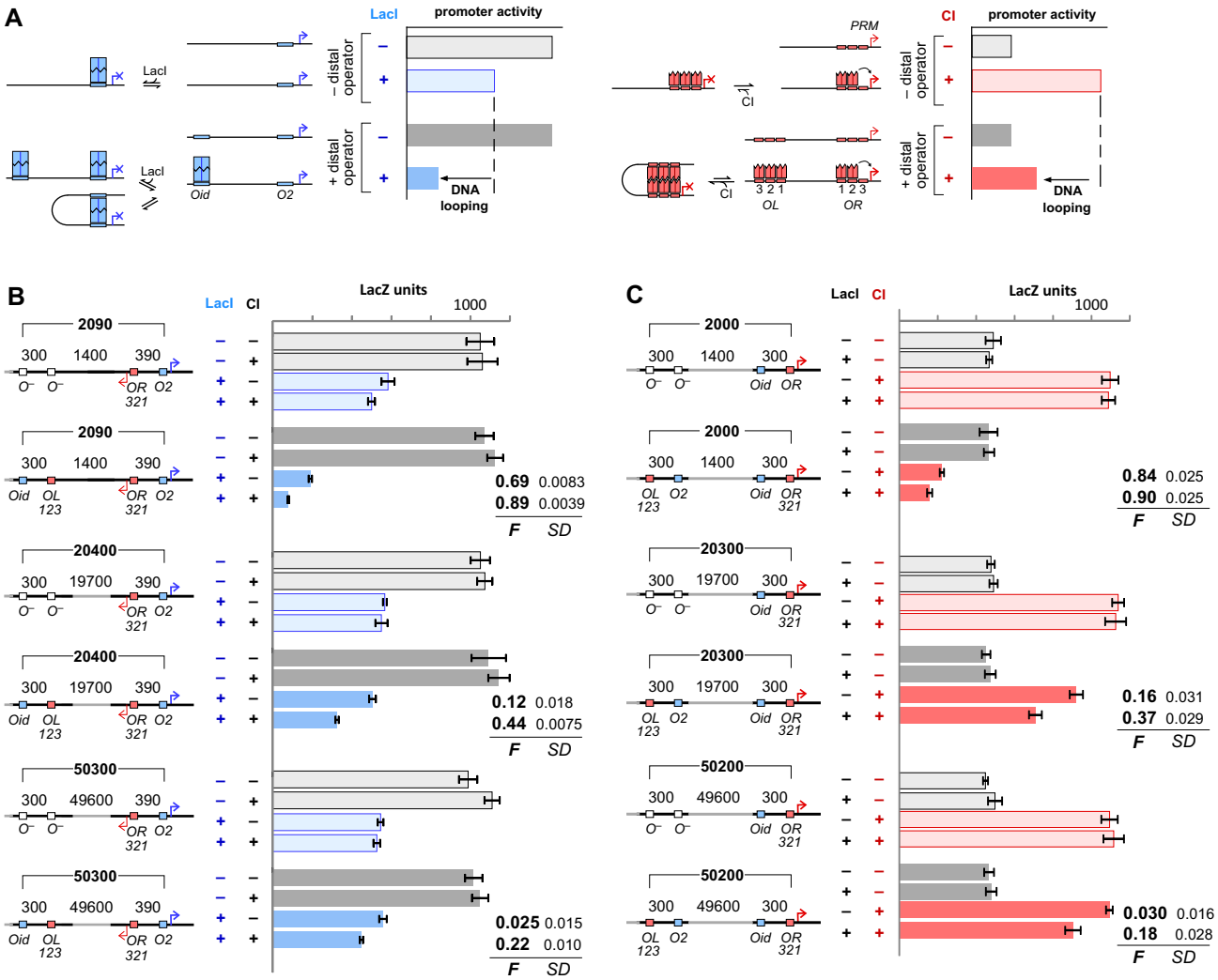

Figure 4. Loop assistance can promote very long-range looping. (**A**) Rationale of the looping assay. See text for details. (**B**) Looping by LacI between *Oid* and *O2* operators separated by 2, 20 or 50 kbp is increased by the presence of CI and internally nested *OL* and *OR* operators. (**C**) Looping by CI between *OL* and *OR* operators separated by 2, 20 or 50 kbp is increased by the presence of LacI and internally nested *Oid* and *O2* operators. (B and C) The promoter proximal segment was fixed on the chromosome at λ*attB*, while the distal segment was inserted into the chromosome ∼2, 20 or 50 kb upstream of the distal site by recombineering (see Supplementary Material and Supplementary Figure S6). Data error bars are 95% confidence intervals, $n = 9$. Calculated fractional looping F values are ± standard deviations.

Addition of LacI increased CI looping slightly to $F = 0.90 \pm 0.03$. Looping by CI alone was considerably weaker at the 20 and 50 kb spacings, $F = 0.16 \pm 0.03$ and $F = 0.03 \pm 0.02$, respectively. In the presence of the internal LacI loop, CI looping was increased 2.3-fold to $F = 0.37 \pm 0.03$ for the 20 kb spacing, and 6-fold to $F = 0.18 \pm 0.03$ for the 50 kb spacing (Figure 4C).

Thus, as predicted by the model, looping by a single protein at very long DNA distances can be strongly assisted by a nested loop caused by a second protein, with the loop assistance effect becoming stronger as unassisted looping gets weaker. Although our assay cannot measure F for the internal loops, it is reasonable to assume that the 49 600 bp internal CI or LacI loops, like the 50 200 external loops, form rarely by themselves. Thus, two very weak DNA loops can combine synergistically to drive substantial looping at 50 kb.

Application of the two-loop model

The results of the 2, 20 and 50 kb loop reporters conform generally with the expectations of the two-loop model. However, the data allow a more detailed examination of the predictive value of the theory, as they provide reasonable estimates for most of the terms in the equation for loop assistance (Equation 2, Figure 3A).

As well as measuring looped fractions F_1 and $F_{1(2)}$, the analysis of the data of Figure 4B and C also provides direct estimates of $J_{asa'}$ for the single, unassisted external LacI or CI loops, along with I_{LacI} or I_{CI} values (values listed in Figure 5A and B). The data do not provide measurements of J_s (for the internal loop) but estimates of these values can be made from the measured $J_{asa'}$ values and using the exponent of the power law for our previously observed J versus d relationship (Figure 5A and B). This leaves $J_{aa'}$ as the only unknown in Equation 2. We used an iterative fitting proce-

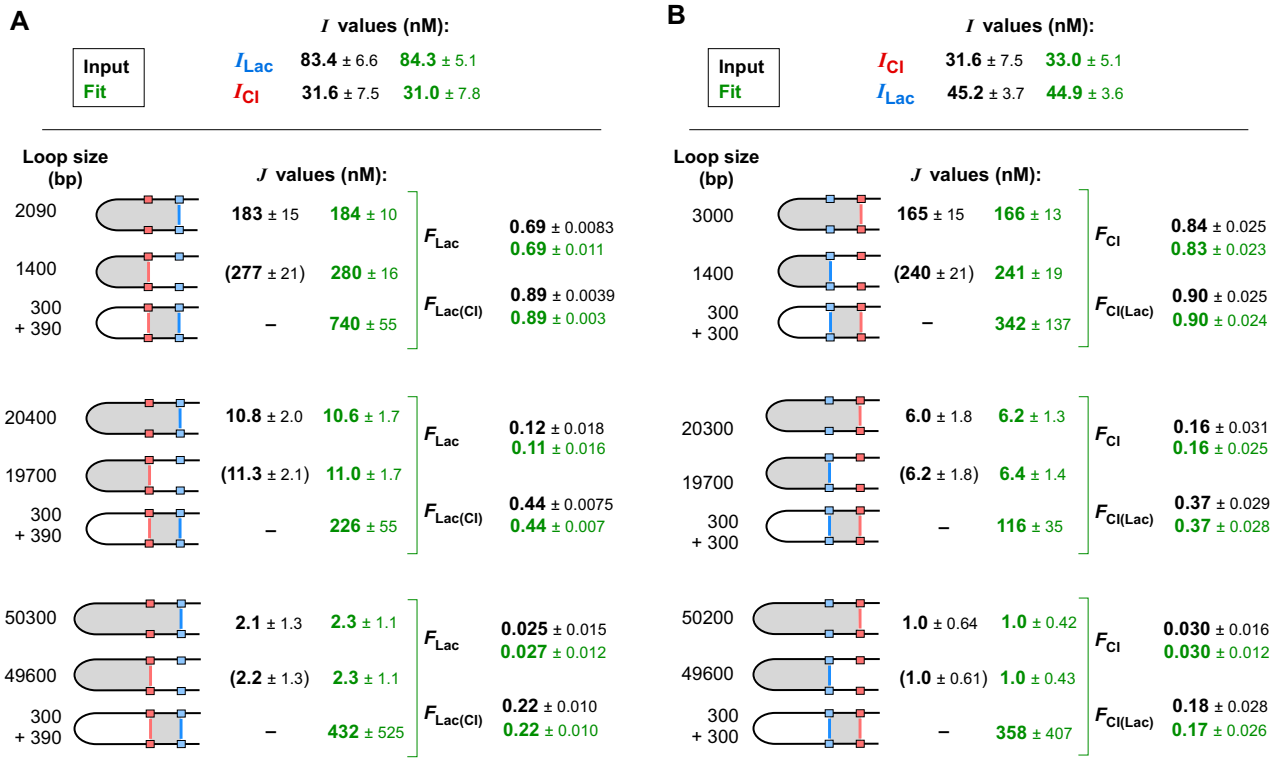

Figure 5. Application of the two-loop model to loop assistance at 2, 20 and 50 kb. Estimation of $J_{aa'}$ for the 300 + 390 bp loop (**A**) or the 300 + 300 bp loop (**B**) by Monte-Carlo fitting of the F_{LacI} and $F_{LacI(CI)}$ values (A) or the F_{CI} and $F_{CI(LacI)}$ values (B) from the 2, 20 and 50 kb LacI/CI 2-loop reporters (Figure 4) using the 2-loop model (Supplementary Material). *J* values in black are inputs to the fitting: $J_{asa'}$ values were obtained directly from the analysis of the data of Figure 4, and J_s values (in parentheses) were estimated from the $J_{asa'}$ values using the power law exponent of Supplementary Figure S1. The input *I* values for LacI (A) and CI (A and B) were obtained from the analysis of Figure 4A and B, respectively. The input *I* values for LacI (B) was calculated from that in A to account for the lack of the promoter in the internal O2 element ($R = 0$). Parameter values in green/gray are fitted values that optimize the match between the observed and model-predicted F_{LacI}, F_{CI}, $F_{LacI(CI)}$ and $F_{CI(LacI)}$ values, with $J_{aa'}$ for each reporter the only free parameter. All values are ± standard deviations.

dure to find values of $J_{aa'}$ that best reproduce the observed $F_{1(2)}$ values and were consistent with the various other *J* and *I* estimates (Figure 5A and B). These predicted $J_{aa'}$ values were subject to two deviations from simple expectation.

First, none of the $J_{aa'}$ estimates of Figure 5A and B conform precisely to the expected *J* value for the simple sum of the lengths of the a + a' arms, ranging from 0.19- to 1.4-fold the expected values of $J = 531$ nM for the 690 bp loop in the LacI looping reporters and $J = 616$ nM for the 600 bp loop in the CI looping reporters (Supplementary Figure S1). This suggests that the DNA arms are not sufficiently long to avoid effects of DNA stiffness. Instead, it appears that the orientations of each DNA helix as it exits the *OL*-CI-*OR* complex or the *Oid*-LacI-*O2* complex affect, positively or negatively, the likelihood of the DNA coming together again in a manner to suit formation of the other looping complex.

Second, and more difficult to understand, different $J_{aa'}$ values were needed to fit the data for the 2 and 20 kb spacings within the LacI or CI loop reporters (Figure 5A and B; note that $J_{aa'}$ estimates for the 50 kb reporters are poorly defined). Because the 300 + 390 bp a + a' loops in the LacI looping reporters and the 300 + 300 bp a + a' loops in the CI reporters are unchanged between the 2, 20 and 50 kb spacings, we expected that the $J_{aa'}$ estimate should be constant across the different spacings. However, fitting the data from three spacings simultaneously by constraining all the 300 + 390 $J_{aa'}$ values or all the 300 + 300 $J_{aa'}$ values to be the same was unable to produce good matches to the *F* values. In both the LacI and CI looping reporters, the $J_{aa'}$ fitted for the 2 kb spacing was roughly 3-fold higher than for the 20 kb spacing, suggesting more efficient looping assistance for the shorter distance. Although we cannot exclude that this difference in the fitted $J_{aa'}$ values results from errors in the estimates of the other parameters, the results suggest that the length of the internal s loop is somehow affecting the a + a' loop.

One possibility is that there is some stiffness of the 1400 bp internal loops that restricts how the DNA arms enter the internal DNA:protein complex, and which somehow propagates through the complex to affect the orientation of the DNA of the a and a' arms to enhance formation of the external DNA:protein complex. For example, stiffness in the 1400 bp loop might favor parallel or antiparallel orientations of the DNA at the internal bridging complex, which in turn could affect the ability of the outer loop to form. This implies that even a 1400 bp loop is not entirely in the entropic range. Another possibility is that differences in excluded volume effects caused by the internal 1400 bp loop

compared to the 20 kb loop may affect the formation of the outer loop.

These results highlight difficulties in making precise predictions of loop assistance effects *in vivo*, even in a reasonably well defined system. If the $J_{aa'}$ values were those expected for the sum of the a + a′ arms and were independent of the length of the internal loops, then the CI-assisted LacI looping efficiencies $F_{LacI(CI)}$ would be 0.86 and 0.64 for the 2 and 20 kb cases, rather than the 0.89 and 0.44 seen, while the LacI-assisted CI looping efficiencies $F_{CI(LacI)}$ would be 0.95 and 0.72 for the 2kb and 20 kb cases, rather than the 0.90 and 0.37 seen.

Enhancing DNA looping by reducing *I*: strengthening the looping interaction

Loop assistance can enhance long distance DNA looping by reducing the effective *J* between the interacting sites. Equation 1 (Figure 2A) shows that an alternative approach to increase looping is to increase the strength of the interaction itself, that is to decrease *I*.

We thus tried to increase the strength of the interactions between the lambda CI sites that provide loop assistance in the 50 kb LacI looping reporter. Our approach was to increase the number of CI operators to provide additional protein bridges between the distant sites. We placed the operators close together to try to minimize looping interactions between neighboring CI sites, which would compete with the long-range interactions. We optimized this strategy empirically, testing various combinations and spacings of CI operators for their ability to assist LacI looping (Supplementary Figure S3). Of these, a four operator binding site, *OL12*-42 bp-*OR12* (Figure 6A), gave the greatest assistance. At the 50 kb spacing, loop assistance with these *CIx4* sites, gave $F_{LacI(CI)} = 0.41 \pm 0.01$ (Figure 6B), compared with 0.22 ± 0.01 obtained with the standard *OL* and *OR* sites (Figure 4B).

We then used these enhanced CI looping elements in LacI looping reporters with 100 and 200 kb spacings (Figure 6B). At these sub-megabasepair distances, looping by LacI alone was essentially undetectable. However, the presence of CI and the *CIx4* sites gave $F_{LacI(CI)} = 0.26 \pm 0.01$ and 0.24 ± 0.01 for the 100 and 200 kb spacings, respectively (Figure 6B). Thus, a simple increase in the number of CI operators was able to drive efficient looping at up to at least 200 kb.

The two-loop model was used to analyse these results in order to extract an estimate of *I* for the looping *CIx4* element. We imagine that each operator pair (*OR1.OR2* or *OL1.OL2*) binds a CI tetramer, which can then interact with a non-neighboring tetramer. Thus, the four pairs of operators in the two *CIx4* sites should be able to interact to form four different single long-range loops or two different double long-range loops. Each *CIx4* element could therefore be considered a pair of looping elements separated by loop arms of 34 bp, and it would be possible to model the looping of the *CIx4*-LacI looping reporters as a three loop system (see below). However, it is difficult to predict the *J* values for the 34 bp loop arms. Instead, a simpler approach, and one that allows us to apply the 2-loop model, is to consider each *CIx4* element as a single unit, with looping between two *CIx4* elements characterized by a single value of *I*.

We made the assumption that $J_{aa'}$ for the 300 + 390 bp external loop in these reporters is the same as for the 20 kb *OL-OR* assisted reporter (226 ± 55 nM; Figure 5A). Fitting with the 2-loop model effectively uses the 50 kb *F* measurements, the I_{LacI} value and this $J_{aa'}$ value to extract a compatible estimate of I_{CIx4}. The obtained $I_{CIx4} = 3.6 \pm 2.7$ (Figure 6C) indicates that the *CIx4* interaction is substantially stronger than the $I_{CI} = 31.6 \pm 7.5$ estimate for the *OL-OR* interaction.

An extended *J* versus DNA distance relationship

The two-loop model can also be applied to the results of Figure 6B to extract estimates of *J* for the 100 and 200 kb spacings. Our previous analysis of single LacI loops allowed us to estimate the *J* versus *d* relationship for *d* ranging from 250 to 5600 bp (23). The weak but measurable unassisted looping by LacI or CI (Figure 5) add direct estimates of *J* ~6–11 nM and ~1–2 nM for 20 and 50 kb distances, respectively (Figure 7A). At the 100 and 200 kb spacings, unassisted LacI looping is too weak to measure *J* directly (Figure 6B). However, using the assisted *F* values at 100 and 200 kb, and the I_{LacI}, I_{CIx4} and $J_{aa'}$ estimates, it was possible with the 2-loop model to obtain $J_{asa'}$ estimates of ~0.5 nM for these distances (Figure 6C).

The 20 and 50 kb distance *J* estimates conform reasonably well to the power law obtained for the shorter distances (Figure 7A). A power law fit to all of the directly measured *J* values for distances up to 50 kb gives a slightly steeper decay of *J* with *d* than obtained from the shorter distances only, with an exponent of −1.14 (Figure 7A). The 100 and 200 kb *J* estimates (not used in the fit) align reasonably well with this power law.

If a DNA site is able to move freely in the cell, then one would expect the lowest possible *J* value to be equal to the concentration of a single molecule in the cell, ~1 nM in *E. coli*. Our lowest *J* estimates are close to this value (Figure 7A). The loop assistance approach should allow estimation of *J* at even greater distances. *J* values well below 1 nM would suggest restrictions to free movement of DNA.

These estimates of *J* ranging over almost three orders of magnitude of DNA distance, though not entirely smooth, do not show evidence of major discontinuities (Figure 7A). This is in accordance with other studies in bacteria, which most often show smooth decreases in interaction efficiency with distance (19,21). However, some exceptions to a smooth decay have been observed at fixed locations termed macrodomain boundaries that inhibit the interaction of λ attachment sites when these sites are placed either side of a boundary (21). We note that our *lacZ* reporter is located at the λ *attB* site at 17.4 centisomes and the distal LacI and CI sites are placed in the DNA further away from the origin, with the 200 kb sites at 21.7 centisomes (Supplementary Figure S5). Thus, our sites lie wholly within the right macrodomain (13–26 centisomes) (21) and do not cross identified macrodomain boundaries.

Model for loop assistance with more than two looping elements

Decreasing *I* by adding protein binding sites, as we have done with CI, may also be feasible for increasing LacI loop-

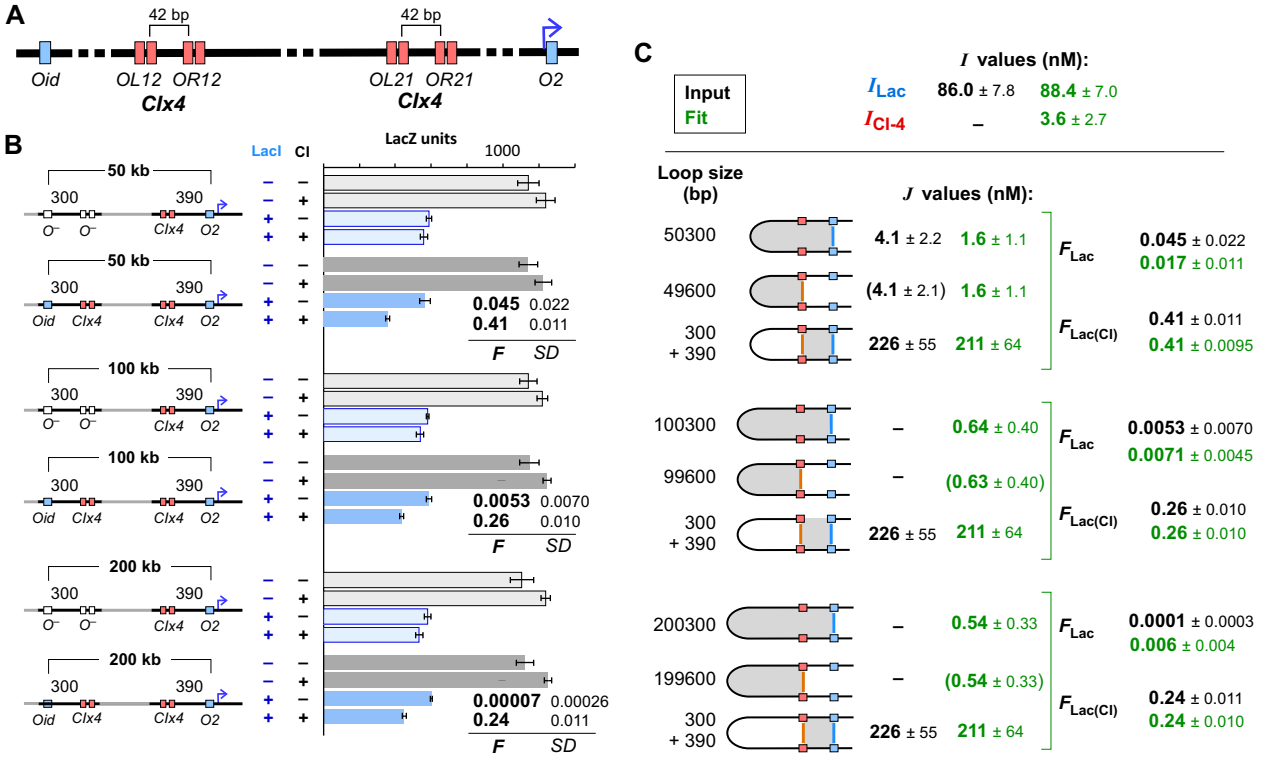

Figure 6. Loop assistance by four pairs of CI operators can drive looping to 200 kb. (**A**) The *CIx4* looping elements. (**B**) Looping by LacI between *Oid* and *O2* operators separated by 50, 100 or 200 kb is increased by the presence of CI and internally nested *CIx4* operators. Data presentation and analysis was as for Figure 4A. (**C**) Treatment of the *CIx4* interaction as a single looping element allows estimation of I_{CIx4} and $J_{asa'}$ values for the 100 and 200 kb distances by fitting of the observed F_{LacI} and $F_{LacI(CI)}$ values from (B) using the 2-loop model. Unassisted looping for the 100 and 200 kb reporters was too weak to reliably estimate $J_{asa'}$ directly. The parameter values in black are inputs to the fitting: the I value for LacI was obtained from the data in (B); the $J_{asa'}$ value for the 50 kb spacing was obtained directly from the analysis of the data in (B); the $J_{aa'}$ value used was that for the 300 + 390 bp aa' loop for the 20 kb CI OL/OR-assisted LacI reporter (Figure 5A) and was assumed to be the same for all three reporters. Parameter values in green/gray, including I_{CIx4} and $J_{asa'}$ values for the 100 and 200 kb distances, are fitted values that optimize the match between the observed and model-predicted F_{LacI} and $F_{LacI(CI)}$ values. J_s values were estimated from the corresponding $J_{asa'}$ value using the power law (values in parentheses). All values are ± standard deviations.

ing. However, we expect that the benefit of increasing the number of binding sites will become limited by short-range interactions within the extended looping elements. An alternative strategy to increase long-range looping efficiency without using strongly interacting looping elements would be to increase loop assistance by using more than two looping proteins. The 2-loop model can be readily extended to cases with more than two simple looping elements (Supplementary Figure S4 describes its application to a 3-loop system). We found that even with fairly moderate individual looping interactions ($I = 100$) and quite large distances between nested sites (500 bp), combining multiple looping elements should be able to drive DNA looping at very large distances (Figure 7B).

DISCUSSION

Toward predictable *invivo* DNA looping

A generalizable theory for protein-mediated DNA looping would aid understanding, prediction and manipulation of DNA looping in a variety of situations. Our simple framework separates the factors controlling DNA looping into two independent factors, I and J, allowing modular analysis and design of a variety of DNA looping situations, as long as the relevant I and J values can be reliably determined.

Our measurements of loop efficiencies for single loops ranging from 250 bp to 50 kb show that loop length is a strong determinant of J, with the trend described reasonably well by a simple power law. Our analysis supports the extension of this power law to 200 kb (Figure 7A). We note however that DNA segments of similar lengths can give quite different J values (Figure 7), and it is likely that differences in the activity of *in vivo* factors on different DNA sequences (e.g. the presence of DNA-bound proteins) affect J in an unpredictable way. The J versus d relationship may also be different for different regions of the *E. coli* chromosome or under different growth conditions and is likely to be different in different cell types. The LacI looping reporter system can be used to test the generality of the J versus d relationship in *E. coli* and should be adaptable to measure J in other bacteria. Use of a similar transcription-based measurement of J in eukaryotic systems using LacI seems possible if a clearly defined relationship between promoter activity and the fractional occupation of a binding site for LacI at the promoter can be established.

Direct determination of I is likely to be difficult for most DNA looping elements. Our ability to obtain I for LacI

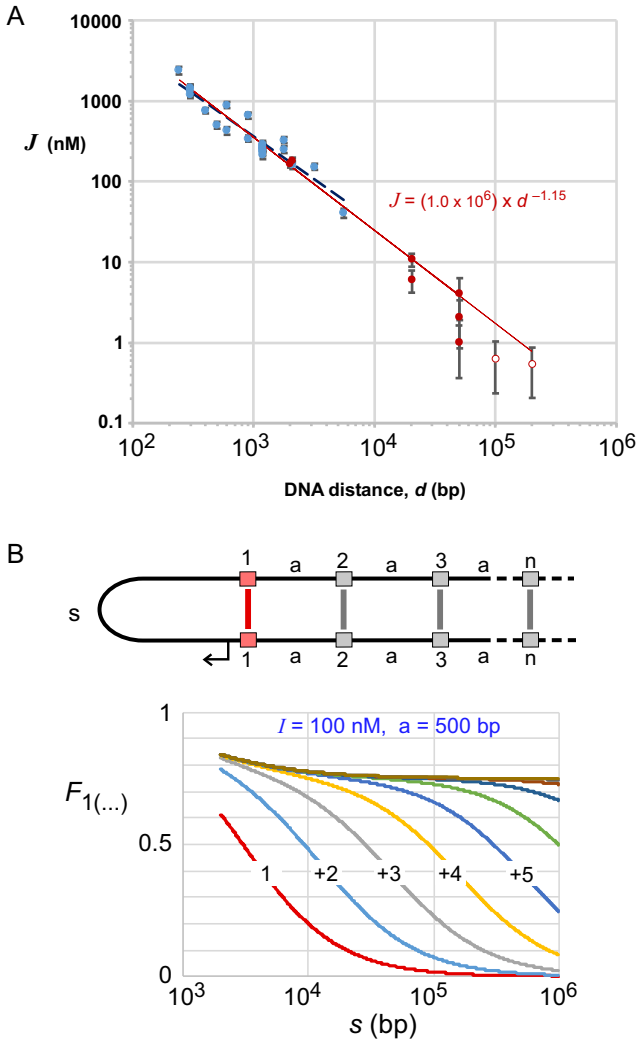

Figure 7. (**A**) Extended relationship between J and DNA distance d from 242 bp to 200 kb in *Escherichia coli* cells. Results from previously obtained data are shown in light shading, with their power law fit dashed (from Supplementary Figure S1). Filled darker shaded points are directly measured J values obtained from the reporter data of Figures 4 and 6. Unfilled points are the indirectly estimated J values from Figure 6. Each J value is an average of multiple fits; error bars are standard deviations. The power law fit to all the data up to 50 kb (filled circles) is shown as a solid line. (**B**) Enhancing looping with more than two nested looping elements. A multiple loop model was developed (Supplementary Material) to calculate the predicted effect of additional looping elements on the loop fraction of loop 1. The example shows the case of equally sized assisting loop arms $a = 500$ bp, and $I = 100$ nM for each of 9 looping elements, for different spacer distances, s. The power law in (A) was used to calculate J for each loop based on the sum of the lengths of the DNA segments in the loop.

and CI relies on numerous measurements of DNA-looping dependent promoter repression combined with substantial biochemical knowledge of the mechanisms and parameters for regulation by these well-characterized proteins. However, as we have shown for the *4xCI* looping elements, it is possible to estimate I indirectly through loop assistance (Figure 6). In this case, I for the internal looping element was determined from the measured increase in external LacI looping and estimates of $J_{asa'}$, J_s, $J_{aa'}$ and I_{LacI}, using Equation 2 (Figure 3A). Thus, measurement and optimization of I (for example by altering sites or protein concentration) can be achieved without knowledge of the looping mechanism.

Our theory had mixed success predicting the loop assistance effect of nested loops. Though we saw the expected strong loop assistance for large loops, the fitted $J_{aa'}$ values for the short DNA loops closed by LacI at one side and CI on the other were different from those expected from the simple sum of the arm lengths, and were not consistent between the 2 and 20 kb spacings (Figure 5). The model relies on the separation of I and J—the I value for a protein–DNA bridge and the J values of the surrounding loops should be independent of each other. The deviations from expectations may be due to a breakdown of this independence that can occur when the DNA distances are not in the entropic range i.e. the loops are not fully flexible. Alternatively, the model may be ignoring other kinds of interactions between nearby loops, such as excluded volume effects.

We note that our model for nested loops does not apply directly to alternating loops—where the loop formed by one pair of elements sequesters one of the elements for the other loop. We have demonstrated in *E. coli* that such arrangements can produce loop interference, with the formation of one loop inhibiting the other (30). Such loop interference is important, as it is the favored mechanism for insulator elements that inhibit enhancer–promoter contact (14,15). Further quantitative data will be needed to test whether our modeling approach can be adapted to alternating loops.

Efficiency and specificity in enhancer–promoter interaction

Our analysis offers insights into how loop assistance could provide efficiency and specificity for the very long-range interactions between enhancers and promoters in eukaryotic genomes.

The ability of an inactive promoter to be activated by contact with an enhancer implies that bound proteins or protein binding sites at the enhancer aid the binding or activity of the transcriptional apparatus at the promoter. It is important for specificity that these interactions between the enhancer factors and the basal promoter factors are weak (I is high) to ensure that activation by the enhancer is restricted to nearby promoters (those where the loop has a high J). These weak interactions would explain the ability of enhancers to activate non-cognate promoters when moved close to them (when J is high), as in enhancer trap assays (33). An ability to interact with nearby basal promoters probably also underlies the weak transcription and RNAP localization often found near enhancers (34).

Thus, when an enhancer is far from its cognate promoter, additional mechanisms are needed to make the interaction efficient. This can be achieved by placing additional, stronger looping elements near the enhancer and the promoter to provide loop assistance. The formation of the double-looped species is determined by the product of two factors (Figure 3A): (1) the propensity to form the *assisting* loop, given by J/I for that loop (e.g. J_s/I_2 in Figure 3A), and (2) the propensity to form the smaller a + a' loop, given by J/I for that loop (e.g. $J_{a+a'}/I_1$ in Figure 3A). Maximizing this product for a given long-range enhancer–promoter pairing, requires a strong interaction (low I) for the assisting loop, and a high J for the a + a' loop. We have shown that

multiple operators for a single looping protein can decrease *I* to levels that make it capable by itself of significant looping at a distance of 100 kb (Figure 2B). High *J* values for the a + a' loop should also be achievable by placing the looping elements close to the promoter and enhancer to keep the a + a' loop as small as possible.

The combination of these approaches should be capable of markedly increasing the range of efficient enhancer–promoter interactions. However, if the same loop assistance elements are shared by different enhancers and promoters, then this increased efficiency carries the risk of low enhancer–promoter specificity. Specificity could be maintained if different looping elements were used for each enhancer–promoter pair, but it is likely that there are many more enhancer–promoter pairs than there are specific looping elements. Our modeling indicates that use of a few moderately strong looping elements can provide efficient looping by combined loop assistance (Figure 7B). Thus, an alternative is to use a smaller set of different looping elements but in different combinations for each enhancer–promoter pair.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Sandeep Krishna and members of the Shearwin laboratory for discussions.

FUNDING

Australian National Health and Medical Research Council [GNT1100653]; Australian Research Council via a Discovery Early Career Researcher Award [DE150100091 to N.H.]; WH Elliott Fellowship in Biochemistry (to I.B.D.); Center for Models of Life (to K.S., I.B.D). Funding for open access charge: National Health and Medical Research Council [GNT1100653].
Conflict of interest statement. None declared.

REFERENCES

1. Schleif,R. (1992) DNA looping. *Annu. Rev. Biochem.*, **61**, 199–223.
2. Matthews,K.S. (1992) DNA looping. *Microbiol. Rev.*, **56**, 123–136.
3. Ptashne,M. (1986) Gene regulation by proteins acting nearby and at a distance. *Nature*, **322**, 697–701.
4. Vilar,J.M.G. and Leibler,S. (2003) DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.*, **331**, 981–989.
5. Bulger,M. and Groudine,M. (2011) Functional and mechanistic diversity of distal transcription enhancers. *Cell*, **144**, 327–339.
6. Dröge,P. and Müller-Hill,B. (2001) High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. *Bioessays*, **23**, 179–183.
7. Rippe,K. (2001) Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.*, **26**, 733–740.
8. Lettice,L.A., Heaney,S.J.H., Purdie,L.A., Li,L., de Beer,P., Oostra,B.A., Goode,D., Elgar,G., Hill,R.E. and de Graaff,E. (2003) A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.*, **12**, 1725–1735.
9. Herranz,D., Ambesi-Impiombato,A., Palomero,T., Schnell,S.A., Belver,L., Wendorff,A.A., Xu,L., Castillo-Martin,M., Llobet-Navas,D., Cordon-Cardo,C. *et al.* (2014) A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat. Med.*, **20**, 1130–1137.
10. Gibcus,J.H. and Dekker,J. (2013) The Hierarchy of the 3D Genome. *Mol. Cell*, **49**, 773–782.
11. Sexton,T., Yaffe,E., Kenigsberg,E., Bantignies,F., Leblanc,B., Hoichman,M., Parrinello,H., Tanay,A. and Cavalli,G. (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*, **148**, 458–472.
12. Le,T.B., Imakaev,M.V., Mirny,L.A. and Laub,M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science*, **342**, 731–734.
13. Cagliero,C., Grand,R.S., Jones,M.B., Jin,D.J. and O'Sullivan,J.M. (2013) Genome conformation capture reveals that the Escherichia coli chromosome is organized by replication and transcription. *Nucleic Acids Res.*, **41**, 6058–6071.
14. Rao,S.S., Huntley,M.H., Durand,N.C., Stamenova,E.K., Bochkov,I.D., Robinson,J.T., Sanborn,A.L., Machol,I., Omer,A.D., Lander,E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
15. Phillips-Cremins,J.E. and Corces,V.G. (2013) Chromatin insulators: linking genome organization to cellular function. *Mol. Cell*, **50**, 461–474.
16. Sanborn,A.L., Rao,S.S., Huang,S.C., Durand,N.C., Huntley,M.H., Jewett,A.I., Bochkov,I.D., Chinnappan,D., Cutkosky,A., Li,J. *et al.* (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6456–E6465.
17. Muller,J., Oehler,S. and Muller-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.
18. Dandanell,G., Norris,K. and Hammer,K. (1991) Long-distance deoR regulation of gene expression in Escherichia coli. *Ann. N. Y. Acad. Sci.*, **646**, 19–30.
19. Higgins,N.P., Yang,X., Fu,Q. and Roth,J.R. (1996) Surveying a supercoil domain by using the gamma delta resolution system in Salmonella typhimurium. *J. Bacteriol.*, **178**, 2825–2835.
20. Ringrose,L., Chabanis,S., Angrand,P.O., Woodroofe,C. and Stewart,A.F. (1999) Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *EMBO J.*, **18**, 6630–6641.
21. Valens,M., Penaud,S., Rossignol,M., Cornet,F. and Boccard,F. (2004) Macrodomain organization of the Escherichia coli chromosome. *EMBO J.*, **23**, 4330–4341.
22. Cui,L., Murchland,I., Shearwin,K.E. and Dodd,I.B. (2013) Enhancer-like long-range transcriptional activation by λ CI-mediated DNA looping. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 2922–2927.
23. Priest,D.G., Cui,L., Kumar,S., Dunlap,D.D., Dodd,I.B. and Shearwin,K.E. (2014) Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and λ repressors. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 349–354.
24. Mahmoudi,T., Katsani,K.R. and Verrijzer,C.P. (2002) GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J.*, **21**, 1775–1781.
25. Nolis,I.K., McKay,D.J., Mantouvalou,E., Lomvardas,S., Merika,M. and Thanos,D. (2009) Transcription factors mediate long-range enhancer-promoter interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 20222–20227.
26. Calhoun,V.C., Stathopoulos,A. and Levine,M. (2002) Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 9243–9247.
27. Kwon,D., Mucci,D., Langlais,K.K., Americo,J.L., DeVido,S.K., Cheng,Y. and Kassis,J.A. (2009) Enhancer-promoter communication at the Drosophila engrailed locus. *Development*, **136**, 3067–3075.
28. Deng,W., Rupon,J.W., Krivega,I., Breda,L., Motta,I., Jahn,K.S., Reik,A., Gregory,P.D., Rivella,S., Dean,A. *et al.* (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell*, **158**, 849–860.
29. Fujioka,M., Mistry,H., Schedl,P. and Jaynes,J.B. (2016) Determinants of Chromosome Architecture: Insulator Pairing in cis and in trans. *PLoS Genet.*, **12**, e1005889.
30. Priest,D.G., Kumar,S., Yan,Y., Dunlap,D.D., Dodd,I.B. and Shearwin,K.E. (2014) Quantitation of interactions between two DNA

loops demonstrates loop domain insulation in E. coli cells. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E4449–E4457.
31. Garcia,H.G. and Phillips,R. (2011) Quantitative dissection of the simple repression input-output function. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 12173–12178.
32. Rusinova,E., Ross,J.B., Laue,T.M., Sowers,L.C. and Senear,D.F. (1997) Linkage between operator binding and dimer to octamer self-assembly of bacteriophage lambda cI repressor. *Biochemistry*, **36**, 12994–13003.
33. Ruf,S., Symmons,O., Uslu,V.V., Dolle,D., Hot,C., Ettwiller,L. and Spitz,F. (2011) Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat. G enet.*, **43**, 379–386.
34. Li,W., Lam,M.T.Y. and Notani,D. (2014) Enhancer RNAs. *Cell Cycle*, **13**, 3151–3152.