The loopometer: a quantitative in vivo assay for DNA-looping proteins

Nan Hao, Adrienne E. Sullivan, Keith E. Shearwin and Ian B. Dodd

ABSTRACT

Proteins that can bring together separate DNA sites, either on the same or on different DNA molecules, are critical for a variety of DNA-based processes. However, there are no general and technically simple assays to detect proteins capable of DNA looping in vivo nor to quantitate their in vivo looping efficiency. Here, we develop a quantitative in vivo assay for DNA-looping proteins in Escherichia coli that requires only basic DNA cloning techniques and a LacZ assay. The assay is based on loop assistance, where two binding sites for the candidate looping protein are inserted internally to a pair of operators for the E. coli Lac repressor. DNA looping between the sites shortens the effective distance between the lac operators, increasing LacI looping and strengthening its repression of a lacZ reporter gene. Analysis based on a general model for loop assistance enables quantitation of the strength of looping conferred by the protein and its binding sites. We use this ‘loopometer’ assay to measure DNA looping for a variety of bacterial and phage proteins.

INTRODUCTION

Proteins that bind to specific DNA sites and are able to interact with each other to bring together separate DNA molecules or to form loops within the same DNA molecule are critical in essential DNA processes, such as transcription, replication, recombination and DNA organization (1–3). The primary role of the interaction can be to affect the function of the proteins at one or both sites. DNA looping provides cooperative binding over large distances, increasing site occupancy by the proteins (1,2,4), and can be used to target catalytic activities in transcriptional control, such as in activation of σ54 promoters (5) or eukaryotic enhancers (6,7). Bridging interactions between replicases bound to separate plasmids can inhibit replication initiation to provide plasmid copy number control (8,9).

Alternatively, the primary role of the interaction can be the juxtaposition of the two DNA segments. Promoters can be repressed by being enclosed in small DNA loops (10,11), and the large chromatin loops formed by the interaction between insulator elements partition eukaryotic chromosomes into topological domains that are important for enhancer–promoter specificity (12). In many cases, the interaction combines both roles. Site-specific recombinases must juxtapose their DNA sites for correct recombination and the interactions in the synapsed complex also activate the recombinase catalytic steps (13). The long-range interaction between the OL- and OR-binding sites for the phage λ CI repressor increases CI binding at these sites, improving repression of the lytic promoters (14,15), while the DNA loop formed also juxtaposes a distal UP element near to the PRM promoter where it can activate lysogenic transcription (16,17).

Despite the importance of DNA looping and bridging, a general method for identifying proteins capable of bringing DNA together in vivo is not currently available. For most proteins, an in vivo looping capability is first implicated when the function of a protein at one binding site is affected by a distal binding site. For example, in the first described example of DNA looping by a transcription factor, looping by AraC was suspected because deletion of a binding site 270 bp upstream of the araBAD promoter affected AraC regulation of the promoter (18). To further support DNA looping, additional tests, usually specific to each system, are needed to rule out an independent function of the distal site. Helical phasing experiments, where the functional interaction between the sites is sensitive to insertion of nonintegral DNA turns in the intervening DNA, are often used to confirm DNA looping in vivo (18–21). However, helical phasing sensitivity is not present for long DNA loops (>500 bp), or for trans interactions, and may be minimized by protein flexibility. An alternative route to identifying DNA-looping proteins in vivo is provided by ligation-based proximity assays such as 3C (22). These techniques have revealed vast numbers of DNA loops in eukaryotic and bacterial genomes (23,24). However, additional approaches are needed to identify the proteins responsible for these loops.
Thus, confirmation of the DNA looping or bridging activity of a specific protein often requires in vitro techniques such as DNA footprinting (25), electrophoretic mobility shift (26), enhancement of ligation (9) or single-molecule methods such as electron or atomic force microscopy (14,27) and tethered particle motion (28). However, these in vitro approaches need purified active protein and often require specialized expertise and equipment.

In addition, an important limitation of all these approaches is that they do not reveal the in vivo strength of the protein-mediated DNA looping, that is the fraction of time that the sites spend in direct contact in cells. Extraction of looping frequencies from functional in vivo assays has been possible only for highly characterized regulatory proteins, such as the Lac and CI repressors (29). Ligation-mediated proximity assays and techniques such as Dam-C can provide relative but not absolute frequencies of close proximity (30). Absolute proximity measures can be obtained by microscopic imaging of intact cells (31,32), but none of these proximity approaches is able to confirm direct protein-mediated contact. In vitro studies can quantitate looping; however, these measurements are made under conditions that may not reflect in vivo looping. Knowing how often the loops form in vivo provides critical information about the interplay between DNA looping and function, while identification of DNA-looping proteins capable of directing strong DNA looping will also aid in more effective manipulation of DNA looping in cells.

Here, we report the development of an in vivo assay for DNA-looping proteins in Escherichia coli that does not depend on detailed knowledge of the protein’s function and requires only basic DNA cloning techniques and a LacZ assay. Furthermore, the assay provides a quantitative measurement of the strength of looping exerted by the protein and its binding sites. The assay is based on loop assistance, where the formation of one DNA loop assists the formation of another DNA loop (Figure 1A). When a pair of sites for one DNA-looping protein is nested within a pair of sites for another DNA-looping protein, DNA looping by the internal protein assists looping by the external protein by bringing its binding sites closer together (33–36). In the loopometer, the external loop is formed by two operators for the E. coli Lac repressor (LacI) that repress the expression of a lacZ reporter gene (Figure 1B). The proximal operator directly represses the promoter and the distal operator cooperates by DNA looping to increase LacI occupation of the proximal operator. This allows LacI looping to be detected and measured by the increase in repression in the presence of the distal operator (33,34). To assay DNA looping by a candidate protein, a pair of binding sites is inserted internal to the LacI loop and the protein is expressed. DNA looping by the candidate protein improves LacI looping by shortening the effective distance between the lac operators, detected by increased LacI repression of the reporter (Figure 1B). Analysis of the measurements using a general model for loop assistance (35) enables the looping strength of the protein and its binding sites to be estimated. We validate the assay using the λ CI protein and test a variety of bacterial and phage proteins for DNA-looping activity.

**MATERIALS AND METHODS**

### General strains and media

All reporter strains are derivatives of E4643, itself a derivative of BW30270 MG1655 rph+ (CGSC7925) with the lacIZYA (MG1655:360,527–366,797) region removed (29). Plasmid replication of pLOM2–500 and pCYMR, which are derived from the CRIM plasmids of Haldimann and Wanner (37), is dependent on the R6Kγ π replicate protein Pir. We thus maintain these plasmids in the Pir-expressing (pir+) strain E4644 = EC100D (Epicentre) = F mcra Δ(mrr-hsdRMS-mcrBC) q80d(lacZΔM15) ΔlacX74 recA1 endA1 araD139 Δ ara-leu)7697 galU galK rpsL strB supG5 pir+ (DHFR).

LB (1% Bacto-tryptone, 1% NaCl, and 0.5% yeast extract, pH 7.0) is used for routine growth of strains. M9MM-glyc = 1 × M9 salts, 2 mM MgSO4, 0.1 mM CaCl2, 0.01 mM (NH4)2SO4, 6H2O, and 0.4% glycerol [10 × M9 salts = 67.8 g of Na2HPO4, 30.0 g of KH2PO4, 10 g NH4Cl and 5 g NaCl/1 L H2O] were used for LacZ assays. Antibiotics used were (μg/ml for selection of integrated copy/plasmid copies): Ap, ampicillin (-/100); Km, chloramphenicol (20/30); Km, kanamycin (20/50); Sp, spectinomycin (20/50); Tc, tetracycline (3/20). A 100 mM stock of cuminic acid (CA; Sigma-Aldrich Cat. 268402) was prepared in ethanol.

### Constructing loopometer strains

**Cloning protein binding sites into pLOM2-500.** The restriction sites flanking sites 1 and 2 in pLOM2–500 (Figure 2A and Supplementary Figure S1A) allow various strategies for cloning candidate binding sites at these positions. We use Gibson isothermal assembly (38), either by 2-fragment assemblies with a single DNA fragment comprising both inserts and attP, (PCR-generated or purchased commercially) or by 4-fragment assemblies of two polymerase chain reaction (PCR)-generated insert fragments, an attP fragment and digested and isolated plasmid backbone. pLOM2–500 plasmids were maintained in E4644 (LB + Km, 30 μg/ml). Supplementary Figure S2B shows the detailed structure of pLOM2–500 and the reporter after its integration into the loopometer landing pad. Primers #2238 (TGCCGACGCTCATGTATGTG) and #787 (ACCGGTTAATTAACGGCACCACCGAAG) were used for sequencing the site 1–site 2 region.

**Cloning protein genes into pCYMR.** The pCYMR plasmids (Figure 2A and Supplementary Figure S3A) contain a multiple cloning site (SpeI.ApaI.l.XmaI.pVulI.AflII.AccIII.NarI) located downstream of the CymR-binding site (cyMO), followed by transcription terminators. We generally used 2-fragment Gibson assembly to insert PCR-generated fragments of protein genes into the pVulI site. Sequences for translation of the protein message need to be included in the insert. pCYMR plasmids were maintained in E4644 (LB + Km, 50 μg/ml). Inserted sequences were confirmed using primers #2037 (CATGATCCATAGATCTCTTCTC) and #787 (ACCGGTTAATTAACCGCACCACCGAAG). Three variants of pCYMR were made with altered activity.
of the expression promoter: pCYMR.1 (standard vector), pCYMR.4 (higher expression) and pCYMR.6 (lower expression; Supplementary Figure S3). The detailed structure of the pCYMR expression plasmids and the cloning junctions for the genes inserted are given in Supplementary Figures S3 and S4.

Integrating pLOM-500. The pLOM-500 reporter plasmid containing the protein-binding sites is separately integrated into three strains: AH6112 (lacI'Oid'), AH6113 (lacI'Oid') and AH6114 (lacI'Oid'). These strains differ in whether Oid is present at the distal site in the loopometer landing pad (Oid') or absent (Oid; Figure 2, sequences in Supplementary Figure S2D), and whether the pIT3-SH.lacl-rev plasmid (lacI'; Supplementary Figure S2E) or the empty vector pIT3-SH (lacI') is integrated at attPHK022. Strains AH6112, AH6113 and AH6114 also contain the pINTs λ Int helper plasmid (37), which express λ Int at temperatures above 30°C and carry a temperature-sensitive pSC101 repA101 gene needed for its maintenance. For transformation and integration, these cells are grown at 30°C in LB + Ap, 100 μg/ml to OD600 ~0.2–0.4, transferred to 39°C for 20 min, before being made competent and transformed with pLOM-500 vectors by the TSS method (39), with selection for integrants on LB + Cm, 20 μg/ml plates at 37°C. Purified colonies are tested for correct pLOM-500 integration by PCR (40) with a mix of primers #462 (ATGACAGAGGCAGGGAGTGG), #1729 (TTGTGCTTCTCTGGAGTGCG), #1685 (GTATCCCCACTCATTCGGGGCTC), and #1686 (CCAGCCATACGTACGACG), which give 351, 841, 470 and 727 bp attP, attB, attL and attR fragments, respectively.

Integrating pCYMR. The three resulting strains are then transformed with the pAH121φ21 Int helper plasmid (37), with outgrowth in LB at 30°C and selection on LB + Ap, 100 μg/ml plates at 30°C. Integration competent cells of these strains are then made and transformed (as above) with empty pCYMR or the pCYMR plasmid containing the candidate protein gene, with selection on LB + Km, 20 μg/ml plates at 37°C. Purified colonies are tested for correct integration by PCR (40) with a mix of primers #465 (GGGAATTCTTGAAGACG), #467 (ACTTAACGGCTTGACATGG), #871 (ATCGCCTGTATGAACCTG) and #2404 (GCAGCCTAACAAAAAACAACG), which give 1226, 420, 621 and 1025 bp attP, attB, attL and attR fragments, respectively. Alternative methods for expressing the candidate protein may be used, as long as they do not re-
quire selection for chloramphenicol (the reporter module) or spectinomycin (the LacI expression module), and are not under LacI control.

LacZ assays

The standard loopometer assay involves LacZ (β-galactosidase) assay of six strains, all with the same protein-binding sites inserted at loopometer sites 1 and 2. The strains have three different combinations of lacI and Oid: lacI Oid−, lacI Oid′ and lacI Oid−, and carry either the integrated empty pCYMR vector or the pCYMR vector expressing the relevant protein.

Strains were grown and kinetic LacZ assays were performed in 96-well flat-bottomed microtiter plates. Six fresh colonies of each strain on LB + Cm, 20 μg/ml plates (except for ID1292 and ID1293, which were selected with Km20) were picked with a 1000 μl micropipette tip and resuspended in 40 μl M9MM + Cm20. About 5 μl of each resuspension was added to 95 μl M9MM+ Cm20, 20 μg/ml + CA in a microtiter plate, sealed and incubated with shaking at 37°C overnight. These cultures were then diluted 5 μl into 95 μl fresh medium and incubated for ~3 h. The five control strains were grown in 0 and 128 μM CA, while the lacI+ Oid+ strains were grown on 0, 8, 16, 32, and 128 μM CA (see Figure 2B).

Cultures were grown to OD600 0.3–0.8, measured using a Labsystems Multiskan Ascent plate reader with a 620 nm filter. The OD620 values were converted to OD600 (1 cm path length) values using an empirically derived relationship. For the assay, 5 μl of culture was added to 195 μl warmed assay buffer in a fresh microtiter plate, consisting of 88 μl TZ8 buffer, 60 μl 4 mg/ml o-nitrophenyl-β-D galactoside (Sigma-Aldrich Cat. N1127) in TZ8, 2 μl 10 mg/ml chicken egg white lysozyme (Sigma-Aldrich Cat. L6876, 40,000 units/ml) in TZ8, 4 μl 20 mg/ml polymyxin B (Sigma-Aldrich Cat. P-4932) in H2O, 6 μl H2O and 35 μl M9MM. TZ8 buffer is 100 mM Tris-Cl, pH 8.0, 1 mM MgSO4, 10 mM KC1. The plate was incubated at 28°C in the plate reader, with OD414 readings taken every 2 min for 1 h. Enzyme activity was determined as the slope of the line of best fit of OD414 versus time (readings with OD414 > 2.5 were ignored). LacZ units were calculated as 200,000 x (OD414/min)/(OD600 x 5 μl). The improved linearity compared with our previous assay (41) is examined in Supplementary Figure S5.

Quantitation of looping

\( F_{L(X)} \), the fractional LacI looping in the presence of the candidate protein, is obtained from the LacZ activities of the lacI Oid− and lacI Oid′ reporters with the expression of the protein, combined with the background LacZ activity of strain ID1285, using Equation (1) (Figure 4A).

Estimation of \( I_L / I_X \) from this \( F_{L(X)} \) value uses Equation (2) (Figure 5B), which requires additional \( F_L \) measurements to estimate the weights \( p, q \) and \( r \). The derivation of Equation (2) is based on our statistical-mechanical model for loop assistance (Figure 5; 35). The weight \( p \) is obtained from the \( F_L \) measurement of LacI looping of the lacI Oid− and lacI Oid′ versions of the loopometer in the presence of empty pCYMR. The weight \( q \) is estimated from assays of the lacI Oid− and lacI Oid′ strains ID1290 and ID1291, in which Oid and Plac O2 are separated by a b′ sequence similar to that between sites 1 and 2 (b, Figure 5B). The weight \( r \) is estimated from assays of the lacI Oid− and lacI Oid′ strains ID1292 and ID1293, in which Oid and Plac O2 are separated by a concatenation of the 500 bp Oid-site 1 and site 2-Plac O2 a and c arms separated by a 45 bp λ OR21 segment (Figure 5B).

Availabilty

Plasmids pAS0007 (= pLOM2–500 with ϕKO2 operators: Addgene # 164868), pID1072 (= pCYMR-1 with ϕKO2 cB: Addgene # 164869) and plasmid pAH121 (Addgene #164901), as well as strains AH6112–4, ID1285, and ID1290–3 will be deposited in Addgene.

Details of system construction

pLOM plasmids. Plasmid pLOM1–500 is an earlier version of pLOM2–500 that does not carry flanking restriction sites on both sides of sites 1 and 2, and was used for reporters with some binding sites (Supplementary Figure S1B). The sequences of protein binding sites cloned into these vectors are listed in Supplementary Figure S1.

pCYMR plasmids. The CymR repressed promoter in these vectors was derived from the T5 promoter, cymO sequence of pNEW (42), which includes an unwanted overlapping lac operator and also displayed leaky CymR repression. We randomized the sequences between the -35 and -10 hexamers (including the first -10 basepair) to remove this operator and alter promoter activity to generate the pCYMR-1, pCYMR-4 and pCYMR-6 variants (Supplementary Figure S3).

Loopometer recipient strains. The reporter landing pad present in strains AH6112, AH6113 and AH6114 was constructed in two steps by recombineering (Supplementary Figure S2A) using pSIM6 (43). First, a PCR fragment containing a kanamycin resistance gene (KmR) and a distal portion of the lacZ gene was used to replace attB, in the E4643 chromosome to give strain AH6101 (Supplementary Figure S2A). Second, after reintroduction of pSIM6, a PCR fragment carrying Oid− (or Oid′), attBλ, PlacUV5·lacO2 and the proximal portion of the lacZ(O2−) gene (Supplementary Figure S2A) was used to replace KmR, with screening for lacZ′ on LB + X-gal (20 μg/ml) plates. These two strains were transformed with the pAH69 HK022 Int helper plasmid (37) and pIT3-SH.lacZ.rev (Supplementary Figure S2E), or the empty vector pIT3-SH were then integrated at attPHK022. The resulting strains were transformed with the pINTTs λ integrase helper plasmid (37) to give AH6112 (lacI Oid′), AH6113 (lacI Oid−) and AH6114 (lacI Oid−).
cassette flanked by FRT sites, followed by FLP-mediated removal of the KmR gene using the pE.FLP plasmid (40), resulting in an in-frame deletion of almost the entire deoR gene.

The Plac strain for measuring reporter background. A DNA segment spanning the leftward FRT site, chloramphenicol resistance gene (CmR) and PlacUV5 was obtained from AH6114 carrying an integrated pLOM2–500 vector (Supplementary Figure S2B) and cloned into a plasmid, where mutations were introduced into the –35 and –10 sequences of PlacUV5 (Supplementary Figure S2D). A PCR fragment (Supplementary Figure S2C) from this Plac plasmid was used for recombineering into an AH6114 (pLOM2–500) strain in which the sequences between the FRT sites (Supplementary Figure S2B) had been removed by treatment with pE.FLP (40), allowing selection for the CmR gene in the insert. The resulting Plac strain, ID1285 is lacI' Oid', with matS null sequences (Supplementary Figure S1) at sites 1 and 2.

Strains for calibration of LacI looping. The DNA sequences for construction of b' and ac strains (Supplementary Figure S6) were assembled in plasmids, which were then used to generate PCR fragments (Supplementary Figure S6) for recombineering into the chromosome of AH6101 (Supplementary Figure S2A), with selection for CmR (b') or screening for lacZ' (ac). Successive integration of pIT3-SH.lacI-rev and empty pCYMR gave b' strains ID1290/ID1291 (Oid'/Oid') and ac strains ID1292/ID1293 (Oid'/Oid'). The b' segment shares 1870 bp with the loopometer b segment, with divergent sequences at each end near Oid and Plac O2. The ac strain carries a 45 bp λ OR21 insert between the fused sites 1 and 2.

Extended reporters were made by insertion of plasmid pID1302 into the attB3 site of the b' reporters ID1290/ID1291, and full-length lacI' Oid' reporters (abc) carrying the matS null site (Supplementary Figure S1) at both sites 1 and 2. Plasmid pID1302 (Supplementary Figure S6B) carries the attP3 site, the pir-dependent R6Kγ origin, a tetracycline resistance gene (TcR) and ~6 kb of spacer sequence derived from the E. coli ftsK and me genes. The int3 gene and the Int3 attachment sites are from the collection of Yang et al. (44). Integration of pID1302 (Supplementary Figure S6B) was mediated by expression of Int3 from the helper plasmid pAH6046, which was derived from pINTTs (37) by replacing λ int with the int3 gene.

Sequences of strains and plasmids are available on request.

RESULTS AND DISCUSSION

Procedure for testing DNA looping by a candidate protein

The standard assay uses a looping reporter and a separate module for the controlled expression of the protein, both of which are integrated into the bacterial chromosome (Figure 2A).

The looping reporters are constructed by insertion of two potential binding sites for the candidate protein at sites 1 and 2 on either side of the bacteriophage λ attP attachment site (attP3) in plasmid pLOM2–500, using standard cloning techniques (‘Materials and Methods’ section). The resulting plasmid is then integrated into the chromosome of three different E. coli strains by recombination into an attB3 site on a specially constructed ‘landing pad’ (Figure 2A). Integration is catalyzed by λ integrase expressed from a separate helper plasmid (37). Integration results in the two binding sites being located ~2 kb apart, separated by the inactive plasmid replication origin and the chloramphenicol resistance gene, with the proximal binding site located ~500 bp upstream of a lacO2 operator controlling a promoter (PlacUV5) for a lacZ reporter gene. Integration into strain AH6114 produces the intact loop reporter, which has a second lac operator, Oid, located ~500 bp upstream of the distal binding site, and a lacI gene elsewhere on the chromosome. The AH6112 and AH6113 strains lack either lacI or Oid, respectively, and serve as controls.

For expression of the candidate looping protein, we routinely use a chemically inducible expression system based on the CymR repressor from Pseudomonas putida F1 (42), which is inactivated by added CA. This allows testing of a range of protein concentrations, which may be important if high protein levels give submaximal looping, as seen for LacI (29), or are toxic. The gene for the candidate looping protein is inserted into the pCYMR plasmid, and the resulting plasmid is transformed into the three reporter strains previously transformed with the phage ϕ21 integrase helper plasmid (37), for catalyzing integration at the ϕ21 attachment site on the bacterial chromosome (attB21) (Figure 2A).

Our standard pCYMR.1 vector gives a low level of uninduced ‘leak’, with strong induction by CA (Supplementary Figure S3B). Nevertheless, we generally also integrate an empty pCYMR.1 plasmid into the reporter strains to provide control strains with no candidate protein.

Once the strains have been constructed, LacZ assays are done to test whether CA-induced expression of the candidate protein gives an Oid-dependent increase in LacI repression of the lacZ gene (Figure 1B). We use a microtiter plate-based version of the basic Miller LacZ assay (‘Materials and Methods’ section); however, standard Miller assays are adequate.

Validation of the assay with λ CI

We tested the assay with the λ CI repressor (45). Binding of CI dimers to individual operators via the N-terminal domains, and further association of these dimers to tetramers and octamers mediated by the C-terminal domain (46,47), is able to produce a variety of DNA loops. The natural cooperative binding of two CI dimers to pairs of adjacent operators at its OR- and OL-binding sites is due to mini-loops, with the operators, which are spaced ~2 DNA turns apart, being ‘looped’ by a CI tetramer (17,48). DNA looping has also been seen between single operators spaced 5 or 6 DNA turns apart (25). Longer DNA loops—beyond 2 kb in vitro and in vivo—have been observed for the interaction of two tetramer-bound sites, resulting in four operators linked by a CI octamer (14,17,28). The natural three-operator groupings at OL and OR are capable of even stronger looping (28,35), presumably because the two operators not bound by the octamer are bridged by an additional CI tetramer (49) (Figure 2B). We have shown that the full OL and OR

Downloaded from https://academic.oup.com/nar/article/49/7/e39/6123373 by University of Adelaide user on 22 July 2021
Figure 2. The loopometer system. (A) The basic procedure for the construction of the looping reporter strains. Protein-binding sites are cloned into the pLOM2–500 plasmid and the resulting plasmid is inserted by integrase into a reporter landing pad in the *Escherichia coli* chromosome, causing the two sites to be nested inside a pair of *lac* operators, *Oid* and *O2*, which control the PlacUV5 promoter for a *lacZ* reporter gene. LacI is supplied by a PlacI* lacI*+ gene inserted elsewhere in the chromosome. The pLOM2–500 plasmid is also integrated into control strains lacking *lacI*+ or *Oid*. The gene for the candidate protein is cloned into the pCYMR expression vector, under the control of the CymR repressor that can be inactivated by cumic acid (CA). This expression vector or the empty control is inserted by integrase-mediated recombination into the chromosome of the three strains carrying pLOM2–500. (B) LacZ assay results for *Cl* and *Cl*– reporter DNA looping is indicated by a LacI-dependent, Oid-dependent, candidate protein-dependent decrease in units. Error bars are 95% confidence limits (Student’s *t*, *n* = 6). The insert shows DNA looping by a *Cl* octamer and tetramer at the native *OL* and *OR* sites. (C) Graphic display of results for *Cl* and *Cl*– reporter DNA looping is indicated by a LacI-dependent, Oid-dependent, candidate protein-dependent decrease in units. Error bars are 95% confidence limits (Student’s *t*, *n* = 6). The insert shows DNA looping by a *Cl* octamer and tetramer at the native *OL* and *OR* sites.
sequences can form a 20 kb loop with \( \sim 16\% \) efficiency in vivo (35).

Figure 2B shows the results of CA-controlled expression of \( \lambda \) CI protein with \( \lambda \) OR and OL inserted at sites 1 and 2 in the reporter. In the lacI background, the reporter is unpressed, giving \( \sim 3000 \) LacZ units, and this activity is unaffected by the CI expression module, with or without full induction by CA (128 \( \mu \)M). In the lacI background but in the absence of the upstream Oid operator, the reporter is repressed \( \sim 5\) -fold due to weak binding of LacI to the proximal O2 site. Again, this activity is unaffected by CI expression. These controls check that any binding of the candidate protein to its sites does not directly affect the promoter or affect LacI repression in the absence of the upstream Oid. In the presence of Oid and the absence of CI, LacI repression of the reporter increases \( \sim 4\)-fold. This loop-dependent repression effect is due to the strong Oid site binding a LacI tetramer and thus fixing the free DNA-binding domain of the tetramer at the end of a 3 kb tether attached to O2. At this distance, the effective concentration of the LacI DNA-binding domain seen by O2 is substantially higher than the concentration of the LacI DNA-binding domains of free LacI tetramers, resulting in frequent DNA looping and increased occupation of O2.

In the presence of CI, repression in the lacI Oid\( ^{+} \) reporter is increased even further (Figure 2B). Given the results of the control strains, this increased repression can be interpreted as DNA looping by CI between OL and OR that, at least some of the time, shortens the distance between Oid and O2 to increase their relative concentration and increase LacI looping (Figure 1).

We generally test a range of induction levels for the candidate protein, as DNA looping generally has a concentration optimum (29,34). Too low protein concentration gives insufficient occupation of the DNA sites; too high concentration causes loop blockage, where the sites are occupied by higher order multimers that are unable to interact further (e.g. LacI tetramers at both lac operators).

To test whether the assay could display differences in the strength of looping, we compared \( \lambda \) CI looping of the native 3-operator OL and OR sites with looping by 2-operator or single operator sites. The equivalent assay to Figure 2B was done for reporters with these sites, and a subset of the results for all three combinations is plotted in Figure 2C. These plots are normalized to the average units obtained for each Oid reporter \((\pm \text{CI}, 0/128 \ \mu \text{M CA})\) to account for day-to-day variation in the LacZ assay.

As expected, the single operator OL1-OR1 reporter displayed weaker looping than the multioptimer combinations, with CI expression giving less assistance to LacI repression at all induction levels. We note that this is the first time that looping has been observed for single CI operators spaced more than a few DNA turns apart. Looping by the 2-operator OL12-OR12 combination was also weaker than the 3-operator pair at most CI induction levels. Interestingly, the 1- and 2-operator combinations gave maximal looping at low CI and less looping as CI levels increased. Formally, this can be explained as being due to the formation of higher order looping-incompetent CI multimers at each site. However, we were surprised to see this effect for \( \lambda \) CI, as it implies that CI octamers are forming independently at 1- or 2-operator sites. This effect was not seen for the 3-operator OL and OR sites, presumably because loop blockage is counterbalanced by loop-promoting bridging between the third operators (49).

Having validated the assay, we then used it to test other known or suspected DNA-looping proteins.

DeoR

The E. coli DeoR protein utilizes DNA looping in its repression of an operon for catabolism of nucleosides. DeoR binds to single operators \((\text{O1 and O2})\) spaced \( \sim 600 \) bp apart at each of a pair of tandem promoters, with a third site \((\text{O3})\) located \( \sim 300 \) bp further upstream. Repression at either promoter is improved in the presence of the other sites and is detectable when O1 is placed \( \sim 5 \) kb away from O2 (50). DNA loops between all sites have been observed by electron microscopy (51).

We inserted the deoO1 and O2 operators into sites 1 and 2 of the loopometer reporter, the deoR gene into the pCYMR-1 expression module and tested loop assistance in strains in which the endogenous deoR gene was deleted (‘Materials and Methods’ section). A clear Oid-dependent increase in LacI repression was seen with increasing expression of DeoR (Figure 3A). A similar degree of looping was seen for the reporter exposed to the endogenous level of DeoR protein in a deoR\( ^{+} \) strain (Figure 3A).

Hybrid \( \lambda -186 \) CI repressor

We have previously shown that a hybrid repressor in which the \( \lambda \) CI DNA-binding N-terminal domain (NTD) is fused to the phage 186 CI oligomerization C-terminal domain (CTD) is able to regulate the \( \lambda \) PR and PRM promoters at OR in a way that responds to the presence of OL located 3.8 kb away (52). The intact 186 CI protein is a nonlambloid repressor that uses distal binding sites for regulation in vivo (53) and forms DNA loops in vitro (27). The 186 CI CTD forms a wheel-like 14-mer structure that should present attached NTDs on its rim for DNA binding (52). It is thus very likely that this hybrid protein can form DNA loops; however, this has not been independently confirmed.

When we expressed the \( \lambda -186 \) CI hybrid from integrated pCYMR-1 (‘Materials and Methods’ section) in the full \( \lambda \) OL-OR loopometer reporter, we saw an increase in Oid-dependent LacI repression indicative of DNA looping (Figure 3B). Looping was weaker than we saw with intact \( \lambda \) CI (Figure 2C).

Lambda phage repressors: HK022 CI and \( \varphi \)KO2 CB

\( \varphi \)HK022 is a lambdoid phage with a similar basic genomic arrangement to \( \lambda \), including divergent lytic operons controlled by OL and OR elements separated by the immunity region, including the cl gene (54). The natural distance between OL and OR is \( \sim 700 \) bp in HK022, compared to 2.3 kb in \( \lambda \). The HK022 CI repressor shows homology with other lambdoid repressors and is known to bind cooperatively to adjacent operators within OR and OL (54). DNA looping by HK022 CI seems possible given these similarities to \( \lambda \), and is further suggested by loop-like DNase I sensitivities.
Figure 3. Testing candidate proteins and sites for DNA looping. LacZ units produced by loopometer reporters in lacI+Oid+ and lacI+Oid– strains, normalized to the average of the Oid– values. Protein expression was induced with CA from chromosomally integrated pCYMR vectors or an empty vector control. Measurements in lacI–Oid+ strains showed no substantial effects of candidate proteins in the absence of LacI and are omitted to save space, except for (G) and (H). Error bars are 95% confidence limits (Student’s t), n = 6 unless otherwise stated. Sequences of inserts at loopometer sites 1 and 2 are given in Supplementary Figure S1; protein sequences inserted in pCYMR are given in Supplementary Figure S4. (A) DNA looping between deoO1 and deoO2 operators by endogenous or exogenous E. coli DeoR. (B) DNA looping between full λ 186 CI and OR123 sites by a hybrid protein with the λ CI DNA-binding N-terminal domain and the phage 186 CI C-terminal association domain. (C) DNA looping between phage HK022 OL and OR by HK022 CI. (D) DNA looping between phage κKO2 OL and OR by κKO2 CB. (E) DNA looping by integrase dInt7 (catalytically inactivated) between its attB and attP sites (left), but a lack of looping between its attL and attR sites (right). (F) A lack of looping between matS-binding sites by endogenous E. coli MatP is indicated by the absence of decreased LacZ units compared to a control reporter with only a single matS site. (G) and (H) Inconclusive looping assays for phage 186 CI looping between its pR sites and regulatory FL site, and for phage P1 RepA looping between P1 triple-iteron sites. In both cases, the expressed proteins have confounding lacI-independent and Oid–-independent effects.
induced by CI in the 53 bp between \textit{OL2} and \textit{OL3}, and by
the location of a CI operator 300 bp downstream of \textit{OR} that
is involved in immunity (54,55). However, DNA looping has
not been tested.

The entire HK022 \textit{OR} and \textit{OL} regions (with mutations to
inactivate the \textit{pL} and \textit{pR} promoters; Supplementary Figure
S1A) were inserted at the loopometer sites 1 and 2 (Supple-
mentary Figure S2). Expression of HK022 CI produced an
Oid-dependent increase in LacI repression of the reporter,
clearly showing DNA looping (Figure 3C).

\textit{Klebsiella oxytoca} phage \textit{qKO2} is another lambdoid
phage, closely related to N15. These ‘telomeric’ phages are
unusual in forming nonintegrated linear prophages and
are classed as lambdoid based on their operon structures
and sequence similarities (56). The \textit{qKO2} immunity repres-
sor \textit{CB} shows homology with lambdoid repressors and re-
presses lytic transcription from \textit{OL} and \textit{OR} regions located
\~630 bp apart, either side of the \textit{cB} gene (57). The simila-
rities to \textit{\lambda} led Hammerl \textit{et al.} (57) to speculate that \textit{qKO2}
\textit{CB} protein might form a DNA loop between \textit{OL} and \textit{OR}.
However, this has not been experimentally validated.

We inserted \textit{qKO2} \textit{OL} and \textit{OR} (with mutations to in-
activate promoters; Supplementary Figure S1A) into the
loopometer and expressed the \textit{qKO2} \textit{CB} protein. Increased
LacI repression seen in the presence of \textit{CB} induction con-
firmed \textit{OL}-\textit{OR} looping by \textit{CB} (Figure 3D).

Thus, all three of the lambdoid phage repressors
tested—\textit{\lambda}, HK022 and \textit{qKO2}—display DNA looping be-
tween the operators that control early lytic transcription.
The genomic arrangement of divergent lytic promoters sep-
parated by a short immunity region containing the repres-
sor gene is a module common to many lambdoid phages,
suggesting that long-range repressor looping is widespread.
In \textit{\lambda}, CI looping increases repression of the lytic promot-
ers (14,15), and also provides complex control of the im-

\textit{\lambda} promoter (17). Looping-dependent cooperative re-
pression of the lytic promoters seems likely to be a shared
feature; however, differences in the operator arrangements
in HK022 and \textit{qKO2} suggest that these and other lambdoid
phages may use looping differently in control of the immu-
nity promoter.

\textbf{dInt7—an inactivated serine integrase}

Site-specific recombination between sites either on the same
or different DNA molecules must involve the formation of
a protein bridge between the two DNA sites. However, the
strength of this DNA bridging capability is not clear, since
the catalytic steps of recombination may be sufficiently fast
that only a transient DNA contact is required.

Serine integrases are a large class of site-specific re-
combinases that, in the absence of a recombination di-
rectionality factor (RDF), catalyze a unidirectional reac-
tion between their attachment sites—\textit{attP} + \textit{attB} \rightarrow \textit{attL} +
\textit{attR}—without the need of cofactors such as IHF (58). The
bridging complex is an Int tetramer with the two DNA sites
each bound to an Int dimer.

To test the DNA-looping capacity of serine integrases, we
used \textit{Int7}, one of a set of 34 large serine integrases isolated
and characterized by Yang \textit{et al.} (44). We made an inactive
variant—\textit{dInt7}—by mutating the catalytic serine (Ser10) to
alanine, and inserted this into the pCYMR-1 expression
module. The cognate \textit{attB} and \textit{attP} sites were inserted into
sites 1 and 2 of the loopometer reporter, in an inverted ori-
dentation. We saw a clear \textit{Oid}-dependent increase in LacI re-
pression when dInt7 was expressed (Figure 3E), indicative
of substantial DNA looping between \textit{attB} and \textit{attP}.

The inability of the serine integrases to recombine \textit{attL}
and \textit{attR} is thought to be due to poor synopsis between In-
bound \textit{attL} and \textit{attR} sites. Electrophoretic mobility shift as-

ays with \textit{\phi}C31 and BxB1 integrases found synaptic com-
plexes only between \textit{attP} and \textit{attB} (59,60). A model stimu-
lated by structural studies of the L1 integrase bound to \textit{attP}
DNA suggests that upon completion of the DNA strand
cleavage and rejoining reactions between \textit{attP} and \textit{attB}, new
cis interactions form between the Int monomers bound at
\textit{attB} and \textit{attR} that prevent Int tetramerization (61).

To test the expected lack of synopsis between the \textit{attL7}
and \textit{attR7} sites, we used transient expression of active Int7
to recombine \textit{attP7} and \textit{attR7} in the loopometer reporter,

generating reporters with an inverted 2 kb internal segment
flanked by \textit{attL7} and \textit{attR7} sites. Expression of dInt7 did not
result in detectable DNA looping in these reporters (Figure
3E). This result provides \textit{in vivo} confirmation of an inability
of the integrase alone to bring \textit{attL} and \textit{attR} together, at
least for \textit{Int7}. However, the assay does not by itself show
whether the looping defect is due to a lack of Int bind-
ing to \textit{attL} or \textit{attR} or whether Int is bound but cannot
loop.

In general, it is not possible to be certain from a nega-
tive loopometer result that a protein and sites are incapable
of looping, because the assay does not provide independent
confirmation that the protein is expressed, is active and is
binding its DNA sites. In the case of dInt7, we at least know
that the expressed protein is active by its looping of \textit{attP}–
\textit{attR}, but while other serine integrases are known to bind
to \textit{attL} and \textit{attR} sites \textit{in vitro} (59,60), it is possible that
this binding is not occurring for dInt7 \textit{in vivo}.

\textbf{A negative result for MatP:matS}

Another example of a negative result is provided by our
examination of looping by MatP. The \textit{E. coli} MatP pro-
tein binds to 13 bp sites termed \textit{matS} that are clustered
around the replication terminus of the chromosome. Some
23 \textit{matS} sites are distributed over an \~800 kb region cor-
responding to the Ter macrodomain, with MatP and \textit{matS}
sites functioning to structure and insulate this domain (62).
X-ray crystallography of MatP–\textit{matS} complexes showed a
MatP tetramer forming a bridge between two separate \textit{matS}
DNAs, with DNA looping also observed \textit{in vitro} by elec-
tron and atomic force microscopy (63), leading to the idea
that the Ter domain might involve a network of MatP–\textit{matS}
looping. However, more recent Hi-C measurements did not
reveal strong contacts between \textit{matS} sites. This and other
results indicate that MatP exerts its effects without DNA
looping (64).

To test MatP/\textit{matS} looping, we constructed loopometer
reporters carrying a \textit{matS} sequence at sites 1 and 2. To avoid
effects on the cells due to the elimination of the endogenous
MatP protein (62), we made control reporters without \textit{matS}
at site 1. We found that LacI looping was not enhanced by
the presence of the second matS site, indicating a lack of looping (Figure 3F).

This result supports the idea that MatP does not loop matS sites in vivo. However, our assay does not confirm the binding of MatP to our matS sites, and it is possible that the DNA sequence context or the cellular location somehow reduces binding at our sites relative to the natural sites at Ter. The conflict between the in vitro evidence for looping and the in vivo results may be due to a variety of factors. The in vivo concentration of MatP may be far from optimal for looping; either too low to give significant occupancy of matS sites (note that ChIP assays (62) can be positive even when occupancy is low) or so high that MatP tetramers bind independently to each matS site. Bound MatP dimers may also interact with other proteins that block dimer-dimer contacts in vivo.

Inconclusive results: 186 CI and P1 RepA

In two cases, testing of DNA looping in the loopometer gave inconclusive results because the expression of the candidate protein affected LacZ units in the control strains lacking either LacI or Oid.

In attempting to test DNA looping by intact phage 186 CI repressor between its FL and pR regulatory sites (53), we found that while expression of 186 CI reduced LacZ units in the lacI Oid site, it also reduced LacZ units in the lacI Oid and lacI Oid control strains. This effect was also seen (Figure 3G) when low levels of expression of 186 CI were used to avoid potential cellular effects by the use of the pCYMR-4 variant, which uses a weaker expression promoter (Supplementary Figure S3). Expression of 186 CI also inhibited LacZ activity in reporters that contain non-186 sequences at sites 1 and 2, suggesting that 186 CI is somehow affecting Lac expression or its repression of Plac. The cause of this effect is currently unresolved and prevents conclusions about 186 CI DNA looping from the assay.

We also attempted to test looping by the phage P1 RepA DNA replicase protein, which has been proposed to interact when bound to iteron sequences located on different DNA molecules in a ‘handcuffing’ interaction that is thought to help control the copy number of the plasmid P1 prophage (8). However, we found that the expression of P1 RepA also affected LacZ activity of the control strains (Figure 3H).

These results show that the loopometer is not a foolproof assay for DNA looping and highlight the importance of using the control reporter strains.

It has been pointed out to us that these control strains would not reveal the effects of an expressed protein on general factors that could specifically affect DNA looping, such as levels of DNA supercoiling (65) or nucleoid-associated proteins (21, 66). In cases where such effects are suspected, we recommend testing the effect of the expressed protein on Oid and Oid reporters that do not contain binding sites for the protein, in order to detect any confounding effects on LacI looping.

Model-based quantitation of looping strength

A simple display of whether or not a candidate protein and its sites interact to loop DNA provides useful information about the mechanism of action of the protein. The fractional reduction in LacZ units in the lacI Oid reporter relative to the lacI Oid reporter (Figure 3) allows a comparison of the looping effect of the different proteins, sites and concentrations, but does not provide quantitation of the relative strengths of looping. Quantitation of looping strength can provide more information about the effect of concentrations or site variants, the likely importance of looping in the protein’s activity and allows better comparison of DNA looping by different proteins.

To quantitate looping strength, the first step is to calculate the fractional LacI looping, that is, the fraction of time that Oid and O2 are looped by LacI. This is obtained using Equation (1) (Figure 4A; (34)), which takes into account background LacZ units that result from low level lacZ expression from sources other than Plac. This background can be measured using strain ID1285, which carries a mutated Plac promoter (‘Materials and Methods’ section), giving ~15 LacZ units with our assay. The fractional decrease in units seen with the lacI Oid reporter relative to the lacI Oid reporter gives a linear read-out of the fractional looping: when there is no looping, the Oid and Oid units are the same, giving F1 = 0; if looping were 100%, Plac would be fully repressed and the Oid units would equal the background, giving F1 = 1.

Figure 4B shows the F1 values calculated from the data of Figures 2 and 3. In the absence of the test proteins, the lacI Oid and lacI Oid reporters give F1 values (for LacI alone) of ~0.74. That is, at the LacI concentration in our reporters, the 3 kb loop between Oid and O2 is formed ~74% of the time. F1(X), the fraction of LacI looping in the presence of the various internal looping proteins, is increased over these F1 values (Figure 4B). Slight increases in looping are seen with uninduced expression (0 μM CA) due to some leak in CymR repression. The maximal values of F1(X), allow the looping strengths of the test proteins and sites to be ranked. In this set, the strongest looping was exhibited by λC1/Oli23-ori23 and the weakest looping by λC1/Oli-ori1, with the other proteins and sites giving intermediate looping strengths. We note that this comparison between different proteins is based on the assumption that differences in the structure of the bridge formed by the different proteins between sites 1 and 2 do not substantially affect LacI looping, that is, that the measured differences in LacI looping are due solely to differences in the frequency of formation of the internal loop. The large size of the external loop, 500 + 500 bp, gives us some confidence in this assumption.

The F1 and F1(X) values can be used to move beyond a simple ranking and to obtain estimates of the relative looping strengths by estimating the strength of looping relative to LacI.

The approach uses a simple model for loop assistance (Figure 5A; 35). The model specifies four species due to the looped or nonlooped state of each pair of sites. Each of these species can be assigned a statistical weight or relative propensity (w1 to w4). The propensity to form a DNA loop is a balance between the energetic cost of bringing together two DNA sites (primarily entropic at these long distances), and the energetic benefit provided by the interaction of the sites due to the protein-DNA and protein-
protein interactions involved. The cost is inversely related to the effective relative concentration of the two DNA sites, $J$, which for sites in *cis* is a function of the distance between them along the DNA (29,35,36). We represent the benefit by the factor $I$, which is effectively a loop dissociation constant, being inversely related to the benefit and having units of concentration (35). $I$ quantitates the ‘looping strength’ of a protein and its DNA sites, with lower $I$ indicating higher looping strength. $I$ is determined in a complex way by the specific DNA:protein and protein:protein binding constants, the protein’s concentration and the various looping and nonlooping complexes formed at the two DNA sites. Each of the weights for the single-looped species is given by the ratio of the $J$ factor for the DNA loop and the $I$ factor for the protein-mediated pairing (Figure 5A).

The weight for the double-looped species is the weight for the internal DNA loop and the $I$ factor for the DNA sites, with lower $I$ indicating higher looping strength. $I$ is determined in a complex way by the specific DNA:protein and protein:protein binding constants, the protein’s concentration and the various looping and nonlooping complexes formed at the two DNA sites. Each of the weights for the single-looped species is given by the ratio of the $J$ factor for the DNA loop and the $I$ factor for the protein-mediated pairing (Figure 5A).

Using the loopometer data and two additional measurements, it is possible with this model to obtain an estimate of the ratio of the $I$ value for the candidate protein ($I_{X}$) relative to that of LacI between *Oid* and *Oid* (IL, IL). The weights $w_{2}$ to $w_{4}$ can be expressed in terms of $I_{L}/I_{X}$ and the three weights for LacI looping of the various DNA segments: $J_{L,ab}/I_{L}, J_{L,bc}/I_{L}$ and $J_{L,ac}/I_{L}$, designated $p, q$ and $r$ for brevity (Figure 5B). Substituting these terms into the equation for $F_{L(X)}$ allows $I_{L}/I_{X}$ to be obtained from $F_{L(X)}$ if $p, q$ and $r$ are known (Figure 5B, Equation 2). Estimates of these three weights can be obtained because the weight for a single DNA loop is related to $F$ for that loop: $J/I = F/(1–F)$ (35). The weight $p$ can be obtained from the $F_{L}$ measurement of LacI looping of the loopometer in the absence of the internal protein (Figure 5B). To estimate $q$, we constructed and assayed *Oid* and *Oid* versions of reporters with *Oid* and O2 separated by a 2 kb b-like DNA segment (Figure 5B; strains ID1290 *lacI* *Oid* and ID1291 *lacI* *Oid*; ‘Materials and Methods’ section ). This b’ segment contains most of the internal loopometer b segment, but with different sequences near *Oid* and *Plac*. Measurement of LacI looping in these reporters measures $F_{L}(b’)$, giving an estimate for $q = J_{b’}/I_{L}$.

We note that equal looping of these b’ and b segments by LacI has not been demonstrated. To estimate $r$, we constructed and assayed reporters in which the internal b segment was removed, leaving *Oid* and *O2* separated by a 1 kb ac segment comprised of the joined 500 bp a and c arms (strains ID1292 *lacI* *Oid* and ID1293 *lacI* *Oid*), allowing measurement of $F_{L}(ac)$. Assuming that the looping of the DNA-joined ac segment is similar to the looping of the a+c segments when they are brought together by the b-looping bridge formed by the candidate protein, we can use $F_{L}(ac)$ as a proxy for $F_{L}(ac+c)$. We note that we saw some divergence from this assumption in a previous study, suggesting that the internal b loop or the protein bridge may affect the looping of the a and c arms (35). Thus, the $F_{L}$ and $F_{L(X)}$ measurements from the loopometer, combined with assays of four pre-made strains allows estimation of $I_{L}/I_{X}$ for the candidate protein and its sites.

The $I_{L}/I_{X}$ values calculated in this way for the tested looping proteins and their sites are given in Figure 4B. Note that higher $I_{L}/I_{X}$ values indicate higher looping strength of the candidate protein relative to LacI/*Oid*–*O2* (as $I$ is inversely related to the interaction strength). The relationship between $I_{L}/I_{X}$ and $F_{L(X)}$ is plotted in Figure 5C, allowing a simple read-out of $I_{L}/I_{X}$ from the loopometer measurement of $F_{L(X)}$. The various errors in the measurements result in substantial uncertainties in these $I_{L}/I_{X}$ values, particularly for the stronger looping proteins. For LacI looping between OL123–OR123, we obtained $I_{L}/I_{X} = 1.1$, with a 95% confidence interval of 0.57–3.17 (Figure 4B). This range spans the $I_{L}/I_{X}$ value of 2.7 obtained previously by our more direct comparison of LacI/OL123–OR123 and LacI/*Oid*–*O2* looping (35). The steepness of the $I_{L}/I_{X}$ versus $F_{L(X)}$ plot at high $F_{L(X)}$ values indicates that the loopometer is not well suited for distinguishing between very strong DNA-looping proteins. However, the curve shows that the loopometer...
Figure 5. Estimating looping strength relative to LacI/Oid-O2. (A) Our statistical–mechanical model for two nested DNA loops applied to the case of the loopometer reporter (see text). (B) Derivation of Equation (2), which relates $I_L/I_X$ (the looping strength of the candidate protein and sites relative to LacI/Oid-O2) to loopometer measurements of $F_L$ and $F_{L(X)}$ and the statistical weights $p$, $q$ and $r$ for looping of particular DNA segments by LacI alone (see text). $F_{L,abc}$ is the mean of the eight $F_L$ values in Figure 4 ($n = 8$), but values for individual reporters could also be used. $F_{L,b'}$ and $F_{L,ac}$ values ($n = 32$ and 28) were obtained from strains ID1290 and ID1291 or ID1292 and ID1293, respectively. Values are ± 95% confidence limits (Student's $t$). (C) The plot of Equation (2), using the $p$, $q$ and $r$ values of (B), showing the measured $F_{L(X)}$ values and derived $I_L/I_X$ values of the proteins/sites tested.
should be capable of detecting and measuring very weak looping, down to looping ~100-fold weaker than LacI/Oid-2O.

The $I_L/I_X$ values do not provide absolute measurements of looping strength for individual proteins, but the internal LacI standard allows the looping strengths of the different proteins to be quantitatively compared. For example, the data indicate that maximal looping by xCl between OL12 and OR12 is some 2-fold weaker than for the full 3-operator sites, with single operator OL1-OR1 looping a further 5-fold weaker. Importantly, the use of these $I_L/I_X$ values should assist the comparison of looping strengths of proteins that may be assayed under different conditions, aiding comparisons between assays done in different labs.

Ideally, the measurement of $I_L/I_X$ in the loopometer would allow estimation of the fractional looping of the protein and its sites at any DNA distance. The fractional looping for a single loop is given by the simple relation: $F = J/(J + I)$, and since we previously estimated $I_L = 83 \text{ nM}$ (35), we should be able to obtain $I_X$ from the $I_L/I_X$ value. We also determined a power law relationship between $J$ and DNA distance $d$ (bp): $J = 1.04 \times 10^6 \times d^{0.15} \text{ nM}$ by measuring LacI looping over distances ranging from 300 to 50 000 bp in the E. coli chromosome (summarized in (36)). Together, this should allow the $F$ versus $d$ relationship to be obtained for any $I_L/I_X$ value. However, these previous measurements were made with an older version of our LacZ assay, which we now know gives an underestimate of LacI looping (Supplementary Figure S5). To recalibrate the relationship between LacI looping and DNA distance, we constructed loopometer reporters in which the Oid-2O distance was increased by 8 kb by the integration of a plasmid into an integrase attachment site that had been included in our original constructs (Figure 6A; ‘Materials and Methods’ section). By assaying Oid" and Oid" versions of these reporters, we thus obtained $F_L$ values for ~10 and 11 kb spacings. Combining these with the $F_L$ measurements for ~1, 2 and 3 kb spacings (Figure 5B), we found a reasonable match between these five values and the predictions of the $F$ versus $d$ relationship (Figure 6B; $I_L/I_X = 1$ curve), if either $I_L$ was about 2-fold lower, or the power law prefactor was about 2-fold higher, than our previously measured values. This recalibration allows the fractional looping of the proteins and their sites at any DNA distance on the E. coli chromosome to be estimated from the $I_L/I_X$ values (Figure 5). For example, the loopometer measurement of $I_L/I_X = 0.16$ for DeoR/deoO1-O2 predicts ~72% looping between these sites at their natural 600 bp spacing (65–82% based on the 95% confidence interval for $I_L/I_X$) and 20% looping at a 5 kb separation. The predicted looping between λ OL123 and OR123 at their natural 2.3 kb spacing is ~79%, in agreement with an independent measurement using high-resolution live-cell microscopy (31); though uncertainty in $I_L/I_X$ gives a large range for predicted OL123-OR123 looping: 48–92%). Note that prediction of looping frequencies for spacings below 500 bp is unreliable due to the onset of helical phasing sensitivity at these distances.

CONCLUSIONS

The loopometer represents a simple assay that can provide strong evidence for DNA looping in vivo. At a minimum, it requires only the construction of two plasmids—one for the binding sites and one to express the candidate

![Figure 6](https://example.com/figure6.png)

**Figure 6.** Estimating fractional looping for different DNA distances. (A) Reporters with increased Oid-Plac.O2 spacings were made by integrating an 8 kb plasmid into the abc and b’ segments of reporter strains to make strains ID1318/ID1319 (abc: Oid’ /Oid") and strains ID1320/ID1321 (b’: Oid’ /Oid”). $F_L$ values (± 95% confidence limits, Student’s t, n = 8) were measured as in Figure 4. (B) Plot of predicted fractional looping by the candidate protein, $F_X$, versus DNA distance, $d$ (bp), for different values of $I_L/I_X$, using the relationships $F = J/(J + I)$ (35) and $J = 1.04 \times 10^6 \times d^{0.15} \text{ nM}$ (36) and assuming $I_L = 38.3 \text{ nM}$. Points are $F_L$ looping values for LacI/Oid-2O from Figure 5 and (A) and are reasonably fitted by the $F_X$ curve for $I_L/I_X = 1$. 

$\text{CONCLUSIONS}$

The loopometer represents a simple assay that can provide strong evidence for DNA looping in vivo. At a minimum, it requires only the construction of two plasmids—one for the binding sites and one to express the candidate
protein—followed by sequential integration of these (and the empty expression plasmid) to make six bacterial strains for the LacZ assay. Quantitation of the looping strength relative to LacI/Oid-O2 is enabled by further assays of five pre-made strains, for measuring background and for LacI looping calibration.

Aside from the detection and measurement of DNA looping, the assay permits examination of the effect of protein concentration on looping, which may be an important factor in regulation. Additional constructions and assays can provide comparisons of looping between different sites, including testing the effect of binding site mutations. Similarly, the effects of protein mutations on looping can readily be examined, though the assay does not by itself distinguish between effects on protein–protein versus protein–DNA interactions.

The in vivo fluorescence microscopy-based assay of DNA looping by Hensel et al. (31) is the only approach that is currently able to provide similar information to the loopometer. In this study, the OL- and OR-binding sites for LacI were placed adjacent to sites bound by fluorescently tagged proteins, and the distances between these sites were measured in live cells in the presence of CI. While this assay could readily be adapted to other proteins, it requires specialized equipment and expertise and is likely to be limited to strong DNA-looping proteins, as only these can efficiently loop DNA over the long distances required to resolve looped and unlooped states microscopically.

The loopometer assay currently has a number of limitations, some of which should be able to be overcome by further development: (i) The assay can fail if there are confounding effects of the expressed protein on looping-independent expression of lacZ. This could potentially be overcome by the use of a different reporter or by replacing LacI with a different looping protein. (ii) Comparison of looping strength between different proteins relies on the untested assumption that the precise structure of the bridge formed by the different proteins between sites 1 and 2 does not substantially affect LacI’s ability to loop the two 500 bp DNA arms. We suspect that DNA arms of different lengths could be used to test and validate this assumption. (iii) The assay only tests homotypic interactions, that is looping by a single protein. However, the assay could be easily modified to detect heterotypic interactions by introducing a second protein expression system. (iv) The assay can only test proteins that have specific binding sites and can be expressed in the active form in E. coli, which may limit testing of eukaryotic proteins. By itself, the assay does not provide evidence for the activity of the protein unless looping is observed. The use of fusions to bacterial DNA-binding domains (e.g. the λ CI NTD) may extend the range of testable proteins, since the ability to fold an active DNA-binding domain would not be required. We note that such fusions could provide an alternative bacterial one- or two-hybrid assay, allowing detection of interacting domains by DNA looping. (v) A bacterial assay may not be appropriate for eukaryotic looping proteins because of the lack of necessary accessory factors (e.g. nucleosomes). However, the principle of detecting DNA-looping proteins by loop assistance should also work in eukaryotic cells, as studies have shown that DNA-looping interactions can increase gene expression by bringing an enhancer closer to the promoter (67–69). Thus, insertion of a DNA sequence flanked by binding sites for a candidate DNA-looping protein between a promoter and enhancer, combined with an inducible protein expression system, could be used to generate a qualitative eukaryotic loopometer.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Kim Sneppen, Sandeep Krishna and members of the Shearwin lab for discussions.

FUNDING

National Health and Medical Research Council [GNT1100651]; Australian Research Council [DP160101450]; CSIRO (to N.H., in part via Synthetic Biology Future Science Platform. Funding for open access charge: University of Adelaide [75125608].

Conflict of interest statement. None declared.

REFERENCES

1. Schleif, R. (1992) DNA looping. Ann. Rev. Biochem., 61, 199–223.
2. Matthews, K.S. (1992) DNA looping. Microbiol. Rev., 56, 123–136.
3. Cournac, A. and Plumbridge, J. (2013) DNA looping in prokaryotes: experimental and theoretical approaches. J. Bacteriol., 195, 1109–1119.
4. Dröge, P. and Müller-Hill, B. (2001) High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. BioEssays, 23, 179–183.
5. Bush, M. and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. Microbiol. Mol. Biol. Rev., 76, 497–529.
6. Long, H.K., Prescott, S.L. and Wysocka, J. (2016) Ever-changing landscapes: transcriptional enhancers in development and evolution. Cell, 167, 1170–1187.
7. Schoenfelder, S. and Fraser, P. (2019) Long-range enhancer–promoter contacts in gene expression control. Nat. Rev. Genet., 20, 437–455.
8. Das, N. and Chatteraj, D.K. (2004) Origin pairing (‘handcuffing’) and unpairing in the control of Pl plasmid replication. Mol. Microbiol., 54, 836–849.
9. Schumacher, M.A., Tonthat, N.K., Kwong, S.M., Chinnam, N.B., Liu, M.A., Skurray, R.A. and Firth, N. (2014) Mechanism of staphylococcal multiresistance plasmid replication origin assembly by the RepA protein. Proc. Natl. Acad. Sci. U.S.A., 111, 9121–9126.
10. Becker, N.A., Greiner, A.M., Peters, J.P. and Maher, L.J. (2014) Bacterial promoter repression by DNA looping without protein–protein binding competition. Nucleic Acids Res., 42, 5485–5504.
11. Choy, H.E., Park, S.W., Parrack, P. and Adhya, S. (1995) Transcription regulation by inflexibility of promoter DNA in a looped complex. Proc. Natl. Acad. Sci. U.S.A., 92, 7327–7331.
12. Chetverina, D., Aoki, T., Erokhin, M., Georgiev, P. and Schedl, P. (2014) Making connections: insulators organize eukaryotic chromosomes into independent cis-regulatory networks. BioEssays, 36, 163–172.
13. Grindley, N.D.F., Whiteson, K.L. and Rice, P.A. (2006) Mechanisms of site-specific recombination. Annu. Rev. Biochem., 75, 567–605.
14. Révet, B., Von Wilcken-Bergmann, B., Bessert, H., Barker, A. and Müller-Hill, B. (1999) Four dimers of a repressor bound to two suitably spaced pairs of operators form octamers and DNA loops over large distances. Curr. Biol., 9, 151–154.
15. Svenningsen, S.L., Costantino, N., Court, D.L. and Adhya, S. (2005) On the role of Cro in lambda prophage induction. Proc. Natl. Acad. Sci. U.S.A., 102, 4465–4469.
61. Rutherford, K. and Van Duyne, G.D. (2014) The ins and outs of serine integrase site-specific recombination. Curr. Opin. Struct. Biol., 24, 125–131.
62. Mercier, R., Petit, M.A., Schbath, S., Robin, S., El Karoui, M., Boccard, F. and Espézi, O. (2008) The MatP/matS site-specific system organizes the terminus region of the E. coli chromosome into a macrodomain. Cell, 135, 475–485.
63. Dupaigne, P., Tonthat, N.K., Espézi, O., Whitfill, T., Boccard, F. and Schumacher, M.A. (2012) Molecular basis for a protein-mediated DNA-bridging mechanism that functions in condensation of the E. coli chromosome. Mol. Cell, 48, 560–571.
64. Lioy, V.S., Cournac, A., Marbouy, M., Duigou, S., Mozziconacci, J., Espézi, O., Boccard, F. and Koszul, R. (2018) Multiscale structuring of the E. coli chromosome by nucleoid-associated and condensin proteins. Cell, 172, 771–783.
65. Lewis, D.E.A., Geanacopoulos, M. and Adhya, S. (1999) Role of HU and DNA supercoiling in transcription repression: specialized nucleoprotein repression complex at gal promoters in Escherichia coli. Mol. Microbiol., 31, 451–461.
66. Becker, N.A., Kahn, J.D. and Maher, L.J. (2007) Effects of nucleoid proteins on DNA repression loop formation in Escherichia coli. Nucleic Acids Res., 35, 3988–4000.
67. 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., 106, 20222–20227.
68. 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.
69. Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.F. et al. (2017) Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat. Commun., 8, 15993.