

# DNA sequence-dependent mechanics and protein-assisted bending in repressor-mediated loop formation

## SUPPLEMENTARY INFORMATION

James Q. Boedicker, Hernan G. Garcia, Stephanie Johnson, Rob Phillips

### Table of Contents:

- I. Experimental procedures
  - a. Strain Construction
  - b. Supplementary Table 1: Host strains used in this study
  - c. Supplementary Table 2: Primers used in this study
  - d. Supplementary Tables 3 and 4: Sequences of constructs
  - e. Measuring Gene Expression
  - f. HU purification
  
- II. Theoretical derivations
  - a. The sensitivity of repression to changes in the looping energy
  - b. Calculating looping energies for E8 and TA from cyclization data
  - c. Supplementary Table 5: Parameters used in calculations
  - d. Thermodynamic equilibrium model for Lac repressor-mediated gene regulation involving loop formation
  - e. A model incorporating both unassisted and assisted loop formation
  - f. Incorporation of HU into the looping model
  
- III. Additional results
  - a. HU compacts DNA in the range of concentrations used *in vitro* in this work
  - b. HU changes the looping J-factor of a DNA but not repressor-operator dissociation constants
  - c. HU dramatically increases the looping probabilities of both in-phase and out-of-phase operator constructs
  - d. HU alters the phasing of Lac repressor-mediated looping *in vitro*
  - e. HU does not preferentially stabilize a particular looped “state” *in vitro*

## **I. Experimental Procedures:**

### **a. Strain Construction**

We created a library of different realizations of the *lac* operon harboring two operators, as seen in Figure 1A. Previously we constructed a series of constructs containing a random looping sequence called E8. Here we extend this library by replacing the looping sequence with a flexible sequence called TA. See [1] for further information on strain construction. The sequences corresponding to these constructs and their variable looping sequences can be found in Supplementary Tables 3 and 4. All constructs were verified by sequencing, and constructs and sequences are available upon request. These constructs were integrated into the genome of *E. coli* strain HG104 with the wild-type *lacI* background using recombineering as described in Garcia and Phillips [2]. All constructs were integrated into the *galK* gene.

Strain HG105 containing a deletion of *lacI* was used to obtain the unregulated level of expression of our constructs. DNA looping constructs were moved into this strain using P1 transduction ([http://openwetware.org/wiki/Sauer:P1vir\\_phage\\_transduction](http://openwetware.org/wiki/Sauer:P1vir_phage_transduction)). Host strains containing deletions of *hupA* and *hupB* were obtained from the Keio collection. Strains with deletions of both *hupA* and *hupB* are called  $\Delta$ HU. P1 phage transduction was used to transfer looping constructs into the  $\Delta$ HU versions of HG104 and HG105 to create JB101 and JB102 respectively, see Supplementary Table 1. FLP-FRT recombination was used to remove the *kan* cassette of the *hupA* and *hupB* deletions after transduction as described previously [3], see Figure S14. Looping constructs were selected using kanamycin.

### **b. Supplementary Table 1: Host strains used in this study.**

Host Strains	Genotype	Source or reference
<i>E. coli</i>	MG1655	
HG104	MG1655 $\Delta$ <i>lacZYA</i>	[2]
HG105	MG1655 $\Delta$ <i>lacIZYA</i>	[2]
JB101	HG104 $\Delta$ <i>hupA</i> , $\Delta$ <i>hupB</i>	This study
JB102	HG105 $\Delta$ <i>hupA</i> , $\Delta$ <i>hupB</i>	This study

### **c. Supplementary Table 2: Primers used in this study.**

Primers	Sequence	Comments
3.1	GTGCAATCCATCTTGTTCAATCAT	Sequence <i>lac</i> constructs in <i>galK</i>
3.2	CCTTCACCCTCTCCACTGACAG	Sequence <i>lac</i> constructs in <i>galK</i>
<i>hupA</i> Fw	CTGATTTGTCGTACCTGGAGTCTTC	Sequence <i>hupA</i> region
<i>hupA</i> Rv	GAAGTGAAGAGTTATGACTACAGGCAGTGAG	Sequence <i>hupA</i> region
<i>hupB</i> Fw	ATTGCCGATCTGGACATTCATCCTGTG	Sequence <i>hupB</i> region
<i>hupB</i> Rv	AGACGATTCAGCACCTGTTGACG	Sequence <i>hupB</i> region

**d. Supplementary Tables 3 and 4: Sequences of constructs.**

**Supplementary Table 3: Regions in both E8 and TA constructs.**

Region	Sequence
Upstream of auxiliary operator	AGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATT CCGACGTC
Auxiliary operator (Oid)	AATTGTGAGCGCTCACAATT
Promoter	TTTACAATTAATGCTTCCGGCTCGTATAATGTGTGG
Main operator (O2)	AAATGTGAGCGAGTAACAACC
Downstream of main operator	AATTCATTAAGAGGAGAAAGGTACCGCATGCGTAAAGGAGAAGA <u>ACTTT</u>

Underlined portion indicates the beginning of the YFP coding region.

**Supplementary Table 4: Sequences of the E8 and TA variable looping regions.**

See attached Boedicker et al Supplementary Table 4

**e. Measuring Gene Expression:**

Cells were revived from -80°C frozen stocks by culturing overnight in Luria Broth (EMD, Gibbstown, NJ) containing kanamycin at 37°C, with shaking at 250 rpm. 1 µL of LB cultures was then used to inoculate 3 mL scale cultures containing M9 (2 mM MgSO<sub>4</sub>, 0.10 mM CaCl<sub>2</sub>, 48 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.6 mM NaCl, 19 mM NH<sub>4</sub>Cl) with 0.5% glucose as a carbon source. Cells were grown in 14 mL Falcon BD tubes with the cap placed loosely. M9 cultures were grown at 37°C with shaking at 250 rpm for approximately 10 generations and were harvested at OD<sub>600</sub> between 0.25 and 0.65. In this range of optical densities the average ratio of fluorescence intensity to OD<sub>600</sub> is approximately insensitive to OD<sub>600</sub> [1].

Repression is defined as the fold change of gene expression levels as the result of Lac repressor, Equation 1. Experimentally, this is the ratio of gene expression in cells with and without Lac repressor. Cells not containing the fluorescent YFP reporter were also grown to determine the background autofluorescence of the cells. 200 µL of culture was loaded into the wells of a 96 well plate (Costar, #3631, Corning, NY). Fluorescence measurements from the bottom of each well were obtained using a Tecan Safire 2, with excitation and emission of 505 and 535 nm respectively, both with a 12 nm bandwidth.

To calculate the repression at a given operator distance, first the background fluorescence corresponding to the media background was subtracted from all measurements. Then fluorescence

measurements were normalized by dividing by the optical density of each culture at 600 nm ( $OD_{600}$ ), and autofluorescence obtained from the cells not containing YFP was subtracted. On each day, all strains were measured in triplicate and a mean and standard deviation for each day was calculated for each strain. Measurements were repeated on multiple days, and the mean and standard error for each construct was calculated from the means of each day weighted by the standard deviation, as described previously [1]. Based on previous characterization of YFP gene reporters in our experimental setup [4], the dynamic range of repression measurements are up to a level of approximately 500.

#### **f. HU purification**

A liter of strain RJ5814 (see Materials and Methods in main text) was grown in LB plus ampicillin and 0.3% glucose at 30°C to  $OD_{600}$  of 0.75. HU expression was induced by shifting to 42°C for 30 minutes with shaking. Cells were collected by centrifugation and resuspended in 50 mL of HK buffer (20 mM HEPES-NaOH, pH 7.5, 60 mM KCl, 1 mM DTT) supplemented with 1mM PMSF and Roche Complete protease inhibitor cocktail, and lysed by microfluidization. Cell debris were pelleted by centrifugation, and then the supernatant fractionated by a two-step ammonium sulfate precipitation, with the first step at 70% saturation, after which the HU was in the supernatant, and then 90% saturation, after which the HU was in the pellet. The pellet was resuspended in 20 mL HK, and the two-step ammonium sulfate precipitation repeated. After the second 90% saturation fractionation, the pellet was again resuspended in 10 mL HK and then passed over two tandem 5 mL Q-Sepharose columns pre-equilibrated with HK buffer. The flow-through from this column was applied directly to a 5 mL heparin column, also pre-equilibrated with HK. The column was first washed with HK buffer with 400 mM KCl, and then the HU eluted with a 60 mL linear gradient from 400 mM to 1.2 M KCl. HU does not absorb at 280 nm [5, 6], so the flow-through was monitored at 250 nm absorbance. HU eluted as two peaks at 590 mM and 793 mM KCl; the second peak was two to three times as concentrated as the first peak, and was taken to be heterodimeric HU (the other peak was assumed to be homodimeric HU [Reid Johnson, University of California, Los Angeles, personal communication]). Protein concentration was determined by a BCA assay. Glycerol was added to 10% and the protein flash-frozen and stored at -80°C.

## II. Theoretical derivations:

### a. The sensitivity of repression to changes in the looping energy

We implement Equation 4 to predict how repression changes as a function of the looping free energy. In Figure 3C we plot by what factor the repression level will increase when the free energy change needed to form the loop is decreased by 0 to 2  $k_B T$  for  $\Delta F_{loop}$  between 0 and 20  $k_B T$ . Specifically the curves correspond to

$$\text{Fold change in } Repression_{loop} = \frac{Repression_{loop}(\Delta F_{loop} - \text{offset})}{Repression_{loop}(\Delta F_{loop})}, \quad (S1)$$

in which  $Repression_{loop}$  as a function of looping energy is calculated using Equation 4 and the parameters listed in Supplementary Table 5. We find that for the range of looping energies measured for the random looping sequence, 8 to 10  $k_B T$ , repression will be sensitive to a small reduction in the looping energy.

### b. Calculating looping energies for E8 and TA from cyclization data

Most DNA cyclization experiments are reported in terms of the J-factor [7]. This magnitude can be understood as the concentration of one end of the DNA molecule in the vicinity of the other one. Analogously, the looping J-factor,  $J_{loop}$ , can be defined in the context of *in vitro* DNA looping experiments [8]. An alternative way of viewing both these experiments is through the prism of the cyclization free energy,  $\Delta F_{cyc}$ . It can be shown that the two magnitudes are related by

$$J = \frac{1}{v} e^{-\beta \Delta F_{cyc}}, \quad (S2)$$

where  $v$  represents the biochemical “standard state” of the reaction. Usually, this is taken to be  $v^{-1} = 1$  M.

The cyclization data reported by [9] is indeed expressed in terms of the J-factor. In order to be able to use it as an input to our model, we need to convert it to a cyclization free energy using Equation S2. The free energy of cyclization calculated from the Cloutier data is shown in Supplementary Figure 1. The overall difference in free energy between the two sequences is about 2.3  $k_B T$ . Also note in Supplementary Figure 1 that changing from the flexible sequence TA to the random sequence E8 results in a vertical shift in the J-factor over the loop lengths measured. The locations of the peaks and troughs as a function of length do not change. This indicates that switching between these sequences results in a change in the bendability of the loop without noticeable changes to the twisting energy. The J-factor or looping energy accounts for both the twisting and bending energies needed to place the operators in alignment. It appears that the sequence dependence of the looping energy results from modulation of the bending energy for these two sequences.

**c. Supplementary Table 5: Parameters used in calculations.**

Parameter	Value	Units	Description	reference
<i>dissociation constants to DNA in vitro</i>				
$K_{d,Ra}$	12±3	pM	LacI dissociation constant for operator Oid	[10]
$K_{d,Rm}$	240±50	pM	LacI dissociation constant for operator O2	[10]
$K_{d,HU}$	480	nM	Dissociation constant for nonspecific binding of HU to dsDNA	[11, 12]
<i>binding energies to DNA in vivo</i>				
$\Delta\epsilon_{rad}$	-17±0.2	$k_B T$	LacI binding energy to operator Oid	[2]
$\Delta\epsilon_{rmd}$	-13.9±0.2	$k_B T$	LacI binding energy to operator O2	[2]
$\Delta\epsilon_{HU}$	-9.7	$k_B T$	HU binding energy to nonspecific dsDNA	Calculated from <i>in vitro</i> value
$\Delta\epsilon_{pd}$	-7	$k_B T$	RNA polymerases binding energy to promoter	[13]
<i>in vivo parameters</i>				
R	11±2	-	number of LacI tetramers per cell	[2]
HU	30,000	-	number of HU dimers per cell	[14]
P	2,000	-	number of RNA polymerases per cell	[15]
$N_{NS}$	$4.6 \times 10^6$	bp	nonspecific binding sites, size of <i>E. coli</i> K12 genome in bp	GenBank: U00096.2

**d. Thermodynamic equilibrium model for Lac repressor-mediated gene regulation involving loop formation**

Thermodynamic equilibrium models of gene regulation have been described extensively in previous works [2, 16-24]. Below is a brief derivation of our version of these kinds of models, whose states and weights are depicted in Figure 2A.

In our experiments we measured repression, which is the fold reduction of gene expression due to the presence of Lac repressor, as defined in Equation 1. Using Equation 3, this definition leads to

$$\text{Repression}_{loop}(L) = \frac{\text{gene expression (R = 0)}}{\text{gene expression (R \neq 0)}} = \frac{\frac{k_2 P}{N_{NS}} e^{-\beta \Delta \epsilon_{pd}}}{1 + \frac{P}{N_{NS}} e^{-\beta \Delta \epsilon_{pd}}} \cdot \frac{\frac{k_2 P}{N_{NS}} e^{-\beta \Delta \epsilon_{pd}} + \frac{k_3 P}{N_{NS}} \frac{2R}{N_{NS}} e^{-\beta(\Delta \epsilon_{pd} + \Delta \epsilon_{rad})}}{1 + \frac{P}{N_{NS}} e^{-\beta \Delta \epsilon_{pd}} \left(1 + \frac{2R}{N_{NS}} e^{-\beta \Delta \epsilon_{rad}}\right) + \frac{2R}{N_{NS}} \left(e^{-\beta \Delta \epsilon_{rad}} + e^{-\beta \Delta \epsilon_{rmd}} + e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rmd} + \Delta F_{loop}(L))}\right) + \frac{2R(R-1)}{N_{NS}^2} e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rmd})}}$$
 , (S3)

where we have explicitly written out the value for Z in the denominators. Because the states in which RNA polymerase is bound are rare, with a probability of approximately  $10^{-4}$ , calculated using Equation 2 and the parameters listed in Supplementary Table 5, we assume that escape from the RNA polymerase bound state by initiation of transcription does not significantly alter the equilibrium distribution of states. Based on recent measurements, the rate constants for states 2 and 3 are equivalent for operator distances more than 80 bp and therefore cancel in the final expression [25]. We also make the weak promoter approximation [2, 26]. This simplifies Equation S3 to the expression shown in Equation 4 in the main text.

In order to calculate the looping energy  $\Delta F_{loop}(L)$  from repression data, we solve Equation 4 in terms of  $\Delta F_{loop}(L)$  to get,

$$\Delta F_{loop}(L) = \frac{-1}{\beta} \ln \frac{\text{Repression}_{loop}(L) \left(1 + \frac{k_3}{k_2} \frac{2R}{N_{NS}} e^{-\beta \Delta \epsilon_{rad}}\right) - 1 - \frac{2R}{N_{NS}} \left(e^{-\beta \Delta \epsilon_{rad}} + e^{-\beta \Delta \epsilon_{rmd}}\right) - \frac{4R(R-1)}{N_{NS}^2} e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rmd})}}{\frac{2R}{N_{NS}} e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rmd})}} . \quad (\text{S4})$$

For further developments of the model, including assisted looping and the role of HU in loop formation, similar equations can be derived starting with a different set of states and associated weighting terms, as discussed below.

#### e. A model incorporating both unassisted and assisted loop formation.

This model looks at the potential influence of assisted DNA-bending in loop formation due to, for example, a DNA-bending protein such as HU [27, 28]. We assume DNA loop formation occurs through two mechanisms, unassisted looping and assisted looping involving DNA-bending proteins. We adjust the thermodynamic model as shown in Figure 7A.

As in Figure 7A,  $\Delta F_{loop,u}(L)$  is the unassisted looping energy and  $\Delta F_{loop,a}(L)$  is the assisted looping energy. Repression for this model can be derived as above, resulting in Equation 5 of the main text. To rewrite Equation 5 for the flexible and random looping sequences, we make following assumptions:

- 1) The assisted looping energy is not sequence-dependent,  $\Delta F_{loop,a,flexible}(L)$  for the flexible sequence is the same as  $\Delta F_{loop,a,random}(L)$  for the random sequence. As a first estimate, we assume assisted looping contributes the most to the looping energy shown in Figure 3B. Therefore, for both sequences  $\Delta F_{loop,a}(L)$  is the looping energy extracted previously from the experimental data for the flexible sequence in the presence of HU.
- 2) For the flexible sequence, the assisted looping energy is less than the unassisted looping energy by an amount  $\delta$  measured in  $k_B T$ .
- 3) The unassisted looping energy for the flexible sequence is lower than for the random sequence by an amount  $\sigma$ , which is a function of the sequence of the looping region and is measured in  $k_B T$ .

These three assumptions result in,

$$\Delta F_{loop,a,flexible}(L) = \Delta F_{loop,a,random}(L) = \Delta F_{loop,u,flexible}(L) - \delta = \Delta F_{loop,u,random}(L) - \delta - \sigma, \quad (S5)$$

leading to Equations 8 and 9 of the main text.

#### f. Incorporation of HU into the looping model

The thermodynamic model of loop-mediated gene regulation was adapted to account for the role of the DNA-bending protein HU in loop formation. Above we have described a model involving assisted and unassisted loop formation; here we extend that model further to explicitly include HU binding in the looping region.

The states and weights for this model are shown in Supplementary Figure 10. First we derived a model for the case of only one HU protein binding in the looping region, using states 1-8 in Supplementary Figure 10. This results in the following expression for the probability of loop formation *in vitro*,

$$p(\text{looping}) = \frac{\frac{R \cdot J_{HU=0}}{2K_{d,Rm} K_{d,Ra}} + \frac{R \cdot HU \cdot J_{HU=1}}{2K_{d,Rm} K_{d,Ra} K_{d,HU}}}{\left(1 + \frac{HU}{K_{d,HU}}\right) \left(1 + \frac{R}{K_{d,Rm}} + \frac{R}{K_{d,Ra}} + \frac{R^2}{K_{d,Rm} K_{d,Ra}}\right) + \frac{R \cdot J_{HU=0}}{2K_{d,Rm} K_{d,Ra}} + \frac{R \cdot HU \cdot J_{HU=1}}{2K_{d,Rm} K_{d,Ra} K_{d,HU}}}, \quad (S6)$$

in which  $J_{HU=i}$  is the  $J_{loop}$  for a loop which contains  $i$  HU proteins bound. Expressions similar to Equation S6 were derived, following Equation 2, to calculate the probability of specific looped states. These probabilities of looped states as a function of *in vitro* HU concentration or *in vivo* number of HU per cell are reported in Supplementary Figure 13.

Using the values in Supplementary Table 5, the *in vitro* data at 0 HU was used to calculate the value of  $J_{HU=0}$ . The *in vitro* TPM data shown in Figure 8 was used to fit for the value of  $J_{HU=1}$  for both the flexible

and random looping sequences. These  $J_{loop}$  values are reported in Figure 8C as looping energies. To convert from *in vitro*  $J_{loop}$  to *in vivo* looping energies we use,

$$\Delta F_{loop} = -\frac{1}{\beta} \ln(J_{loop} \nu) + \Delta G_{background}, \quad (S7)$$

in which  $\beta$  is the inverse of Boltzmann's constant times temperature,  $\nu$  accounts for the biochemical "standard state" of the reaction in  $M^{-1}$ , and  $\Delta G_{background}$  is an energy offset which accounts for a difference in the background free energy of the reference state *in vivo* versus *in vitro*. That is, from the *in vitro*  $J_{loop}$  values we know the relative energy differences between the looped state, but we will need to correct for this background energy difference to find the *in vivo* looping energies.

To calculate the value  $\Delta G_{background}$  we take into account that *in vivo* repression is approximately 200 as in Figure 3A. For the two-state model whose states are shown in Supplementary Figure 10, repression was calculated using,

$$\text{Repression}_{\text{two-state model}} = \frac{1 + \frac{2R}{N_{NS}} (e^{-\beta \Delta \epsilon_{rad}} + e^{-\beta \Delta \epsilon_{rd}}) + \frac{2R(R-1)}{N_{NS}^2} e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rd})} + \frac{2R}{N_{NS}} e^{-\beta(\Delta \epsilon_{rad} + \Delta \epsilon_{rd})} (e^{-\beta \Delta F_{loop}(L)} + \frac{HU}{N_{NS}} e^{-\beta(\Delta F_{loop}(L) + \Delta \epsilon_{HU})})}{1 + \frac{2R}{N_{NS}} e^{-\beta \Delta \epsilon_{rad}}} \cdot \left(1 + \frac{HU}{N_{NS}} e^{-\beta \Delta \epsilon_{HU}}\right). \quad (S8)$$

Using Equation S8 to calculate repression, we find the value of  $\Delta G_{background}$  that corrects the looping energies such that the level of repression is 200.

The differences in the background free energy of the unlooped reference state *in vivo* vs. *in vitro* could have many potential contributions, such as the contribution of nucleoid proteins like IHF or Fis to prebending looping region or differences in the structure of DNA such as differences in the extent of supercoiling. Future work is needed to quantify how factors such as these influence the propensity for loop formation throughout the genome and in *in vitro* studies.

In addition to the two-state model, a three-state looping model was also derived. Equations analogous to S6 and S8 were derived for the three-state model using the states and weights listed in Supplementary Figure 10. These equations are similar in form to S6 and S8, but include an extra term to account for the state in which two HU proteins are bound in the loop.

### **III. Additional results:**

#### **a. HU compacts DNA in the range of concentrations used *in vitro* in this work**

HU is a mostly nonspecific DNA binding protein [29] that has been shown through a variety of *in vitro* assays to have two regimes of action: at relatively low concentrations, it compacts DNA and appears to decrease its persistence length (that is, makes it more flexible), apparently by introducing single flexible hinges or bends; but at high concentrations it extends and stiffens DNA, possibly through the formation

of an HU filament that allows the DNA to wrap around it in a superhelical structure [12, 30-32]. It remains unclear how these two regimes are relevant *in vivo*. The concentration of HU *in vivo* (roughly 10  $\mu$ M for fast-growing cells [14]) should not be enough to allow the formation of filaments at physiological salt concentrations [32]. However, HU is known to bind more strongly to special DNA structures, such as bent or kinked DNA, than to regular duplex DNA [33]. HU may also be recruited to specific locations by interactions with transcription factors (for example, GalR [34]); these and other cellular factors may allow the nucleation of filaments even when the total cellular concentration is below the *in vitro* nucleation threshold [31, 32].

The critical concentration of HU that determines the transition between the compacted, flexible regime and the extended, stiff regime is salt dependent, with the transition occurring at lower HU concentrations for lower salt concentrations [32]. We show in Supplementary Figure 5 that at the concentration of monovalent salt in our TPM assays (200 mM) we are able to access the low-concentration regime in which HU compacts the DNA. Within this low-concentration regime our results are consistent with the HU-induced DNA compaction that has been seen previously in single-molecule assays. For example, we observe a maximal compaction of our DNA tethers to 93% of their starting lengths at 500-1000 nM HU, consistent with the range of HU concentrations over which maximal compaction was observed previously (roughly 40% compaction at 500-800 nM HU at comparable salt concentrations [32]).

The discrepancy between the amount of compaction at the maximum—but not the HU concentration at which this compaction occurs—could be due to the fact that the DNAs we use are much shorter than those of [32], offering far fewer binding sites for HU at a given concentration. It is known from TPM assays that the 45 degree bend induced when Lac repressor binds to the operator Oid compacts the tether approximately 2 nm, and that two Lac repressors bound to the same tether results in 2.5-5 nm compaction of the tether dependent on whether or not the two bends are in or out of phase [10]. HU is able to accommodate a variety of bend angles, potentially causing each HU to compact the tether more than 2 nm. Given that our tethers are approximately 450 bp and the binding footprint of HU is about 20 bp [35], roughly 20 copies of HU could bind to the tether. Supplementary Figure 5 shows that HU compacts our tether nearly 20 nm. Assuming the 20 nm compaction resulted from 20 HU bound on the tether, that gives 1 nm compaction per HU bound, a similar number to the compaction per binding of Lac repressor. It is likely that fewer than 20 HU bind per tether under these conditions, suggesting that each HU protein compacts the tether more than 1 nm.

#### **b. HU changes the looping J-factor of a DNA but not repressor-operator dissociation constants**

In both the two-state and three-state models presented in this work, for the effect of HU on looping *in vitro* we make a fundamental assumption that HU acts only to alter the J-factor of the DNA in the loop (particularly its bending energy, but potentially its twisting energy as well, as discussed in the next sections). That is, we assume HU does not alter the affinity of the Lac repressor for its operators. To verify that this assumption is valid, we explicitly tested whether HU also changes repressor-operator dissociation constants as well as J-factors.

In [10] we showed that Lac repressor concentration titration curves, in which the looping probability is measured as a function of repressor concentration, can be used to extract both looping J-factors and repressor-operator dissociation constants for a given DNA molecule, using a statistical mechanical model analogous to that described in Figure 2A. To ensure that HU affects only the J-factor and not the dissociation constants, we performed a similar repressor titration curve with one of the DNAs that we characterized extensively in [10], but here with the addition of a constant amount of HU. As derived in [10], a change in one or both repressor-dissociation constants would manifest as a shift in the repressor concentration in which looping is maximal; but a change in J-factor alone would manifest as an increase in looping probability at all repressor concentrations.

As shown in Supplementary Figure 6, 10 nM HU increases the apparent looping J-factor by a factor of 2 (which corresponds to a decrease of about half of a  $k_B T$  in the looping energy). However, HU leaves repressor-operator dissociation constants unchanged as the repressor concentration at which looping is maximal is unchanged with or without HU. We note that as derived in [10], data at low repressor concentrations (that is, below the maximum of looping) are most useful for determining dissociation constants, so more data were taken at these low concentrations.

**c. HU dramatically increases the looping probabilities of both in-phase and out-of-phase operator constructs**

In Figure 8 of the main text we show that the addition of HU can dramatically increase the looping probability even of a DNA molecule whose operators are out-of-phase (that is, where the looping probability in the absence of HU is a minimum). In Supplementary Figure 7 we show that HU acts similarly on in-phase operators as well. The addition of 500 nM HU abolishes the sequence dependence to looping that is seen in the absence of HU. Such dramatic increases in apparent flexibility in the presence of HU, over many DNA lengths, have also been demonstrated for J-factors measured in *in vitro* ligase-mediated cyclization assays [36]. We note also that the hint of phasing that may be present in Supplementary Figure 7 does not follow the phasing of the looping probabilities in the absence of HU shown in Figure 4 in the main text; that is, the minimum of looping with 500 nM HU is around 139.5 bp, rather than 141.5 bp as it is in the absence of HU. As discussed in the next section, this is consistent with previous studies on HU that suggest HU unwinds DNA, thereby changing the phasing of the operators, in addition to inducing bends or kinks.

**d. HU alters the phasing of Lac repressor-mediated looping *in vitro***

As discussed in the previous section in reference to Supplementary Figure 7, it appears that HU may alter the apparent helical period of Lac repressor-mediated looping. Such an alteration would be consistent with *in vivo* repression studies in the presence versus the absence of HU [27], with changes to the apparent helical period of J-factors calculated from *in vitro* ligase-mediated cyclization assays [36], and with the negative supercoiling of HU-bound DNA seen in crystal structures [30, 36, 37]. A change in the helical period in the context of Lac repressor-mediated looping *in vitro* would manifest as a change in the operator spacings at maxima and minima (or, with several helical periods of operator spacings, a change in the number of base pairs between successive peaks and troughs of looping).

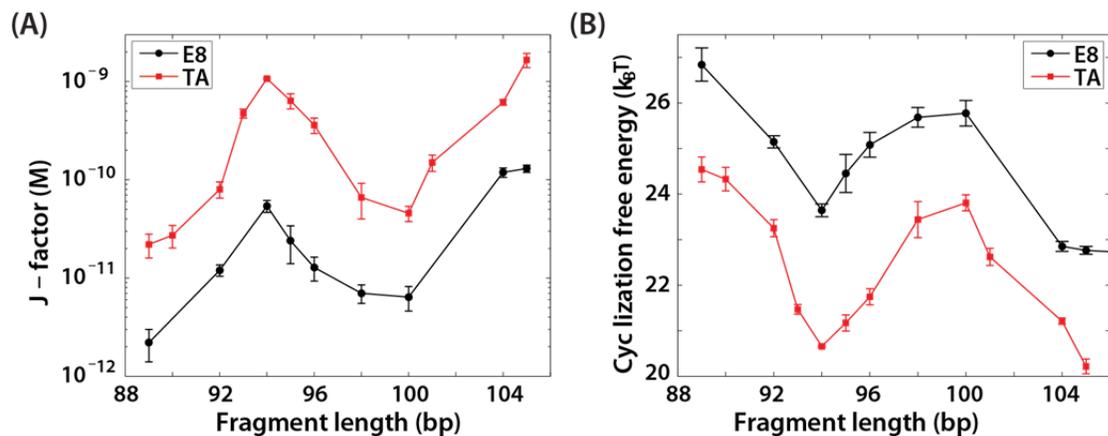
The concentration of HU used in Supplementary Figure 7 is too high to get an accurate sense of the maxima or minima of looping, since all looping probabilities at 500 nM are close to unity. We therefore examined the looping probability as a function of operator spacing at a lower HU concentration. As shown in Supplementary Figure 8, it does appear that the operator spacing at which looping is maximal changes in the presence of 4 nM HU compared to in the absence of HU, although without at least two helical periods of operator spacings we cannot definitively quantify the amount by which HU alters the periodicity of Lac repressor-mediated looping *in vitro*.

**e. HU does not preferentially stabilize a particular looped “state” *in vitro*.**

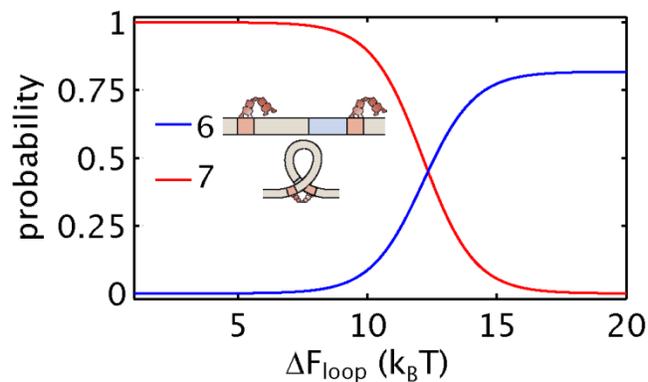
As discussed in [10], in TPM experiments with the Lac repressor we and others observe two looped “states” for any pair of operators, which manifest as two different tether lengths (both shorter than the unlooped state). We call these two states the “bottom” and “middle” looped state, where the bottom state has the smallest root-mean-squared motion of the bead, and the middle state falls between the bottom looped state and the unlooped state. Although it is not yet known what the underlying molecular structures of these two states are, nor whether either or both looped states form *in vivo*, it has been suggested based on *in vivo* repression data that the molecular structures underlying these two states may be responsible for the particular shape of repression-versus-length curves, and that HU may preferentially stabilize some Lac repressor-mediated looped structures compared to others *in vivo* [38]. Here we explore whether HU preferentially stabilizes either of the two looped states that we observe *in vitro*.

The introduction of HU into the TPM experiment can alter the occupation of different looped states. This effect of can be seen in Supplementary Figure 9A, in which the fraction of looping contributed by the middle state is plotted as a function of operator spacing, with and without HU. At all lengths, the inclusion of HU increases looping in the bottom state, leading to a decrease in the fraction of the looping J-factor that is contributed by the middle state. (We note that both with and without HU, the two looped states alternate in prevalence as a function of operator spacing: where the operators are in-phase and looping is maximal, most looping occurs in the middle looped state and the fraction of the total J-factor contributed by the middle state is high; whereas when the operators are out-of-phase, most looping occurs in the bottom looped state and  $J_{\text{loop,M}}/J_{\text{loop,tot}}$  is low.)

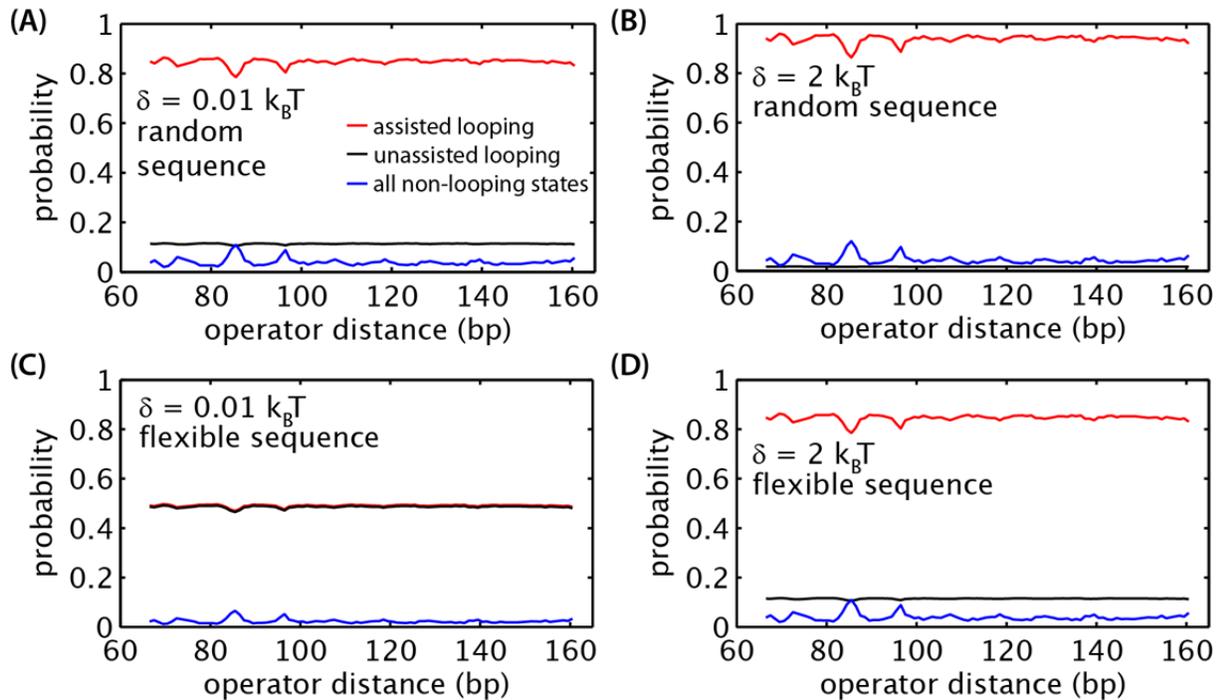
However, we find that HU does not stabilize the bottom state preferentially over the middle state. As shown in Supplementary Figure 9B, as the concentration of HU included in the TPM experiment is increased, the J-factors of the bottom and middle looped states increase to the same degree, such that bottom-state and middle-state J-factors fall on parallel curves.



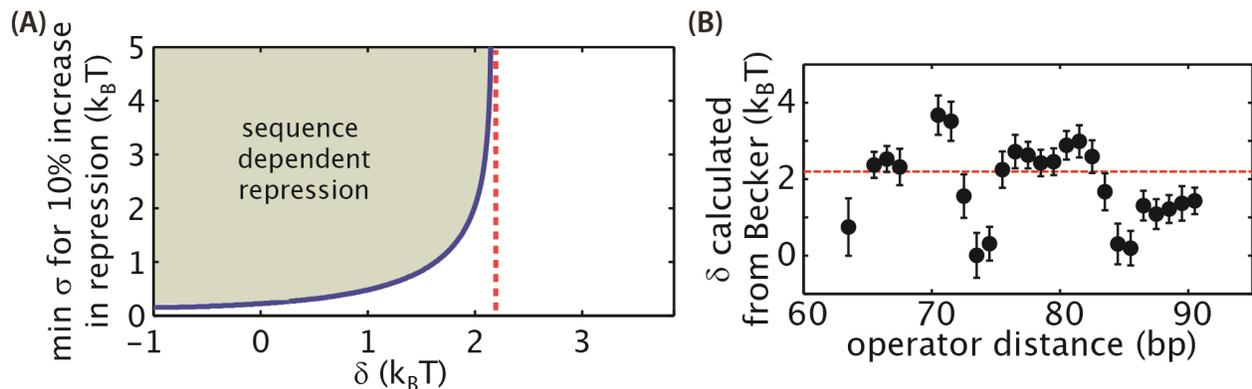
**Supplementary Figure 1:** J-factor measurements for the random sequence E8 and the flexible sequence TA [9]. (B) Looping energies for the sequences E8 and TA extracted from the cyclization measurements using Equation S2.



**Supplementary Figure 2:** Probability of states 6 and 7 from Figure 2A as function of the looping energy calculated using Equation 2 and the parameters in Supplementary Table 5.

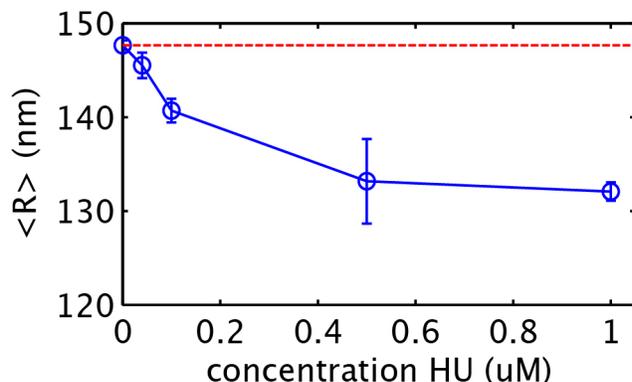


**Supplementary Figure 3:** Predicted probabilities of looped and non-looped states when varying the offset between assisted and unassisted looping energies,  $\delta$ , for the random (A and B) and flexible (C and D) looping sequences. The black lines show the probability of the unassisted looping (state 7 in Figure 7A), the red lines show the probability of the assisted looping (state 8 in Figure 7A), and blue lines show the probability of all unlooped state (the combined probabilities of states 1-6 in Figure 7A).  $\delta=0.01 k_B T$  was used to slightly offset the black and red curves in (C).

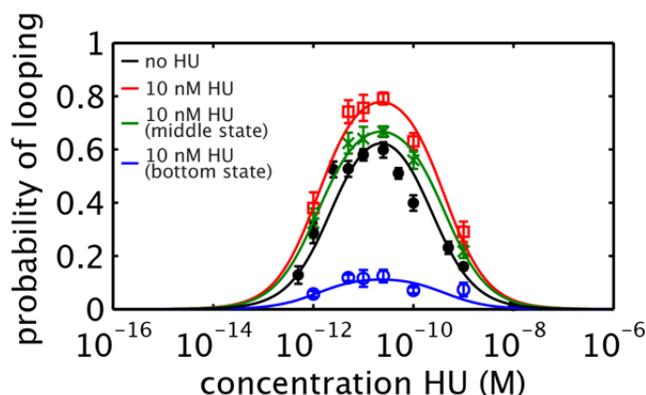


**Supplementary Figure 4:** Prediction of the reduction in looping energy needed to hide sequence dependence. (A) Analysis of the minimal  $\sigma$  needed for a 10% difference in repression as a function of  $\delta$ , notation as in Figure 7. Within the grey shaded region the more flexible sequence will repress at least 10% more than the random sequence. For  $\delta > 2.2 k_B T$  (red dotted line), assisted looping will always mask the sequence dependence of unassisted looping. (B) The value of  $\delta$  as a result of looping assisted by the DNA-bending protein HU calculated from the data of Becker *et. al* [27] shows that HU lowers the looping energy by more than  $2.2 k_B T$  (red dotted line) at most lengths tested. Looping energies were

extracted from measurements reported for wild-type and  $\Delta$ HU strains ( $\Delta$ hupA,  $\Delta$ hupB) as previously described [1]. Error bars are standard deviations.

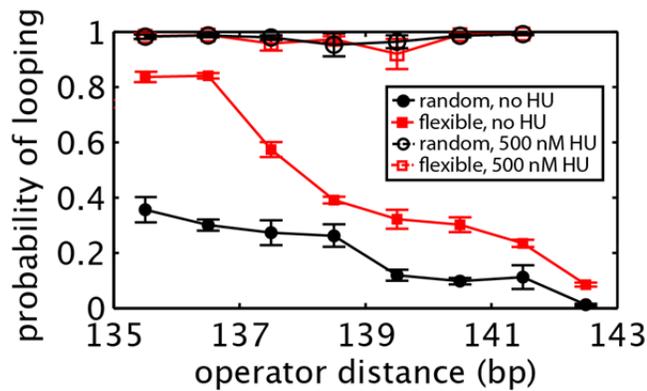


**Supplementary Figure 5:** Our purified HU, in the absence of the Lac repressor, compacts DNA tethers in TPM, consistent with previous single-molecule studies of the effect of HU on DNA tethers [31, 32]. The y-axis gives a measure of effective tether length, the average root-mean-squared motion of beads tethered by a 445 bp DNA, as a function of increasing HU concentration. Horizontal dashed line indicates the average tether length in the absence of HU. Errors bars indicate standard error.

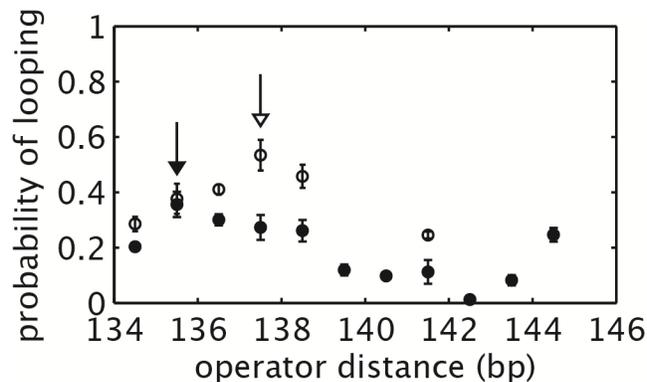


**Supplementary Figure 6:** HU changes only the looping energy, and does not affect repressor-operator affinities *in vitro*. Looping probabilities are shown as a function of repressor concentration, without HU (black data) or at a constant HU concentration of 10 nM (colored data), for a construct whose loop consists of 114.5 bp of the E8 sequence, described in [10]. Unlike the constructs used in this work, this 114.5 bp E8 construct does not contain the promoter as part of the loop; however we have studied this construct extensively in [10] (in which it is called “Oid-E894-O1”), making it the best construct for testing the effect of HU on repressor-operator binding constants. The black data are from [10] and show the 114.5 bp E8 construct lacking a promoter in the absence of HU, which has only one looped state (the “middle” state; see Section IIIe). The red data show the total looping probabilities (that is, the

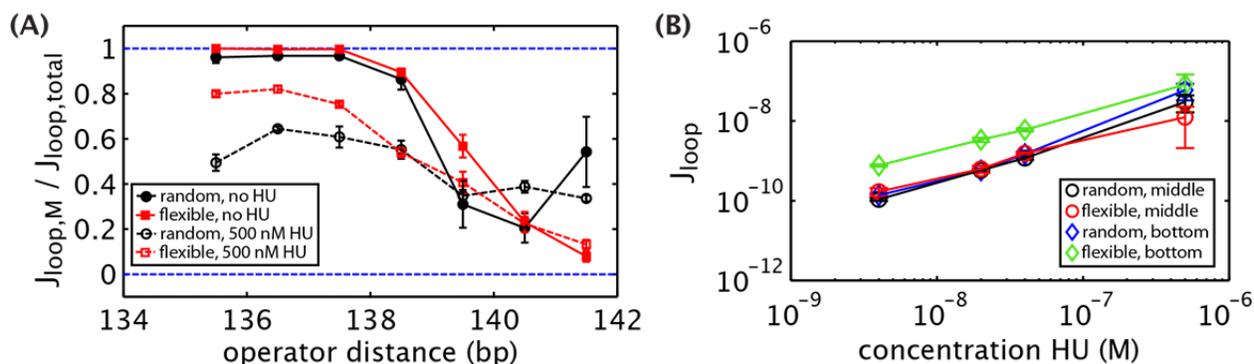
probability of both the “middle” and “bottom” states combined, as shown in the rest of this work) for the same construct but in the presence of 10 nM HU. The addition of HU not only increases the total looping probability but also leads to looping in an additional conformation that we call the bottom looped state (blue) as well as in the middle state (green). Curves show fits to the *in vitro* model for looping probabilities of [10], where the repressor-operator dissociation constants have been fixed to those determined in [10], but the J-factors allowed to vary. All four data sets are well-described by the same dissociation constants, but different J-factors: the J-factor of the middle state in the presence of HU is  $600 \pm 100$  pM, and of the bottom state is  $100 \pm 60$  pM, while the J-factor of the middle state in the absence of HU is  $330 \pm 30$  pM (and the bottom state shows no looping in the absence of HU). The increased error in J in the presence of HU is most likely due to the fewer number of data points used in the fit.



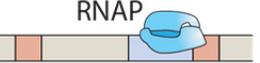
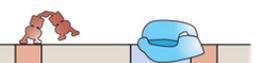
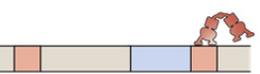
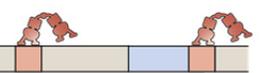
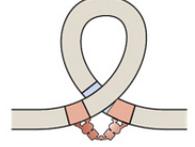
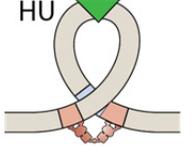
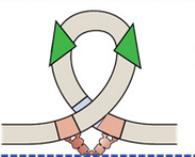
**Supplementary Figure 7:** High concentrations of HU increase looping probabilities regardless of the phasing of the operators. The looping probabilities for random (black circles) and flexible (red squares) looping constructs whose operator distances range from 135.5 bp to 142.5 bp are shown in the presence of 500 nM HU (open symbols) and in the absence of HU (closed symbols, data from Figure 4B). The looping probabilities of the 141.5 bp constructs are shown as a function of increasing HU concentration in Figure 8 of the main text.



**Supplementary Figure 8:** Effect of HU on the operator spacing at which looping is maximal. Looping probabilities of E8-containing constructs whose operator spacings range from 134.5 to 144.5 bp are shown without (closed circles) or with (open circles) 4 nM HU. Here the HU concentration is low enough that modulation of looping probability with operator spacing can be more clearly seen, in contrast to Supplementary Figure 7 where looping probabilities are near unity in the presence of 500 nM HU. As indicated by the arrows, without HU (closed arrow), looping is maximized at 135.5 bp; but in the presence of 4 nM HU (open arrow), looping is maximized at 137.5 bp.

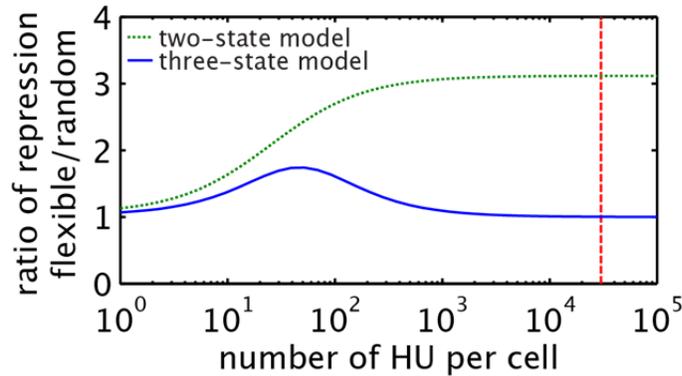


**Supplementary Figure 9:** HU does not preferentially stabilize either of the two looped states. (A) The fraction of the total looping J-factor contributed by the middle state is shown as a function of operator spacing without (closed symbols) or with (open symbols) 500 nM HU. When this ratio is one (indicated by a horizontal blue dashed line), looping occurs only in the middle looped state; when it is zero (indicated by a second blue dashed line), looping occurs only in the bottom looped state. The addition of HU increases the fractional amount of looping in the bottom state. (B) However, HU does not preferentially stabilize either the middle or bottom looped state. Here  $J_{loop}$  for the two states are plotted as a function of increasing HU concentration for constructs containing 141.5 bp of the flexible or random looping sequence. Circles indicate the J-factors for the middle looped state, whereas diamonds indicate J-factors for the bottom looped state. Data for the two looped states fall on roughly parallel lines, indicating that HU does not preferentially stabilize one state over the other.

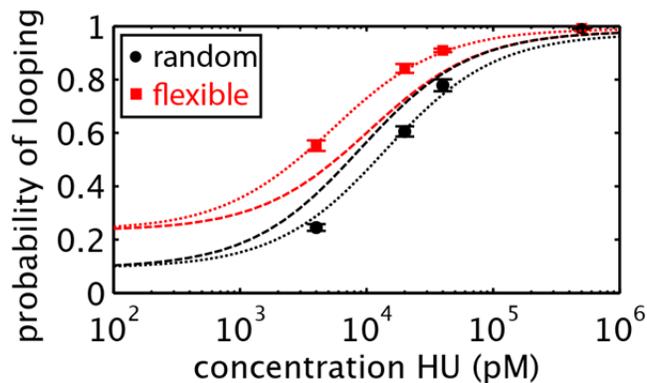
#	STATE	<i>in vivo</i> WEIGHT	<i>in vitro</i> WEIGHT
1		$1+h_{in\ vivo}$	$1+h_{in\ vitro}$
2		$\frac{P}{N_{NS}} e^{-\beta\Delta\varepsilon_{pd}}(1+h_{in\ vivo})$	-
3		$\frac{P}{N_{NS}} \frac{2R}{N_{NS}} e^{-\beta(\Delta\varepsilon_{pd}+\Delta\varepsilon_{rad})}(1+h_{in\ vivo})$	-
4		$\frac{2R}{N_{NS}} e^{-\beta\Delta\varepsilon_{rad}}(1+h_{in\ vivo})$	$\frac{R}{K_{d,Ra}} (1+h_{in\ vitro})$
5		$\frac{2R}{N_{NS}} e^{-\beta\Delta\varepsilon_{rmd}}(1+h_{in\ vivo})$	$\frac{R}{K_{d,Rm}} (1+h_{in\ vitro})$
6		$\frac{4R(R-1)}{N_{NS}^2} e^{-\beta(\Delta\varepsilon_{rmd}+\Delta\varepsilon_{rad})}(1+h_{in\ vivo})$	$\frac{R^2}{K_{d,Rm} K_{d,Ra}} (1+h_{in\ vitro})$
7		$\frac{2R}{N_{NS}} e^{-\beta(\Delta\varepsilon_{rad}+\Delta\varepsilon_{rmd}+\Delta F_{loop}(L))}$	$\frac{R J_{HU=0}}{2 K_{d,Rm} K_{d,Ra}}$
8		$\frac{2R HU}{N_{NS}^2} e^{-\beta(\Delta\varepsilon_{rad}+\Delta\varepsilon_{rmd}+\Delta\varepsilon_{HU}+\Delta F_{loop,HU}(L))}$	$\frac{R HU J_{HU=1}}{2 K_{d,Rm} K_{d,Ra} K_{d,HU}}$
9		$\frac{2R HU^2}{N_{NS}^3} e^{-\beta(\Delta\varepsilon_{rad}+\Delta\varepsilon_{rmd}+2\Delta\varepsilon_{HU}+\Delta F_{loop,2HU}(L))}$	$\frac{R HU^2 J_{HU=2}}{2 K_{d,Rm} K_{d,Ra} K_{d,HU}^2}$

**Supplementary Figure 10:** States and weights of a thermodynamic model with HU binding. The two-state looping model uses states 1-8, with the three-state looping model including the additional state 9 with 2 HU proteins bound in the loop. For simplicity, the non-looping states in which HU binds to the looping region are not depicted; however the presence of HU in non-looped states is accounted for in the weighting terms by  $h_{in\ vivo}$  and  $h_{in\ vitro}$ . In the two-state model  $h_{in\ vivo} = HU/N_{NS} e^{-\beta\Delta\varepsilon_{HU}}$  and  $h_{in\ vitro} = HU/K_{d,HU}$ . In the three-state model  $h_{in\ vivo} = HU/N_{NS} e^{-\beta\Delta\varepsilon_{HU}} + HU^2/N_{NS}^2 e^{-2\beta\Delta\varepsilon_{HU}}$  and  $h_{in\ vitro} = HU/K_{d,HU} + HU^2/K_{d,HU}^2$ . Variables are as described for Figure 2A, with the addition of HU is the concentration of the protein HU,  $\Delta\varepsilon_{HU}$  is the nonspecific binding energy of HU to double-stranded DNA,  $\Delta F_{loop,HU}(L)$  is the free energy needed to form a loop containing a single bound HU,  $\Delta F_{loop,2HU}(L)$  is the free energy needed to form a loop containing two bound HU,  $K_{d,Rm}$  is the dissociation constant for Lac repressor to operator O2,  $K_{d,Ra}$  is the dissociation constant for Lac repressor to operator Oid,  $K_{d,HU}$  is the dissociation constant for the nonspecific binding of HU to double-stranded DNA,  $J_{HU=i}$  is the J-factor for loop formation with i

HU proteins bound in the loop. Looping energies can be calculated from J-factors using Equation S7. Transcription rate constants are as in Figure 2A, with states 8 and 9 having rate constants of 0.

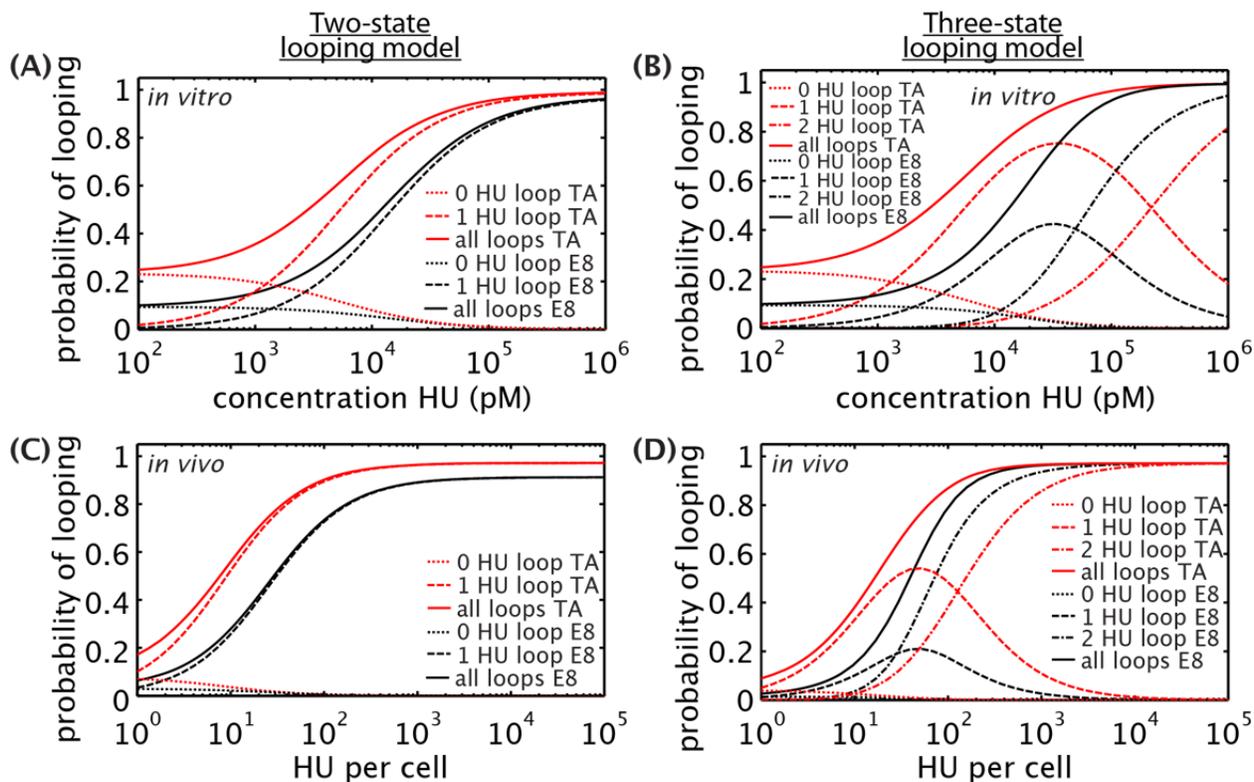


**Supplementary Figure 11:** Predicted ratio of *in vivo* repression for the flexible looping sequence over the random looping sequence calculated using the parameters found in Supplementary Table 5, the looping energies reported in Figure 8C, and Equation S8 or a similarly derived equation for repression in the three-state model. The dashed red line indicates the reported number of HU per cell in wild-type cells [14]. In Figure 5 we report that the ratio of repression for wild-type cells containing the flexible and random looping sequences is approximately 1, which is consistent with the prediction made by the three-state looping model. For the calculation, the looping energies in Figure 8C were converted to *in vivo* looping energies using Equation S7. For the two-state model, *in vivo* looping energies were: TA 0 HU = 14.6  $k_B T$ , TA 1 HU = 8.6  $k_B T$ , E8 0 HU = 15.6  $k_B T$ , and E8 1 HU = 9.8  $k_B T$ . For the three-state model, *in vivo* looping energies used were: TA 0 HU = 15.3  $k_B T$ , TA 1 HU = 9.4  $k_B T$ , TA 2 HU = 8.6  $k_B T$ , E8 0 HU = 16.4  $k_B T$ , E8 1 HU = 10.9  $k_B T$ , and E8 2 HU = 8.6  $k_B T$ .

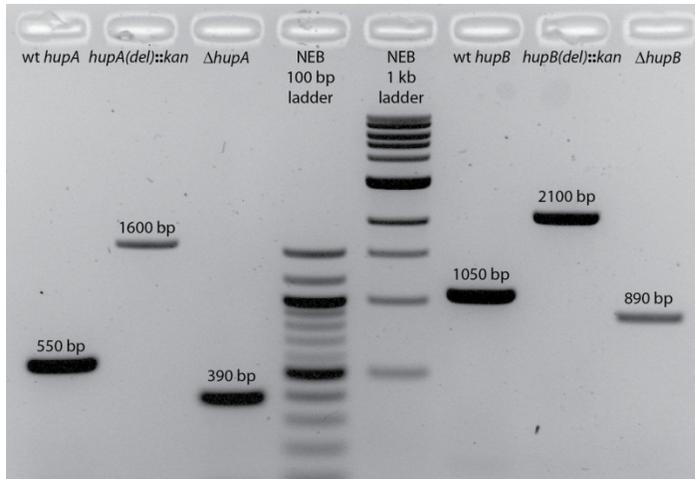


**Supplementary Figure 12:** Fitting the *in vitro* looping data with the two-state looping model where the HU-bound state is forced to have a sequence-independent looping energy. Dotted lines show fits as in Figure 8A, allowing the random and flexible looping sequences to have different values of the looping energy with HU bound. The dashed lines show a HU bound looping energy of 16.9  $k_B T$  for both looping sequences, the best-fit value when the HU bound looping energy is forced to be the same for both looping sequences. When the HU bound looping energy is 16.9  $k_B T$  for both sequences, repression levels *in vivo* will be independent of sequence at 30,000 HU per cell. Such a sequence-independent

assisted looping state would be consistent with what we find *in vivo*; however, it is clear that the *in vitro* data is in poor agreement with a model containing this additional constraint.



**Supplementary Figure 13:** Probability of looping states for the two and three-state looping models both *in vitro* (A and B) and *in vivo* (C and D). Probabilities of individual states in Supplementary Figure 10 are calculated using Equation 2 and parameters found in Supplementary Table 5, Figure 8C, and in the caption of Supplementary Figure 11. The random looping sequence E8 is shown in black and the flexible looping sequence TA is shown in red. For the two-state model it can be seen that the looping states switch between a 0 HU loop (state 7 in Supplementary Figure 10) at low concentrations of HU and a 1 HU loop (state 8 in Supplementary Figure 10) at higher concentrations of HU. In the three-state model the 1 HU loop is populated only at intermediate levels of HU, with the sequence-independent 2 HU loop dominating at higher HU concentrations.



**Supplementary Figure 14:** Agarose gel electrophoresis confirmed the deletion of *hupA* and *hupB* in the chromosome by insertion of *kan* (adding approximately 1050 bp) and the removal of *kan* by FLP-FRT recombination (resulting in bands shortened by approximately 1210 bp). Band lengths of wt *hupA*, wt *hupB*,  $\Delta$ *hupA*, and  $\Delta$ *hupB* are similar to calculated lengths based on the MG1655 genome sequence and a 102 bp scar due to FLP recombination. Gel is 1.5% agarose, stained with SYBR Safe (Life Technologies). DNA ladders were purchased from New England Biolabs, Inc. Length of the PCR product labeled above each band is approximate.

#### Supplemental References:

- [1] Boedicker J Q, Garcia H G and Phillips R 2013 Theoretical and experimental dissection of DNA loop-mediated repression *Phys Rev Lett* **110** 018101-
- [2] Garcia H G and Phillips R 2011 Quantitative dissection of the simple repression input-output function *P Natl Acad Sci USA* **108** 12173-8
- [3] Datsenko K A and Wanner B L 2000 One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products *P Natl Acad Sci USA* **97** 6640-5
- [4] Garcia H G, Lee H J, Boedicker J Q and Phillips R 2011 Comparison and calibration of different reporters for quantitative analysis of gene expression *Biophys J* **101** 535-44
- [5] Rouvireyaniv J and Gros F 1975 Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli* *P Natl Acad Sci USA* **72** 3428-32
- [6] Broyles S S and Pettijohn D E 1986 Interaction of the *Escherichia coli* HU protein with DNA - Evidence for formation of nucleosome-like structures with altered DNA helical pitch *J Mol Biol* **187** 47-60
- [7] Shore D, Langowski J and Baldwin R L 1981 DNA flexibility studied by covalent closure of short fragments into circles *P Natl Acad Sci USA* **78** 4833-7
- [8] Han L, Garcia H G, S. B, Towles K B, Beausang J F, Nelson P C and Phillips R 2009 Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS ONE* **4** e5621
- [9] Cloutier T E and Widom J 2005 DNA twisting flexibility and the formation of sharply looped protein-DNA complexes *P Natl Acad Sci USA* **102** 3645-50
- [10] Johnson S, Lindén M and Phillips R 2012 Sequence dependence of transcription factor-mediated DNA looping. *Nucleic Acids Res* **40** 7728-38

- [11] Koh J, Saecker R M and Record M T, Jr. 2008 DNA binding mode transitions of *Escherichia coli* HU alpha beta: Evidence for formation of a bent DNA - protein complex on intact, linear duplex DNA *J Mol Biol* **383** 324-46
- [12] Xiao B T, Zhang H Y, Johnson R C and Marko J F 2011 Force-driven unbinding of proteins HU and Fis from DNA quantified using a thermodynamic Maxwell relation *Nucleic Acids Res* **39** 5568-77
- [13] Brewster R C, Jones D L and Phillips R 2012 Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli* *Plos Comput Biol* **8** e1002811
- [14] Azam T A, Iwata A, Nishimura A, Ueda S and Ishihama A 1999 Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid *J Bacteriol* **181** 6361-70
- [15] Klumpp S and Hwa T 2008 Growth-rate-dependent partitioning of RNA polymerases in bacteria *P Natl Acad Sci USA* **105** 20245-50
- [16] Becker N A, Kahn J D and Maher L J, III 2008 Eukaryotic HMGB proteins as replacements for HU in *E. coli* repression loop formation *Nucleic Acids Res* **36** 4009-21
- [17] Gill S C, Yager T D and von Hippel P H 1990 Thermodynamic analysis of the transcription cycle in *E. coli* *Biophys Chem* **37** 239-50
- [18] Law S M, Bellomy G R, Schlax P J and J. Record M T 1993 *In vivo* thermodynamic analysis of repression with and without looping in *lac* constructs. Estimates of free and local *lac* repressor concentrations and of physical properties of a region of supercoiled plasmid DNA *in vivo* *J Mol Biol* **230** 161-73
- [19] Rippe K, von Hippel P H and Langowski J 1995 Action at a distance: DNA-looping and initiation of transcription *Trends Biochem Sci* **20** 500-6
- [20] Saiz L, Rubi J M and Vilar J M 2005 Inferring the *in vivo* looping properties of DNA *P Natl Acad Sci USA* **102** 17642-5
- [21] Saiz L and Vilar J M 2008 Ab initio thermodynamic modeling of distal multisite transcription regulation *Nucleic Acids Res* **36** 726-31
- [22] Schlax P J, Capp M W and J. Record M T 1995 Inhibition of transcription initiation by *lac* repressor *J Mol Biol* **245** 331-50
- [23] Vilar J M G and Leibler S 2003 DNA looping and physical constraints on transcriptional regulation. *J Mol Biol* **331** 981-89
- [24] Bintu L, Buchler N E, Garcia H G, Gerland U, Hwa T, Kondev J and Phillips R 2005 Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev* **15** 116-24
- [25] Garcia H G, Sanchez A, Boedicker J Q, Osborne M L, Gelles J, Kondev J and Phillips R 2012 Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. *Cell Rep* **2** 150-61
- [26] Bintu L, Buchler N E, Garcia H G, Gerland U, Hwa T, Kondev J, Kuhlman T and Phillips R 2005 Transcriptional regulation by the numbers: applications *Curr Opin Genet Dev* **15** 125-35
- [27] Becker N A, Kahn J D and Maher L J 2005 Bacterial repression loops require enhanced DNA flexibility *J Mol Biol* **349** 716-30
- [28] 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
- [29] Krylov A S, Zasedateleva O A, Prokopenko D V, Rouviere-Yaniv J and Mirzabekov A D 2001 Massive parallel analysis of the binding specificity of histone-like protein HU to single- and double-stranded DNA with generic oligodeoxyribonucleotide microchips *Nucleic Acids Res* **29** 2654-60
- [30] Guo F and Adhya S 2007 Spiral structure of *Escherichia coli* HU alpha beta provides foundation for DNA supercoiling *P Natl Acad Sci USA* **104** 4309-14
- [31] van Noort J, Verbrugge S, Goosen N, Dekker C and Dame R T 2004 Dual architectural roles of HU: Formation of flexible hinges and rigid filaments *P Natl Acad Sci USA* **101** 6969-74

- [32] Xiao B, Johnson R C and Marko J F 2010 Modulation of HU-DNA interactions by salt concentration and applied force *Nucleic Acids Res* **38** 6176-85
- [33] Pontiggia A, Negri A, Beltrame M and Bianchi M E 1993 Protein HU binds specifically to kinked DNA *Mol Microbiol* **7** 343-50
- [34] Kar S and Adhya S 2001 Recruitment of HU by piggyback: a special role of GalR in repressosome assembly *Genes Dev* **15** 2273-81
- [35] Luijsterburg M S, Noom M C, Wuite G J L and Dame R T 2006 The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: A molecular perspective *J Struct Biol* **156** 262-72
- [36] Czapla L, Peters J P, Rueter E M, Olson W K and Maher L J, III 2011 Understanding Apparent DNA Flexibility Enhancement by HU and HMGB Architectural Proteins *J Mol Biol* **409** 278-89
- [37] Swinger K K, Lemberg K M, Zhang Y and Rice P A 2003 Flexible DNA bending in HU-DNA cocrystal structures *EMBO J* **22** 3749-60
- [38] Saiz L and Vilar J M G 2007 Multilevel deconstruction of the in vivo behavior of looped DNA-protein complexes *PLoS ONE* **2** e355

Boedicker et al, Supplementary Table 4

Operator distance (bp)	Sequence of variable looping region for E8	Sequence of variable looping region for TA
80.5	GTTATCTCGAGTTAGTACGACGTC	CGCCAATAGGATTACTTACTAGTC
81.5	T+80.5	C+80.5
82.5	GT+80.5	AC+80.5
83.5	CGT+80.5	AAC+80.5
84.5	CCGT+80.5	TAAC+80.5
85.5	GCCGT+80.5	TTAAC+80.5
86.5	GGCCGT+80.5	TTAAC+80.5
87.5	TGGCCGT+80.5	TTTTAAC+80.5
88.5	TTGGCCGT+80.5	GTTTTAAC+80.5
89.5	GTTGGCCGT+80.5	CGTTTTAAC +80.5
90.5	TGTTGGCCGT+80.5	GCGTTTTAAC+80.5
91.5	CTGTTGGCCGT+80.5	CGCGTTTTAAC+80.5
92.5	GCTGTTGGCCGT+80.5	CCGCGTTTTAAC+80.5
93.5	TGCTGTTGGCCGT+80.5	-
94.5	GTGCTGTTGGCCGT+80.5	TACCGGTTTTAAC+80.5
95.5	TGTGCTGTTGGCCGT+80.5	CTACCGGTTTTAAC +80.5
96.5	CTGTGCTGTTGGCCGT+80.5	TCTACCGGTTTTAAC+80.5
97.5	CCTGTGCTGTTGGCCGT+80.5	GTCTACCGGTTTTAAC+80.5
98.5	CCCTGTGCTGTTGGCCGT+80.5	TGTCTACCGGTTTTAAC+80.5
99.5	CCCCTGTGCTGTTGGCCGT+80.5	CTGTCTACCGGTTTTAAC +80.5
100.5	TCCCCTGTGCTGTTGGCCGT+80.5	GCTGTCTACCGGTTTTAAC+80.5
101.5	ATCCCCTGTGCTGTTGGCCGT+80.5	CGCTGTCTACCGGTTTTAAC+80.5
102.5	-	-
103.5	GCTGATCCCCTGTGCTGTTGGCCGTGTTATCTCGAGTTAGTACGACC	CGCGCTGTCTACCGGTTTTAACGCCAATAGGATTACTTACTAGTC
104.5	T+103.5	A+103.5
105.5	GT+103.5	TA+103.5
106.5	-	GTA +103.5
107.5	CGGT +103.5	CGTA +103.5
108.5	ACGGT +103.5	ACGTA +103.5
109.5	CACGGT +103.5	CACGTA +103.5
110.5	CCACGGT +103.5	GCACGTA+103.5
111.5	TCCACGGT +103.5	CGCACGTA+103.5
112.5	CTCCACGGT+103.5	ACGCACGTA+103.5
113.5	CCTCCACGGT 103.5	AACGCACGTA+103.5
114.5	-	AAACGCACGTA+103.5
115.5	CGCCTCCACGGT +103.5	TAAACGCACGTA+103.5
116.5	TCGCCTCCACGGT +103.5	TTAAACGCACGTA+103.5
117.5	-	CTTAAACGCACGTA+103.5
118.5	TATCGCCTCCACGGT +103.5	GCTTAAACGCACGTA+103.5
119.5	TTATCGCCTCCACGGT +103.5	CGTTAAACGCACGTA+103.5
120.5	TTTATCGCCTCCACGGT +103.5	CCGTTAAACGCACGTA+103.5
121.5	ATTTATCGCCTCCACGGT +103.5	ACCGTTAAACGCACGTA+103.5
122.5	-	-
123.5	TTATTTATCGCCTCCACGGT +103.5	GCACCGTTAAACGCACGTA +103.5
124.5	-	AGCACCGTTAAACGCACGTA +103.5
125.5	TTTTATTTATCGCCTCCACGGT +103.5	TAGCACCGTTAAACGCACGTA +103.5
126.5	CTTTATTTATCGCCTCCACGGT +103.5	CTAGCACCGTTAAACGCACGTA +103.5
127.5	ACTTTATTTATCGCCTCCACGGT +103.5	TCTAGCACCGTTAAACGCACGTA +103.5
128.5	TACTTTATTTATCGCCTCCACGGT +103.5	CTCTAGCACCGTTAAACGCACGTA +103.5
129.5	CTACTTTATTTATCGCCTCCACGGT +103.5	-
130.5	ACTACTTTATTTATCGCCTCCACGGT +103.5	AGCTTAGCACCGTTAAACGCACGTA +103.5
131.5	AACTACTTTATTTATCGCCTCCACGGT +103.5	AAGCTTAGCACCGTTAAACGCACGTA+103.5
132.5	GAACTACTTTATTTATCGCCTCCACGGT +103.5	CAAGCTTAGCACCGTTAAACGCACGTA+103.5
133.5	AGAACTACTTTATTTATCGCCTCCACGGT +103.5	GCAAGCTTAGCACCGTTAAACGCACGTA+103.5
134.5	TAGAACTACTTTATTTATCGCCTCCACGGT +103.5	AGCAAGCTTAGCACCGTTAAACGCACGTA+103.5
135.5	G+ 134.5	T+134.5

136.5	CG+134.5	GT+134.5
137.5	GCG +134.5	CGT+134.5
138.5	TGCG +134.5	TCGT+134.5
139.5	CTGCG +134.5	GTCGT+134.5
140.5	GCTGCGT +134.5	GGTCGT+134.5
141.5	GGCTGCGGT +134.5	GGGTCGT+134.5
142.5	CGGCTGCGGGT +134.5	CGGGTCGT+134.5
143.5	CCGGCTGCGCGGT +134.5	CCGGGTCGT+134.5
144.5	GCCGGCTGCGACGGT +134.5	GCCGGGTCGT+134.5
145.5	-	GGCCGGGTCGT+134.5