www.sciencemag.org/cgi/content/full/339/6121/816/DC1

Supplementary Material for

Probing Allostery Through DNA

Sangjin Kim, Erik Broströmer, Dong Xing, Jianshi Jin, Shasha Chong, Hao Ge, Siyuan Wang, Chan Gu, Lijiang Yang, Yi Qin Gao, Xiao-dong Su,* Yujie Sun,* X. Sunney Xie*

*To whom correspondence should be addressed. E-mail: xdsu@pku.edu.cn (X.-D.S.); sun_yujie@pku.edu.cn (Y.S.); xie@chemistry.harvard.edu (X.S.X.)

Published 15 February 2013, *Science* **339**, 816 (2013)
DOI: 10.1126/science.1229223

This PDF file includes:

Materials and Methods

1. Preparation of DNA templates used in the single-molecule experiments

Schematics of the templates are provided in fig. S1-2.

DNA templates for glucocorticoid receptor DNA-binding domain (GRDBD) and BamHI – DNA templates containing a binding site for GRDBD (forward or reverse GRE) (*35*) and 6-bp recognition sequence for BamHI with various inter-site distance (*L*) were made by annealing two complementary oligonucleotides (Shanghai Sangon Biotechnology Company, Shanghai, China). One strand of DNA ('a') was biotinylated at 5' end and had a row of T bases at the 5' end. We positioned the GRE at the free end of DNA and the BamHI binding site at least 25 bp away from the surface-end of the DNA template. See Table S1 for the list of sequences used in the experiments.

DNA templates for GRDBD and hairpin –The DNA templates with a hairpin were designed such that a single-stranded loop (3 or 15 bp) is followed by a varying length of linker DNA and GRE in the double-stranded stem region. The hairpin structure was ligated to a biotinylated DNA using T4 DNA ligase (New England Biolabs) to form the full-length DNA templates. The GRE and spacer sequences used were the same as DNA templates for GRDBD-BamHI. See Table S2 for details.

DNA templates for Lac repressor (LacR) and T7 RNA polymerase (RNAp) – The DNA templates were made by annealing two complementary oligonucleotides of <200 bp (Integrated DNA Technologies, Coralville, IA). One strand of DNA ('a') was biotinylated at its 5' end and had a row of T bases at the 5' end to enhance efficiency of surface attachment. The complementary strand ('b') contained *lac* operator (*lacO*) and a binding site for T7 RNAp. For *lacO*, we used the original operator sequence of 21 bp (*lacO*1, AATTGTGAGCGGATAACAA TT) appearing in the *E. coli* genome. It is positioned at least 25 bp away from the surface-end of the DNA template (fig. S2). As for the T7 promoter, we positioned the promoter at the free end of DNA to aid binding of the RNAp. When *lacO*1 was downstream of the promoter (Fig. 3, fig. S5), we included an upstream AT-rich sequence from -40 (+1 is the transcription start site) to increase T7 promoter strength (*36*). When *lacO*1 was upstream of the promoter (Fig. 4A-B, fig. S15), we used only a 27-bp-long promoter sequence. See Table S3-5 for the list of sequences used in the experiments.

DNA templates for LacR and BamHI or EcoRV – All DNA templates were made by annealing two complementary oligonucleotides of <200 bp (Integrated DNA Technologies, Coralville, IA). Similar to the LacR-T7 RNAp templates, *lacO*1 was positioned >25 bp away from the biotinylated end and a 6-bp-long recognition sequence of BamHI or EcoRV was inserted near the free end of DNA (Table S6).

DNA templates for GRDBD and nucleosome – We used a high-affinity and sequence-specific histone binding sequence, Widom 601 (*34*), as a primary nucleosome reconstituted sequence and subcloned it into pET-28a between EcoRI and XbaI sites. Using this plasmid as a PCR template,

we prepared linear DNA containing W601, spacer sequence and a binding sequence for GRDBD (GRE). For PCR, we used a biotinylated primer together with primers GR-8 to GR-27 (Table S7). Biotinylated primer sequence: biotin-AAGCTTGTCGACGGAGCTCGAATTCGTGAGCGCTC ACAATTTG

2. Preparation of proteins

Purification and fluorescent labeling of GRDBD – The gene sequence encoding GRDBD was synthesized (Beijing AuGCT DNA-SYN Biotechnology, China) and inserted between NdeI and XhoI sites in pET-28a plasmid for the N-terminal fusion of $(His)_6$-tag. GRDBD was overexpressed in *E. coli* BL21 (DE3) strain and induced with 0.5 mM IPTG overnight at 18 °C. The harvested cells were resuspended in buffer A (20 mM Tris, pH 7.5, 1 M sodium chloride) and lysed by sonication. After centrifugation, the supernatant was first loaded on a nickel chelating column (5 ml Hi-Trap HP column, GE Healthcare/Amersham) and then washed with buffer A supplemented with 100mM imidazole. The proteins were further purified by a gel filtration column (120 ml Superdex-75, GE Healthcare/Amersham) in Buffer B (20 mM Tris, pH 7.5, 300 mM sodium chloride, 5 mM DTT) and stored in 50 mM Tris, pH 7.5, 200 mM sodium chloride, 5mM DTT, 50% glycerol at -20 °C. Purified GRDBD was labeled with Cy3b–maleimide at the surface cysteines following the same approach used for the fluorescent labeling of LacR. The labeling efficiency was quantified as a dye-to-monomer ratio of 0.4-0.6 to 1.

Purification and fluorescent labeling of wild-type LacR – A gene encoding wild-type LacR (360 amino acids in monomer) was PCR-amplified from pET-30a plasmid and inserted into BamHI site in the same plasmid for over-expression under T7 promoter and for 6x histidine tag at the N terminus of the repressor. Small-scale overnight culture of *E. coli* BL21 (DE3) was used to inoculate 1-liter of LB with kanamycin in the 37 °C shaker. When OD_{600} reached ~0.6, the culture was induced with 1 mM IPTG. The induction continued for 6 hours at 37 °C. A conventional purification protocol for histidine-tagged proteins was used with lysis buffer (25 mM Tris pH 7.5, 300 mM sodium chloride, 2 mM imidazole). After purification, we removed the N-terminal tag by digestion reaction with enterokinase (New England Biolabs) to restore DNA binding activity of LacR. Finally, LacR was dialyzed to a storage buffer (50 mM potassium phosphate pH 7.2, 100 mM sodium chloride, 2mM EDTA, 1 mM DTT).

We used maleimide chemistry to label cysteine residues in LacR with Cy3b dye (Cy3b-maleimide, GE Healthcare Biosciences, Piscataway, NJ). We performed the protein labeling while LacR was bound to the cobalt resin (*37*). Dye molecules (5-10 molar excess) were incubated with the protein mixture overnight at 4 °C with constant mixing. Next day, we continued with the purification protocol for histidine-tagged proteins, which also removed unreacted free dye molecules. The labeling efficiency was assayed by UV-Vis spectroscopy at 280 nm (for protein) and at 550 nm (for Cy3b). Dye to monomer ratio was 0.3-0.5 to 1, and we expect 1-2 labels per tetrameric LacR.

We verified the DNA binding activity of labeled and unlabeled LacR by electrophoretic mobility shift assay (*38*). We used DNA templates whose sequences were the same as the ones used in the single-molecule experiments. 5' end of a DNA strand was labeled with Cy5 for visualization in the Typhoon imager (GE Healthcare Biosciences). Scanning for Cy5 fluorescence signal, we observed no difference in mobility of DNA bound either with labeled or

unlabeled LacR. We also observed that the shifted band for the DNA-protein complex has both Cy3 and Cy5 signals, from labeled LacR and DNA, respectively. All DNA formed the complex when LacR monomer-to-DNA ratio was 3:1, indicating that most of our LacR proteins are active in binding.

Purification and fluorescent labeling of dimeric LacR – The C-terminal end of LacR participates in protein oligomerization. When YFP was fused to the C-terminal end, LacR was not able to tetramerize but remained as a dimer (*39*). It was also shown that a heptad repeat of leucines at the C terminus is crucial for the assembly of a tetrameric LacR (*40*). Based on these findings, we made a mutant of dimeric LacR by inserting a 6x histidine tag at the C terminus and deleting the last 11 amino acids of LacR. We PCR-amplified the *lacI* gene without the last 36 bases (bp) and cloned in between NcoI and XhoI sites in the pET-28a plasmid. Dimeric LacR was labeled with Cy3b-maleimide and purified following the same experimental procedure used for the wild-type LacR.

T7 RNA polymerase (RNAp), EcoRV and BamHI – We purchased these proteins from New England Biolabs. We checked purity of the stock protein solutions through SDS-PAGE gel analysis. Stock concentration: T7 RNAp = 0.7 μM, EcoRV = 190 nM, BamHI = 2.5 μM. Based on the titration experiments, we used 70 nM of T7 RNAp, 9.5 nM of EcoRV, and 50 nM of BamHI as the amount of proteins resulting in the maximum change in k_{off}, i.e. $k_{3\rightarrow2}$.

Nucleosome reconstitution for GRDBD-nucleosome experiments – Histone core particles were prepared with calf core histones (H5505, Sigma-Aldrich). Nucleosomes were reconstituted by mixing 1 μg of biotinylated DNA and 1 μg of histones in 30 μl of buffer containing 10 mM Tris, pH 7.6, 2 M sodium chloride and then using stepwise dialysis through decreasing salt concentrations to 0.2 M. The final products were incubated at 37 °C for two hours, with the purpose of shifting the population of nucleosome positions on the DNA to a single unique position (*41*).

3. Single-molecule fluorescence assay to measure dissociation of GRDBD or LacR

The single-molecule experiments were performed inside a microfluidic channel. Details of a flow cell construction can be found elsewhere (*11, 42*). The only difference from the previous procedure was passivation of the glass slides with PEG polymers (but no biotinylation) to minimize protein adsorption on the upper surface during protein titration experiments.

After constructing a flow cell with DNA templates tethered on the surface (fig. S1-2), ~1 nM of Cy3b-labeld GRDBD or LacR was infused into the flow cell and incubated for <3 minutes. We washed out the excess proteins by infusing imaging buffer containing oxygen scavengers (see below for the buffer contents). Once enough buffer was infused to remove all the free proteins, we started fluorescence imaging as follows.

Details of the single-molecule imaging can be found in previous publications (*43, 44*). In short, we used an inverted microscope (Olympus IX70, Olympus America Inc., Melville, NY) for objective-type total internal reflection (TIR) illumination. The focused beam of a 532-nm diode laser (UGA-250; LambdaPro, Beijing) was directed at the edge of the back focal plane of a 60x objective lens (numerical aperture 1.45; PlanApo, Olympus). Fluorescence was collected

through the same objective lens. After passing through emission filters, the fluorescence was imaged on a back-illuminated electron multiplying charge coupled device camera (iXon*EM*+897, Andor Technology, South Windsor, CT). Typically, the field-of-view contained 250-450 diffraction limited spots in ~75x75 μm^2. Images were acquired at every second for 15-20 minutes. We adjusted the excitation power below 10 W/cm^2 to minimize photobleaching of dyes in the oxygen-free environment created by fresh oxygen scavenger system.

The following are the imaging buffers used. In GRDBD-BamHI experiments, we used 20 mM Tris pH 7.8, 50 mM sodium chloride, 5 mM calcium chloride, 500 μg/mL BSA (New England Biolabs), 10 % (v/v) glycerol, 0.8 % (w/v) glucose, 1 mg/mL glucose oxidase, 0.1 mg/mL catalase, 2 mM trolox. In GRDBD-hairpin and GRDBD-nucleosome experiments, we used 20 mM Tris pH 7.8, 50 mM sodium chloride, 5 mM DTT, 500 μg/mL BSA, 10 % (v/v) glycerol, 0.8 % (w/v) glucose, 1 mg/mL glucose oxidase, 0.1 mg/mL catalase, and 2 mM trolox. In LacR-T7 RNAp experiments, we used 10 mM Tris pH 8.0, 10 mM magnesium acetate, 10 mM potassium chloride, 500 μg/mL BSA, 5 % (v/v) glycerol, 0.8 % (w/v) glucose, 1 mg/mL glucose oxidase, 0.1 mg/mL catalase, and 2 mM trolox. In LacR-endonuclease experiments, we used 50 mM Tris pH 7.5, 100 mM sodium chloride, 5 mM calcium chloride, 500 μg/mL BSA, 5 % (v/v) glycerol, 0.8 % (w/v) glucose, 1 mg/mL glucose oxidase, 0.1 mg/mL catalase, and 2 mM trolox (all reagents were purchased from Sigma-Aldrich, otherwise noted).

Imaging buffer was degassed and carefully filtered through 0.2-μm syringe filters on a daily basis. Fresh oxygen scavenger system was added right before each measurement. During imaging, we kept the imaging buffer flowing at 1 mL/hr by a syringe pump (PhD 2000, Harvard Apparatus, Holliston, MA).

4. Single-molecule data analysis to measure dissociation rate constant

Here, we explain the analysis for LacR-T7 RNAp experiments, but the same applies to any *in vitro* experiments conducted with other proteins.

Images were analyzed using an algorithm written in Matlab (The Mathworks, Inc., Natick, MA), in which positions of fluorescently labeled LacR were tracked by calculating the center-of-mass of intensity profiles. Assuming the fluorophores were distributed in space and their positions did not change within 1-2 pixels due to mechanical drift, the algorithm identified a fluorophore in consecutive frames only if its centroid remained within 1-2 pixels. Because the fast processing of ~400 molecules per frame was more important than higher spatial accuracy, the centroid positions of fluorophores were determined by center-of-mass rather than by Gaussian fitting of their intensity profiles. Often, fluorophores appeared at a new position in the middle of the movie and stayed, but we ignored those fluorophores in the analysis. We only considered those fluorophores that appeared from the initial 10 frames of the movie to count how long they stayed at their initial positions. For data acquired in the absence or in the presence of the second DNA binding protein (e.g., T7 RNAp), the histogram of dissociation time displayed a single exponential decay, consistent with the single-step dissociation kinetics of LacR (fig. S3). The dissociation rate (k_{off}) of LacR was taken as the inverse of the average dissociation time or mean first passage time, τ (see Supporting Text). τ/sqrt(N), where N is the number of LacR molecules analyzed, was taken as a standard error for τ and propagated to errors for dissociation rate (error bars in main-text plots showing normalized k_{off}). For data acquired at intermediate T7 RNAp concentration, errors were calculated as standard error of means. We noticed these errors

are comparable to those determined by bootstrapping. Hyperbolic fitting of the titration curves was performed in Igor Pro (WaveMetrics, Lake Oswego, OR). For the curve fitting, we used a function, $1 + ax/(b + x)$, where $(1 + a)$ corresponds to the plateau values, or $k_{3\rightarrow2}$, and b corresponds to K_d of T7 RNAp ($K_{d,B}^{A}$) reported in Fig. 3D (middle). The curve fitting was based on the Levenberg-Marquardt algorithm for the chi-square fit and the estimated standard deviations of the fit coefficients (a, b) were used as error bars.

5. *In vivo* experiment based on the plasmid (data presented in fig. S16)

The plasmid construct was based on pET-21b vector (Novagen, EMD Biosciences, Inc., Darmstadt, Germany) with ampicillin resistance. The gene encoding rapidly maturing (~7 min) yellow fluorescent protein, Venus, was inserted in between NdeI and XhoI recognition sites on the plasmid. We removed the original T7 promoter and *lacO* region via the BglII and XbaI restriction and inserted a synthesized sequence carrying $lacO_{id}$ and T7 promoter. An EcoRI restriction site was left between $lacO_{id}$ and T7 promoter and used to add DNA bases between the two protein binding sites. The sequences of DNA spacers were random with about 60% AT in the base composition. See Table S8 for details. To increase LacR expression level (*45*), LacR gene was mutated to LacRq through QuikChange site-directed mutagenesis (Agilent Technologies, Santa Clara, CA). T7 RNAp was expressed from the genome in E. *coli* strain BL21-AI (Invitrogen, Carlsbad, CA). Expression of T7 RNAp from the genome offered tighter control of its expression because it is under the arabinose-inducible *araBAD* promoter.

To measure the expression level of Venus, monoclones of transformed BL21AI carrying plasmids with different lengths of DNA spacers were picked and grown overnight (12 hours) in 1 mL LB medium with 100 μg/mL ampicillin at 37 ºC. The overnight cultures were diluted 1:2000 and grown to OD600 ~0.2 in LB cultures containing 100 μg/mL ampicillin and 0.08 % L-arabinose at 30 °C. The cells were then pelleted and resuspended in PBS buffer. After measuring OD600, the cells were lysed by adding 1 mg/mL cetyltrimethylamnionium bromide (CTAB) and 0.5 mg/mL sodium deoxycholate. The lysed cells were spun at 13000 rpm, 4 °C for 10 minutes and 100 μL of supernatant was taken for fluorescence measurement at 527 nm using a MD5 microplate reader (Molecular Devices, Sunnyvale, CA). All the chemicals were purchased from Sigma-Aldrich.

6. *In vivo* experiment based on the chromosome (data presented in fig. S15-16)

BL21-AI strains with different DNA spacers between *lacO* and T7 promoter on the chromosome were created through a two-step homologous recombination method using a *cat-sacB* counter-selectable cassette. Substitutions of genes were made using lambda Red recombinase method (*46*). First, the *cat-sacB* gene was amplified from plasmid pEL04 (a generous gift from Dr. Donald Court, National Cancer Institute at Frederick (*47*)) using primers A1 and A2, and inserted into the chromosome to replace the region between *lacI* and the end of *lacO*2. As a result, all three native LacR binding sites were removed. The chromosomal insertion was done by electroporation of the purified PCR product into the Bl21-AI strain carrying pKD46 and screening by chloromycetin. Second, *lacZ* gene with inactivated *lacO*2, or $lacZO2^{-}$ (*48*), was inserted into pET-21b between NheI and XhoI recognition sites. Similarly, we placed *lacO* in the

upstream of T7 promoter with a certain separation (Table S8). These plasmids were used as templates for PCR amplification of DNA fragments for second round of recombination, with primers B1 and B2. The purified PCR products were electroporated into the strains created in the first step to replace the *cat-sacB* and counter-selected using 10 % sucrose. A list of primers used in the construction is available in Table S9. To provide LacR, the strains with successful recombination were further transformed with another plasmid containing a *lacIq* gene.

Similar to the plasmid assays, overnight LB cultures of the different strains were diluted 1:2000 and grown to OD600 ~0.2 in LB containing 0.08 % L-arabinose at 30 °C. *lacZ* expression was measured by the Miller assay (*26*). Briefly, cells were pelleted and resuspended in wash buffer (100 mM Na_2HPO_4, 20 mM KCl, and 2 mM $MgSO_4$). After measuring OD600, we lysed the cells using wash buffer containing 0.8 mg/mL CTAB, 0.4 mg/mL sodium deoxycholate, and 5.4 µL/mL beta-mercaptoethanol. 10 µL of cells were added to the substrate solution (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 1 mg/mL o-nitrophenyl-β-D-galactoside (ONPG), and 2.7 µL/mL β-mercaptoethanol) and allowed to react for 50 minutes, then quenched by stop solution (1 M sodium carbonate). The samples were spun at 13000 rpm at 4 °C for 10 minutes and 150 µL of supernatant was taken for OD420 measurements using the MD5 microplate reader.

7. Construction of structural models

To have an overview of the ternary complexes, models of the dsDNA and two binding proteins were generated based on high-resolution crystal structures of individual protein-DNA structures. For each protein, the crystal structure contains a short DNA strand bound to the protein. During the flow-cell experiments, dsDNA templates with a length of ~200 bp were used. Thus, there are regions of dsDNA where experimentally determined structure is not available. These include the region from the surface linker to the first protein (e.g., LacR), the inter-site region between the two proteins, and finally, the DNA from the second protein to the end of the dsDNA. These missing pieces of straight B-form DNA were created using online DNA-tools in the model.it® server based on the exact sequence used during our experiments (*49*). For each strand, additional 6-10 bases were added in both 3' and 5' end, so that they can overlap with the DNA sequences in the DNA-protein binary structures. All structures were built using the software WinCoot (*50*).

Starting from the known crystal structure of the LacR-DNA complex, PDB ID: 1LBG (*16*) and 2PE5 (*33*), the generated B-DNA segment from the surface linker was first added. The orientation of the DNA segments was obtained by superimposing the overlapping bases from the generated DNA with the corresponding bases in the protein-DNA crystal structure, using the LSQ superimpose function in WinCoot. In the same manner, the generated B-DNA for the inter-site region, the second protein-DNA complex, and the remaining segment of dsDNA were added to the model. Images of the finished complex were created using the software Pymol (DeLano Scientific, San Carlos, CA).

Note that in our *in vitro* experiments, T7 RNAp is thought to be in a conformation similar to that of the initiation complex structure (PDB ID: 1CEZ) (*51*). However, because only half of the DNA under the protein is seen in the crystal structure, it is hard to build a complete model using that structure. To avoid this problem, the model was built using the transition complex between initiation and elongation (PDB ID: 3E2E) (*18*). This is the complex when the RNAp has a 3-nucleotide-long nascent RNA in its active site. Yet, the RNAp keeps the same interaction with promoter sequences, and thereby the location of the protein along the DNA would not be vastly

different from the initiation complex. Of course, the exact binding conformation may differ as suggested previously (*24*), with twisting of the downstream dsDNA upon addition of initiating nucleotide. This adds extra uncertainty to the exact binding conformation and location of the protein along the DNA, which in turn would also affect the rotation angle of the bent DNA. Bearing this in mind, the complexes provide a reasonable estimate of intermolecular distances, relative orientations, and relative angles between the molecules in the complexes. The complexes also provide an opportunity to explore electrostatic surface effects, as well as comparison between different DNA-binding proteins.

Electrostatic surface representation was built as follows. Individual proteins in the generated complexes were saved as separate PDB files, where DNA and water molecules were removed. Using the PDB2PQR Server v.1.6 (*52*), we determined the charge of each atom at pH 7.5. Protein pKa values were determined using the software PROFKA (*53*). Electrostatic potentials were determined using the ADBS tools plug-in in Pymol (*54*). All settings were left at their default values, and the ions option was set to account for 20 mM sodium chloride.

We applied the same procedure to build the structural models for GRDBD-BamHI (Fig. 1A) and GRDBD-nucleosome (fig. S17).

8. Molecular dynamics simulation

DNA sequence used in the MD simulation is 5'-GAGATGCTAACCCTGATCGCTGATTCCTT GGAC-3'. Two different systems were simulated including one normal DNA and one constrained DNA, in which a constraint on the 15th base pair (GC, L=0 in Fig. 2B) being added to increase the distance between the N1 atom of the Guanine furan ring and the N3 atom of the Cytosine furan ring by 0.5 Å. The detailed constraint setting of the simulation systems are listed in Table S10.

All MD simulations were conducted using AMBER11 suite of programs (*55*). The SPC/E model (*56*) was used to describe the explicit water solvent. Nucleic acid force field parameters were taken from the AMBER ff10 parameter set (*57*). In each system, the canonical B-form DNA was used as the initial structure and then immersed into a cubic box containing 9929 water and 64 sodium counter ions. The SHAKE algorithm (*58*) was used in order to restrain all covalent bonds involving hydrogen. Therefore, all dynamics employed an integral time step of 2 fs. Periodic boundary conditions were used and the particle-mesh Ewald (PME) (*59*) with a direct space cutoff of 10.0 Å was utilized to treat long-range electrostatic interactions.

For each system, the simulation procedure included the energy minimization, heating up and a production equilibrium simulation using the NPT ensemble. Firstly, the system was minimized through 500 steps of steepest descent minimization and a following 500 steps of conjugate gradient minimization with DNA being fixed using harmonic restraints. Then, the restraints on DNA were released and the system were minimized using 1000 steps of the steepest descent and then 1500 steps of conjugate gradient minimization. In order to further relax the systems (especially the solvent molecules), they were first heated to 360 K and equilibrated at the temperature for 300 ps, then they were cooled down to 300 K. Finally, production runs were performed and the temperature of all systems was maintained at 300K using the Langevin dynamics with a friction coefficient of $5ps^{-1}$. The pressure of the system was adjusted to 1 atm by Berendsen weak-coupling algorithm (*60*) with a relaxation time constants of 2.0 ps.

Supplementary Text

*** Difference between GRDBD-BamHI and GRDBD-hairpin experiment**

GRDBD is positioned downstream of BamHI in the (forward) GRDBD-BamHI experiment, but is upstream of the hairpin loop in the (forward) GRDBD-hairpin experiment. Hence, direct comparison of the phase between two experiments is not advised. Yet, it is worthwhile to note that the phase shift upon reversing GRE is 4 bp in both cases.

*** Specific versus nonspecific binding of T7 RNAp**

Our assay is based on the assumption that the labeled LacR and T7 RNAp are specifically bound to their binding sites, i.e., *lacO* and T7 promoter, respectively. We verified the assumption from a few control experiments. First, we repeated the experiment with 120-bp-long dsDNA that did not contain a *lacO*1 site. We not only observed fewer fluorophores in the same-sized field of view, but the fluorophores stayed less than a minute and washed out with the imaging buffer. Therefore, in our experimental condition, the nonspecific binding of LacR makes little contribution to the final histogram, from which the mean first passage time ($1/k_{off}$) was calculated. Furthermore, when a DNA template with *lacO*1 but no target sequence for the second protein was used for the titration with the second protein, the nonspecific binding of the second protein yielded minimal change in k_{off} of LacR (fig. S4). Hence, the observed change in the k_{off} of LacR resulted from the specific binding of T7 RNAp on its promoter sequence.

*** Definition of *L* for the LacR-T7 RNAp experiments (Fig. 3)**

In the case of LacR, endonucleases, and GRDBD, it is straightforward to identify DNA bases that interact with the proteins in the crystal structures. However, for the T7 RNAp-DNA binary complex, the absence of a crystal structure with clear downstream dsDNA makes it difficult to determine *L* in the LacR-T7 RNAp templates, where *lacO*1 is positioned in the downstream of the T7 promoter. We used a structure showing a transition state complex between the initiation complex and the elongation complex (*18*) to make a crude estimation of the edge of T7 RNAp on the downstream DNA. We decided to define *L* from +20 (+1 is the transcription start site), such that *lacO*1 spans +20~+41 nucleotide positions (+1 is transcription start site for T7 RNAp) on the template of $L = 0$. Although this is a very conservative definition of the end-to-end distance between LacR and T7 RNAp, our structural models and experimental observation indicate that this definition yields $L = 0$ to be the relative distance, where LacR and T7 RNAp are pretty close to each other. In particular, the structural model for $L = -1$ predicts physical interference between two proteins on the DNA. Plus, at $L = -1$, the DNA segment coming out of T7 RNAp seems to crash with the tetrameric LacR. See Supplementary Text on 'Structural Model' for further discussions.

*** Phase shift upon reversing *lacO*1 in LacR-T7 RNAp template**

Similar to the phase shift observed upon reversing non-palindromic GRDBD binding sequence in GRDBD-BamHI (Fig. 1B) and GRDBD-hairpin (Fig. 1D) experiments, we also observed the

phase of the oscillation shifted by ~5 bp when we reversed the non-palindromic *lacO*1 sequence in LacR-T7 RNAp experiments (fig. S5).

Asymmetric interaction between LacR and DNA within the operator region is well documented in several studies (*17, 61, 62*). Especially, the structural data of LacR-*lacO*1 binary complex shows that the right half-site of the operator contains significant local deformations while the left half-site is comparable with undistorted B-form DNA (*17*). Due to such asymmetry in DNA distortion, reversing *lacO*1 would change the phase of the coupling oscillation as a function of the separation between LacR and T7 RNAp.

*** Examining various protein pairs**

After studying allosteric interactions between LacR-T7 RNAp (Fig. 3), we used EcoRV and BamHI in place of T7 RNAp and studied their allosteric coupling with LacR through DNA. EcoRV and BamHI are examples of the most well-studied type II endonucleases, showing high sequence specificity and strong binding affinity (*63*). Although showing similarity in the thermodynamic properties of DNA recognition, EcoRV and BamHI display profound differences in their DNA complexes. First, EcoRV approaches DNA from the minor groove side and narrows the major groove, while BamHI embraces the major groove side of the DNA and widens it (*14, 15*). Moreover, binding of EcoRV introduces a ~50° kink in the center of the cognate DNA sequence, but when bound to a non-cognate sequence, the enzyme merely interacts with the DNA backbone and unwinds the central part of the DNA binding region (*15*). On the contrary, DNA retains a regular B-form-like conformation without major bends or kinks in the BamHI complex. Small deviations from the regular B-form DNA were observed as a small twist at the central base (AT) and widening of the major groove by ~3 Å (*14*).

DNA templates were designed to contain *lacO*1 and 6-bp recognition sequences for EcoRV or BamHI with separation, L (see Table S6). From titration experiments, we chose 9.5 nM of EcoRV and 50 nM of BamHI as the amount of proteins resulting in the maximum change in k_{off}, i.e. $k_{3\rightarrow2}$. To promote specific binding of endonucleases without the DNA cleavage reaction, imaging buffers included 5 mM calcium chloride instead of magnesium (*64*).

Interestingly, k_{off} of LacR oscillated as a function of L in both cases (fig. S6). The oscillation occurs with a periodicity of 10 bp with the amplitude dampening within ~30 bp between two binding sites. These restriction enzymes are smaller in size relative to T7 RNAp (~50 kDa rather than 100 kDa) and differ in surface charge distribution. Also, the conformational changes in DNA induced by specific binding of the proteins are different in each complex (*32, 65*). Hence, it is hard to argue that the observed oscillations are specific to a special protein pair and that it is due to 'apparent' melting or bending of the duplex DNA by T7 RNAp or EcoRV, respectively. We note that the phase of the oscillation differs by ~5 bp between LacR-EcoRV and LacR-BamHI. According to our mechanistic interpretation (Fig. 2C), such a difference may arise from the fact that EcoRV and BamHI bind at the opposite side of the groove and yield different DNA distortions.

We further investigated whether the size of LacR contributed to the observed change in its dissociation rate when the second protein interacted with DNA at a distal site. In particular, tetrameric LacR is rather large (~150 kDa), such that it may exhibit steric hindrance with a neighboring protein on the same DNA. We repeated experiments on LacR-BamHI system using Cy3b-labeled dimeric LacR. Oscillation in k_{off} of dimeric LacR upon binding of BamHI was similar to the one measured with tetrameric LacR (fig. S6).

Taken all together, if any hydrogen bonding, van der Waals, or electrostatic interactions were present between two proteins, they would differ in the various protein pairs we examined. The generality in observing the 10-base oscillations with various protein pairs makes it difficult to conceive that protein-protein interactions make a major contribution to the oscillation.

*** Salt dependence experiment for LacR-T7 RNAp**

As a control experiment, we varied potassium chloride concentrations in the imaging buffer for measuring k_{off} of LacR in the presence of T7 RNAp. Varying the ionic strength may affect protein-protein or protein-DNA interactions as well as physical properties of DNA. Debye screening length, a characteristic distance for electrostatic interactions, is 3-0.3 nm for 10-1000 mM sodium chloride (*66*), indicating that electrostatic interactions within ~3 nm are permitted at 10 mM sodium chloride. Increasing the salt concentration to 1M, electrostatic interactions between proteins, if there was any, will be eliminated and protein-DNA interactions will become weaker. Furthermore, the negative charges in the phosphate backbone of DNA molecules may well be shielded at a high salt concentration. In particular, it is known that the persistence length of DNA decreases with increasing monovalent salt concentrations within a range of 1-1000 mM (*67, 68*).

We observed that increasing the salt concentration from 10 mM to 1000 mM yielded a small reduction in k_{off} on a given DNA template but produced a minimal effect on the observed oscillation in the k_{off} (or $k_{3\rightarrow2}$) as a function of separation (L) (fig. S7).

*** Tuning allosteric coupling with different linker DNA for LacR and T7 RNAp**

First, a nick was created in the middle of the linker region. Similar to the single-molecule work on gyrase (*69*), in which a nick was created on a DNA strand to release supercoiling produced by a gyrase activity, a nick on the linker DNA would offer a rotational freedom between two DNA-protein interaction units. The design of DNA template and their sequences are available in Table S4. We observed that k_{off} remained constant with increasing concentration of T7 RNAp on $L = 5$ and 10 templates (fig. S8). It did not originate from the inability of T7 RNAp to bind on the nicked DNA, as we confirmed with separate experiments that T7 RNAp was able to bind on the DNA. Hence, the result suggests that k_{off} of LacR was no longer affected by the presence of T7 RNAp on the other end of the DNA.

Second, the spacing region of DNA was forced to form a 'bubble' by mismatching DNA base sequences on two DNA strands (fig. S9, Table S4). The oscillation in k_{off} as a function of L flattened out as the number of mismatched bases in the linker region increased. Overall, including mismatched bases in between two protein-binding sites produced similar results as obtained using the nicked DNA templates. However, the effect was smaller than that of nicked DNA case, considering error bars. It is difficult to discuss bending and twisting deformations in the mismatched DNA region, yet it is possible that any mechanical stress would be stored or buffered in a different manner in the mismatched dsDNA, compared to a regular dsDNA. In fact, these data suggest that having an intact phosphate backbone and successive base pairing are crucial in observing the long-range allosteric interactions between two proteins.

Lastly, sequence composition of the linker DNA was adjusted (Table S4). Our initial measurement (Fig. 3) was based on the random choice of sequences in the linker region. When this region of DNA was composed of all A or T's, k_{off} was similar to or greater than the value

measured from arbitrary linker sequence. On the contrary, when the linker was composed entirely of G's and C's, the oscillation in k_{off} disappeared, regardless of the separation, L (fig. S10). Again, the diminishing effect was smaller than that of the nicked DNA case. Recent studies on the mechanical properties of DNA illustrated that DNA with high GC content has longer persistence length than DNA with 50% GC content (*70*). However, considering the fact that the linker DNA is shorter than the persistence length, it is more likely that higher stacking energy in GC-rich linker DNA functions as a blockage in the transmission of distortion as explained in Fig. 2C. The information about sequence-dependent properties of DNA will be helpful to deepen our understanding of allosteric interaction between two proteins through DNA.

*** Structural models**

In order to elucidate oscillations occurring every 10 bases of separation (L), we built structural models for the ternary complexes. Crystal structures of each protein-DNA complex and models of B-form DNA composed of exact base sequences used in our experiments were employed to build the models. Because a crystal structure of T7 RNAp initiation complex does not contain information about downstream dsDNA, we used another crystal structure that most likely resembles the complex present during our experiment. Also, the structural models are limited to B-form DNA and, hence, are unable to display any conformational changes in the linker DNA. Nevertheless, the models provide us with a visual aid to examine different geometries in different complexes with varying separation distances (L). For instance, when T7 RNAp moves away from LacR (increasing L), the model predicts that T7 RNAp rotates around the helical pitch of DNA. The models also highlight differences in the protein shape and the conformation of DNA when bound with two proteins.

We calculated the electrostatic surface charge of the proteins in order to indicate different patterns of surface charges in each protein pair. The electrostatic surface charge was calculated considering ionic strength in the imaging buffer used in each experiment (details are provided in Materials and Methods). Under our experimental conditions, the surface of LacR is mostly negatively charged, except the DNA-binding domain which is positively charged. The surface of EcoRV is mostly positively charged, and that of BamHI is mostly negatively charged (fig. S11). The opposite electrostatic surface charges of EcoRV and BamHI make it difficult to argue that the electrostatic interaction between any of these proteins and LacR resulted in the similar oscillations in the stability of LacR as a function of separation, L (fig. S6).

Lastly, we explain our definition of L based on structural models. Structural models are based on dimeric LacR because dimeric LacR bound to DNA has a better structure data than that of the tetrameric LacR. Even considering this difference, the structural model shows little physical collision between DNA segment upstream of T7 RNAp and tetrameric LacR at $L = 0$.

*** A mechanistic interpretation of DNA allostery based on DNA distortion**

1. Data fitting

The damped oscillations of k_{off} are fitted by $f(x) = a\exp(-x/b)\cos(cx+d)+e$ using a nonlinear least square fit in Matlab (Mathworks, Inc). Fig. S12 shows the fitting to the measured k_{off} of GRDBD as a function of its distance from a DNA hairpin, with forward GRE (red) and reverse GRE (blue) sequences (shown in Fig. 1D). As the reverse GRE data have only one period, it is not possible to evaluate its decay constant b by free fitting. In this case, we assume its decay

rate is the same as that of the forward GRE, and leave the rest of the parameters for free fitting. The fitting results are given in Table S11. Both oscillations show a period close to 10 bp, the helical pitch of double-stranded DNA. Their phase shift is 153°, corresponding to a difference of ~4 bp.

2. Why phase shift upon reversing a DNA-binding protein with non-palindromic binding sequence?

We follow the notations used in the main text. Assume protein A and B are separated by L (bp) on DNA and DNA-bound protein B spans N (bp). Then, for protein B, the energetic coupling term at its binding site is $\sum_{i=1}^{N} \delta R_B^A(L+i)\delta R_B^B(i)$, where i denote the i-th base pair within the binding site. After reversing the binding sequence of B, this energetic coupling term becomes $\sum_{i=1}^{N} \delta R_B^A(L+i)\delta R_B^B(N+1-i)$. When the binding sequence of B is nonpalindromic, $\delta R_B^B(i) \neq \delta R_B^B(N+1-i)$, at least for some i. Therefore, the phase of the oscillation would shift.

3. Relating to the measured coupling energetics

The energetic coupling $\Delta\Delta G_{AB}^{\circ}(L)$ between protein A and B is given by the sum of two terms:

$$\Delta\Delta G_{AB}^{\circ}(L) = f(\delta R_A^A, \delta R_A^B(L)) + f(\delta R_B^A(L), \delta R_B^B)$$

where $f(\delta R_A^A, \delta R_A^B(L))$ represents the coupling at the binding site of A, *i.e.*, the interaction between the distortions induced by binding of A and binding of B at L distance away. Likewise, $f(\delta R_B^A(L), \delta R_B^B)$ represents the coupling at the binding site of B. In both terms, the subscripts indicate where the distortion occurs (*i.e.*, binding site of Protein A or B) and the superscripts indicate the binding protein which induced the distortion.

When the energy coupling term $\Delta\Delta G_{AB}^{\circ}(L)$ is small, according to the Taylor expansion, we have

$$f(\delta R_A^A, \delta R_A^B(L)) \approx f(0,0) + \frac{\partial f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^A}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \delta R_A^A + \frac{\partial f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^B(L)}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \delta R_A^B(L)$$
$$+ \frac{1}{2}\frac{\partial^2 f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^{A^2}}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \delta R_A^{A^2} + \frac{1}{2}\frac{\partial^2 f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^B(L)^2}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \delta R_A^B(L)^2$$
$$+ \frac{\partial^2 f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^A \partial \delta R_A^B(L)}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \cdot \delta R_A^A \delta R_A^B(L)$$

The interaction function f should vanish when each of the distortion δR vanishes, *i.e.*

$$f(\delta R_A^A, 0) = 0,\ f(0, \delta R_A^B(L)) = 0.$$

Taking the derivative with respect to the other variable of the function, we also prove that the 2nd to 5th terms in the Taylor expansion equal 0. Therefore, we have:

$$f(\delta R_A^A, \delta R_A^B(L)) \approx \frac{\partial^2 f(\delta R_A^A, \delta R_A^B(L))}{\partial \delta R_A^A \partial \delta R_A^B(L)}\Big|_{\delta R_A^A=\delta R_A^B(L)=0} \cdot \delta R_A^A \delta R_A^B(L).$$

Similarly,

$$f(\delta R_B^A(L), \delta R_B^B) \approx \frac{\partial^2 f(\delta R_B^A(L), \delta R_B^B)}{\partial \delta R_B^A(L) \partial \delta R_B^B}\Big|_{\delta R_B^A(L)=\delta R_B^B=0} \cdot \delta R_B^A(L) \delta R_B^B .$$

Therefore, we obtain Eq. 3:

$$\Delta\Delta G_{AB}^{\circ}(L) \propto [\delta R_A^A \delta R_A^B(L) + \delta R_B^A(L) \delta R_B^B]$$

*** Thermodynamic model**

The thermodynamic model explains the underlying principles in our measurements. We keep the same notations as indicated in Fig. 3C.

1. Mean first passage time

We need to calculate the mean first passage time, $\langle T_1 \rangle$, hitting State 0 or State 2, starting from State 1. Starting from State 1, it will first wait for an exponential time averaged, $1/(k_{1\to0} + k_{1\to3}\cdot[B])$, then jump to either State 0 or State 3 according to the ratio of their reaction constants. If it jumps to State 3, then it still needs to spend the amount of time, $\langle T_3 \rangle$, which is just the mean first passage time hitting State 2 or State 0, starting from State 3. Hence, $\langle T_1 \rangle$ would be the summation of $1/(k_{1\to0} + k_{1\to3}\cdot[B])$ and the probability-weighted mean first hitting time starting from State 3.

$$\langle T_1 \rangle = \frac{1}{k_{1\to0} + k_{1\to3} \cdot [B]} + \frac{k_{1\to3} \cdot [B]}{k_{1\to0} + k_{1\to3} \cdot [B]} \langle T_3 \rangle$$

On the other hand, starting State 3, we could get another equation, following the similar calculation:

$$\langle T_3 \rangle = \frac{1}{k_{3\to2} + k_{3\to1}} + \frac{k_{3\to1}}{k_{3\to2} + k_{3\to1}} \langle T_1 \rangle$$

Hence, the two equations yield:

$$\langle T_1 \rangle = \frac{k_{1\to3} \cdot [B] + (k_{3\to1} + k_{3\to2})}{(k_{3\to1} + k_{3\to2}) \cdot k_{1\to0} + k_{1\to3} k_{3\to2} \cdot [B]}$$

$$k_{off} = \frac{1}{\langle T_1 \rangle} = \frac{(k_{3\to1} + k_{3\to2})(k_{1\to0} - k_{3\to2})}{k_{1\to3} \cdot [B] + (k_{3\to1} + k_{3\to2})} + k_{3\to2}$$

$$\langle T_3 \rangle = \frac{k_{1\to3} \cdot [B] + (k_{3\to1} + k_{1\to0})}{(k_{3\to1} + k_{3\to2}) \cdot k_{1\to0} + k_{1\to3} k_{3\to2} \cdot [B]}$$

The only assumption here is that the reactions between different states are all first-order or pseudo-first-order.

2. Michaelis-Menten-like kinetics

Note that the k_{off} $(1/\langle T_1 \rangle)$ function derived above shows Michaelis-Menten-like kinetics as a function of $[B]$. As shown in fig. S13, k_{off} equals to $k_{1\to0}$ when $[B] = 0$ and to $k_{3\to2}$ when [B] is saturating. When k_{off} is in between these two values $(1/2(k_{1\to0} + k_{3\to2}))$, [B] equals to:

$$B_{\frac{1}{2}} = \frac{k_{3\to2} + k_{3\to1}}{k_{1\to3}}$$

If $k_{3\to1}$ is much greater than $k_{3\to2}$, $B_{1/2}$ can be regarded as K_d of B in the presence of A ($k_{3\to1}$/ $k_{1\to3}$). In our experiments, the second DNA binding proteins (T7 RNAp, EcoRV, and BamHI) exhibit larger off-rates from DNA than LacR does. Hence, it is reasonable to take $B_{1/2}$ value as a thermodynamic dissociation constant of the second protein from the LacR-bound DNA template.

3. Positive and negative cooperativity

The equilibrium constants for each step are (fig. S13):

$$K_{0\to1}=\frac{k_{0\to1}}{k_{1\to0}},\ K_{1\to3}=\frac{k_{1\to3}}{k_{3\to1}},\ K_{2\to3}=\frac{k_{2\to3}}{k_{3\to2}},\ K_{0\to2}=\frac{k_{0\to2}}{k_{2\to0}}$$

Their corresponding intrinsic free energy differences are:

$$\Delta G^\circ_{0\to1}=G^\circ_1-G^\circ_0=-k_BT\ln K_{0\to1},\ \Delta G^\circ_{1\to3}=G^\circ_3-G^\circ_1=-k_BT\ln K_{1\to3},$$
$$\Delta G^\circ_{2\to3}=G^\circ_3-G^\circ_2=-k_BT\ln K_{2\to3},\ \Delta G^\circ_{0\to2}=G^\circ_2-G^\circ_0=-k_BT\ln K_{0\to2}.$$

Then, we have

$$\Delta G^\circ_{total}=G^\circ_3-G^\circ_0=\Delta G^\circ_{0\to1}+\Delta G^\circ_{1\to3}=\Delta G^\circ_{0\to2}+\Delta G^\circ_{2\to3}$$
$$K_{0\to1}K_{1\to3}=K_{0\to2}K_{2\to3}.$$

Hence, $K_{2\to3}>K_{0\to1}\Leftrightarrow K_{1\to3}>K_{0\to2}$, which implies the binding of one protein would enhance the binding of the other protein simultaneously (positive cooperativity); and $K_{2\to3}<K_{0\to1}\Leftrightarrow K_{1\to3}<K_{0\to2}$, which implies the binding of one protein would weaken the binding of the other protein (negative cooperativity).

4. Coupling energy $\Delta\Delta G^\circ_{AB}(L)$ for protein A and protein B

In the main text, we introduced the energetic coupling term, $\Delta\Delta G^\circ_{AB}(L)$, for the protein A and B. Here we explain how to calculate $\Delta\Delta G^\circ_{AB}(L)$ based on the experimentally measured $k_{3\to2}/k_{1\to0}$ values. By definition:

$$\Delta\Delta G^\circ_{AB}(L)=\Delta G^\circ_{0\to3}(L)-\Delta G^\circ_A-\Delta G^\circ_B$$
$$\Delta G^\circ_A=\Delta G^\circ_{0\to1}=G^\circ_1-G^\circ_0=-k_BT\ln K_{0\to1}$$
$$\Delta G^\circ_B=\Delta G^\circ_{0\to2}=G^\circ_2-G^\circ_0=-k_BT\ln K_{0\to2}$$

Hence,

$$\Delta\Delta G^\circ_{AB}(L)=\Delta G^\circ_{0\to3}(L)-\Delta G^\circ_{0\to1}-\Delta G^\circ_{0\to2}=\Delta G^\circ_{2\to3}(L)-\Delta G^\circ_{0\to1}=\Delta G^\circ_{1\to3}(L)-\Delta G^\circ_{0\to2}$$
$$=k_BT\ln\left(\frac{k_{3\to2}}{k_{1\to0}}\cdot\frac{k_{0\to1}}{k_{2\to3}}\right)$$

Or,

$$\frac{k_{3\to2}}{k_{1\to0}}=\frac{k_{2\to3}}{k_{0\to1}}\cdot\exp(\Delta\Delta G^\circ_{AB}(L)/k_BT)\propto\exp(\Delta\Delta G^\circ_{AB}(L)/k_BT)$$

Assuming that the bimolecular rate constants, $k_{2\to3}$ and $k_{0\to1}$, are both independent of L, the oscillation of $\Delta\Delta G^\circ_{AB}(L)$ could be shown by the oscillation of $\frac{k_{3\to2}}{k_{1\to0}}$. If we further assume $k_{2\to3}$ and $k_{0\to1}$ are very close to each other, $\Delta\Delta G^\circ_{AB}(L)$ would be exactly equal to $k_BT\ln\left(\frac{k_{3\to2}}{k_{1\to0}}\right)$.

For LacR and T7 RNAp: LacR is protein A and T7 RNAp is protein B. Using $k_{3\to2}/k_{1\to0}$ shown in Fig. 3D, top, we calculated $\Delta\Delta G^{\circ}_{AB}(L)$ (fig. S14A). The energetic coupling oscillates with amplitude of ~2 k_BT, similar to that of $\Delta G^{\circ}_{0\to3}(L)$. This comparison confirms that the previous assumption that $\Delta G^{\circ}_{0\to1}$ and $\Delta G^{\circ}_{0\to2}$ as well as $k_{2\to3}$ and $k_{0\to1}$ are independent of L.

For GRDBD and BamHI: GRDBD is protein A and BamHI is protein B. Using $k_{3\to2}/k_{1\to0}$ shown in Fig. 1B, we calculated $\Delta\Delta G^{\circ}_{AB}(L)$ (fig. S14B). The energetic coupling oscillates with amplitude of ~1.5 k_BT, similar to that of LacR and T7 RNAp.

*** Comparing *in vivo* results (Fig. 4B) with *in vitro* results (fig. S15)**

The *in vivo* data were directly compared with *in vitro* results, obtained using the same sequence spanning *lacO*1 and T7 promoter as in the chromosomal live cell experiment. Interestingly, a phase shift of ~3 bp was observed between the oscillations measured *in vitro* and *in vivo* (fig. S15). The phase shift may be explained by the following three reasons: First, in the single-molecule assay, T7 RNAp was stalled as the open complex without NTPs. The transition from open complex to elongation complex accompanies large conformational changes in both the protein and DNA (*24, 71*), which can change where the maximum distortion occurs on DNA and induce the phase shift. Second, the chromosomal DNA in *E. coli* is negatively supercoiled (*72*), which can lead to the phase shift. Third, there are many nucleoid-associated proteins (NAPs) that extensively bind to DNA in *E.coli* (*73*). The binding of NAPs can change the local structure of DNA and cause the phase shift.

*** Result of *in vivo* experiments with $lacO_{id}$**

We investigated allosteric interaction between LacR and T7 RNAp on plasmids in *E. coli*. The construct is built on the pET-21b vector, which carries a T7-promoter sequence, symmetric *lac* operator ($lacO_{id}$) and a LacR coding sequence (*lacIq*). We placed a gene encoding rapidly maturing (~7 min) yellow fluorescent protein, Venus, under T7 promoter (fig. S16A). In this configuration, the presence of LacR on $lacO_{id}$ may affect only the transcription initiation and not transcription elongation of T7 RNAp. We also performed similar experiments using $lacO_{id}$ as operator and $lacZO_2^-$ as reporter on the chromosome, using the same spacer sequence as on the plasmid. By measuring Venus fluorescence or LacZ activity, we monitored any change in promoter-T7 RNAp interactions due to the presence of LacR upstream of the promoter.

We observed variation in the expression level of reporter gene as a function of the separation between $lacO_{id}$ and T7 promoter both on the plasmid and chromosome (fig. S16B). The oscillation in the expression level was observed between 3 and 26 bp of separation, with a periodicity close to 10 bp and the amplitude dampening with increasing DNA spacer length. We also noticed that the oscillation curves of $lacO_{id}$ and *lacO*1 in chromosomal experiments are very similar, both in periodicity and phase.

*** Nucleosome stability assay**

As a control experiment, we used a single-molecule immunofluorescence technique (*74*) to confirm that the nucleosome complex is stable enough to remain bound on DNA during our single-molecule experiments. Nucleosomal DNAs were immobilized on the coverslip through

biotin-streptavidin interaction, as described before. Then, 10 nM anti-H3 (Sigma-Aldrich), labeled with Cy3 NHS-ester (GE Healthcare), was infused into the flow-cell and incubated for 6 minutes. After washing, Cy3 fluorescence images were taken to visualize the nucleosome complex as a diffraction-limited spot. From the fluorescence intensity, we confirmed that most of the spots correspond to only one nucleosome. The number of spots were counted per image (~80x80 μm^2), and standard deviations were calculated for images taken from 10 different regions.

We observed 191.8 $\pm$ 1.3 spots per image, which is similar to the number of GRDBD complex observed in the original experiments described in Fig. 1. The background binding resulted in only 9.2 $\pm$ 0.8 spots per image.

When we kept the 532-nm laser on at 6 W/cm^2 and infused buffer continuously at ~1 uL/sec (same as the original experiments for the nucleosome-GRDBD), the number of spots per image was reduced to 155.8 $\pm$ 1.6 after 15 min. If we only kept flow on without laser illumination, the mean spot count per image was 184.6 $\pm$ 6.1 even after 2 hours. These results suggest that the histone complex is much more stable than GRDBD on DNA, as the latter dissociated completely within 15 min. We believe that reduction of spots in the presence of laser illumination is partially accounted for by photobleaching of Cy3, which is less stable than Cy3b.

We also did the same experiments on naked DNA, *i.e.*, without nucleosome complex. We found that the anti-H3 cannot bind to naked DNA (26.1 $\pm$ 0.8 spots per image). Non-specific binding relative to background produced only 14.1 $\pm$ 0.9 spots per image. Hence, the fluorescence spot is specific to the nucleosomes (fig. S18).

SUPPORTING FIGURES

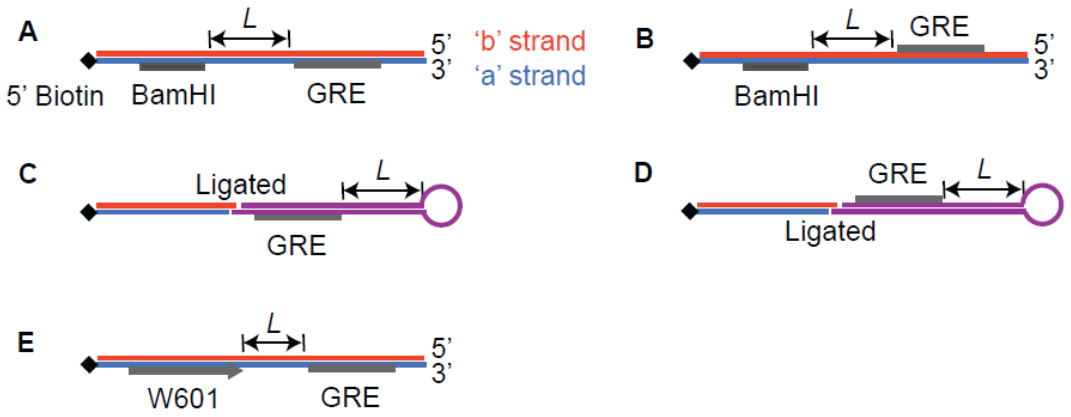

Fig. S1. DNA templates used in measuring GRDBD dissociation rate constant. **(A)** DNA template for the forward GRDBD-BamHI experiments shown in Fig. 1B. **(B)** DNA template for reverse GRDBD-BamHI experiments shown in Fig. 1B. **(C)** DNA template for forward GRDBD-hairpin experiments shown in Fig. 1C. **(D)** DNA template for reverse GRDBD-hairpin experiments shown in Fig. 1D. **(E)** DNA template for GRDBD and nucleosome. This DNA template was made by PCR. See Table S1,2 and 7 for exact sequences of these templates.

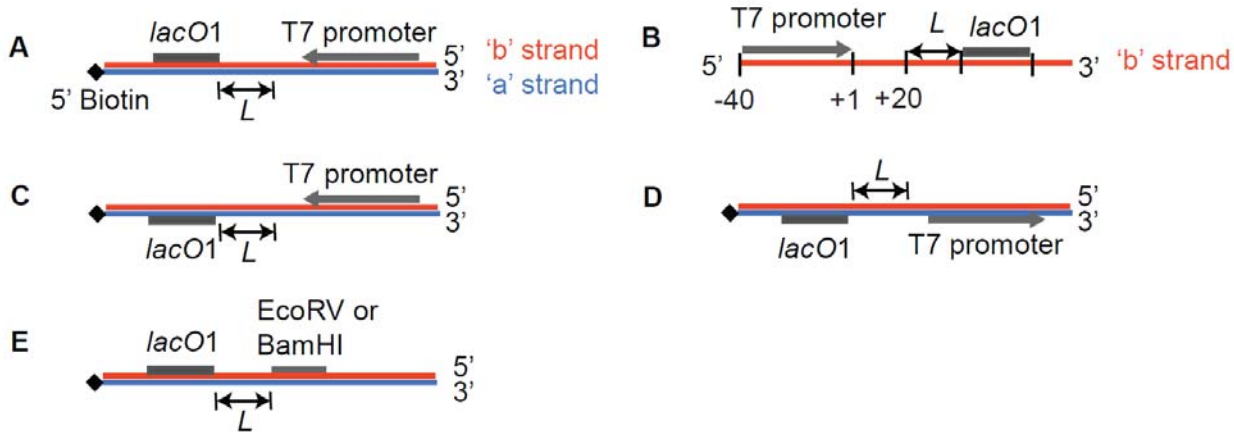

Fig. S2. DNA templates used in measuring LacR dissociation rate constant. **(A)** DNA template for the (forward) *lacO*1 downstream of T7 RNAp (data shown in Fig. 3). Arrow indicates T7 promoter in the direction of transcription. **(B)** Detailed view of the 'b' strand shown in **A**. In the templates shown in **A-C**, we used the same base sequences around T7 promoter (-40 to +20 positions; +1 is the transcription start site for T7 RNAp). We varied *L* length by adding 0-30 bp random sequence (see Table S3). **(C)** DNA templates for the 'reverse' *lacO*1 experiment shown in fig. S5. The DNA sequences were almost the same as **A** except *lacO*1 being on the opposite strand of T7 promoter (see Table S3). **(D)** DNA template for the 'upstream' *lacO*1 experiment shown in fig. S15. To be consistent with DNA sequences used in live cells, we used only 28-bp sequence (-27 to +1 positions) for T7 promoter (see Table S5). **(E)** DNA templates for LacR and endonuclease (e.g. EcoRV and BamHI). Data are shown in fig. S6.

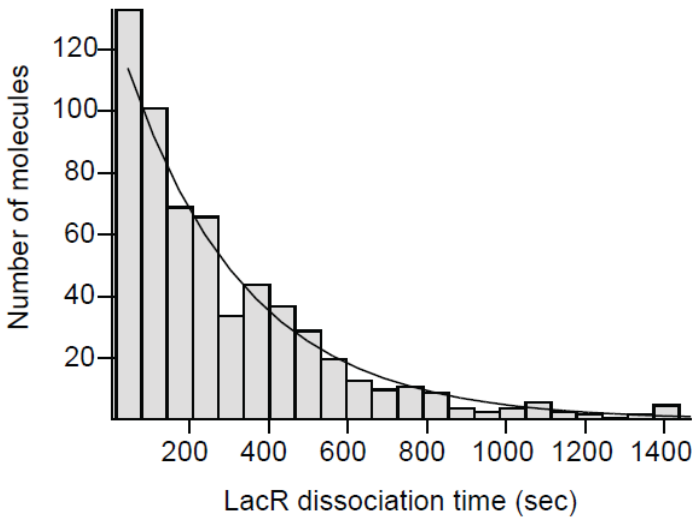

Fig. S3. Histogram of dissociation times of LacR from DNA in the absence of T7 RNAp. We analyzed fluorescence images of single LacR molecules to determine time for dissociation from their initial locations. As for the single-step dissociation kinetics for LacR, the histogram was fit with a single exponential decay curve, $\exp(-t/\tau)$, where τ indicates the mean first passage time ($1/k_{off}$).

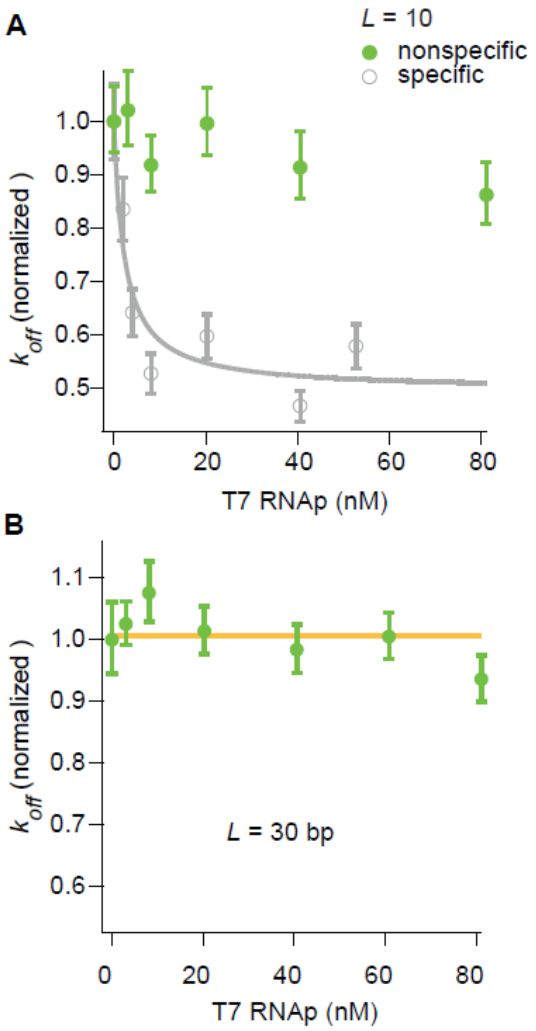

Fig. S4. A few instances where k_{off} of LacR was unaffected by T7 RNAp. **(A)** When we performed the titration experiment using DNA templates with only *lacO*1 but no T7 promoter, we observed no change in the k_{off} of LacR even in the saturating concentration of T7 RNAp, indicating that the change observed with T7 promoter (e.g., L = 10 shown in gray) was due to specific binding of T7 RNAp. **(B)** When the distance between *lacO*1 and T7 promoter was >25 bp, we did not observe any change in the k_{off} of LacR.

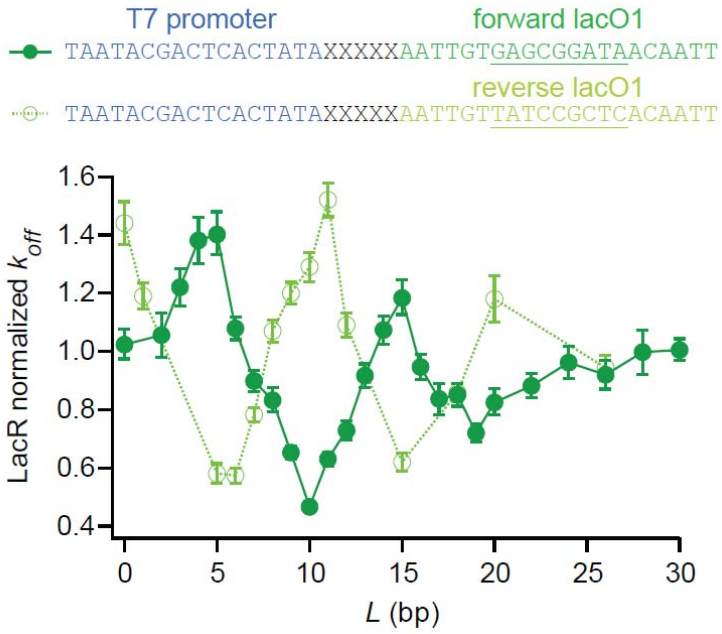

Fig. S5. Allosteric coupling between LacR and T7 RNAP when *lacO*1 is reversed on the DNA. The central non-palindromic region of *lacO*1 is underlined in the sequence. For the reverse template, we simply placed *lacO*1 on the strand opposite to T7 promoter. The oscillation in the maximum k_{off} , $k_{3\to2}$, was shifted by ~5 bp (light green open circles) in comparison to the previous measurement (green filled circles; re-plot of Fig. 3D top).

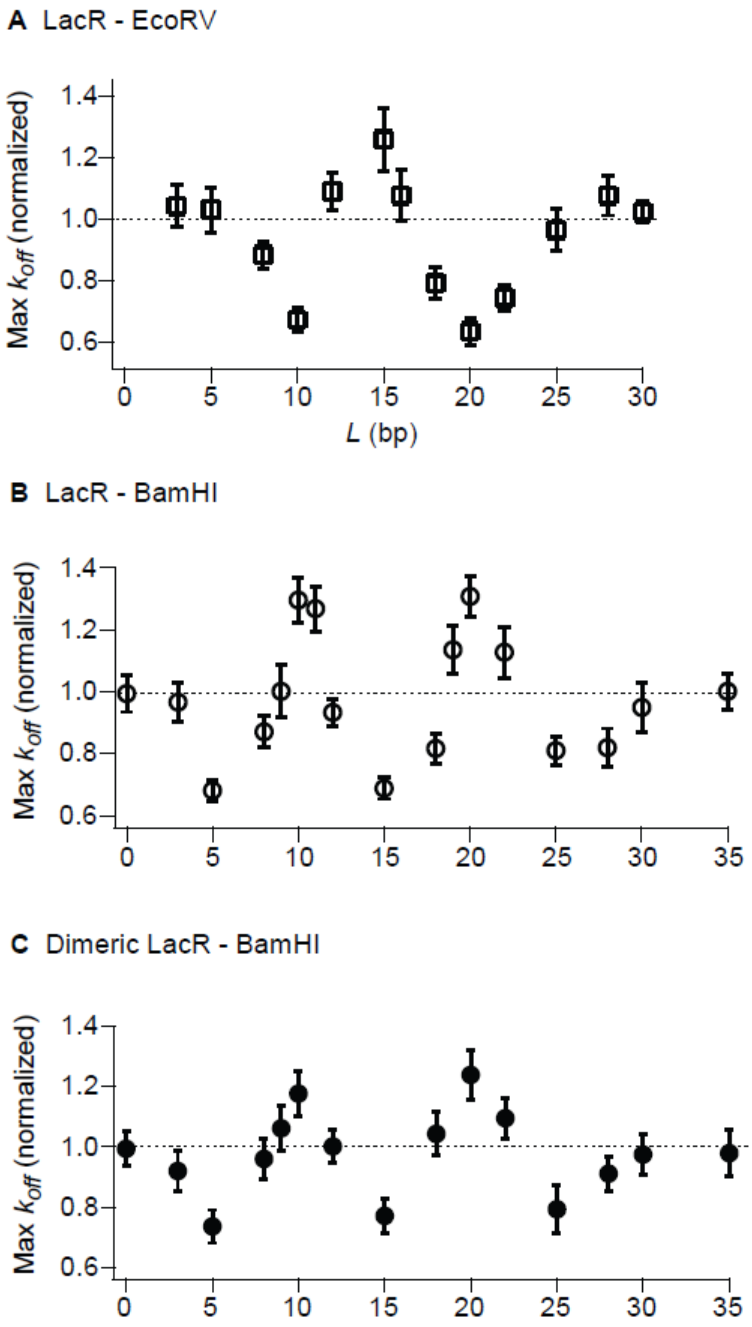

Fig. S6. Allosteric interaction between LacR and endonuclease as a function of their separation on DNA. **(A)** Allosteric interaction between LacR and EcoRV specifically bound to DNA. Similar to LacR-T7 RNAp, the stability of LacR was modulated by EcoRV with the periodicity of 10 bp. **(B)** Allosteric interaction between LacR and BamHI specifically bound to DNA, also occurring with the periodicity of 10 bp. **(C)** Allosteric interaction between dimeric LacR and BamHI specifically bound to DNA. No significant difference was observed between tetrameric and dimeric LacR.

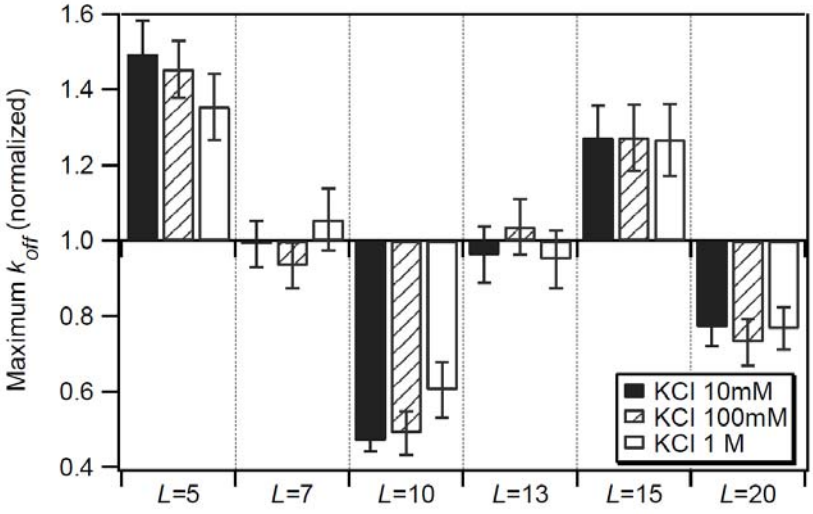

Fig. S7. Salt dependence experiment for LacR-T7 RNAp. Concentration of potassium chloride in the imaging buffer was varied from 10 mM (black bars) to 100 mM (dashed bars) and 1000 mM (white bars). The oscillatory behavior remains unaffected by the salt concentrations.

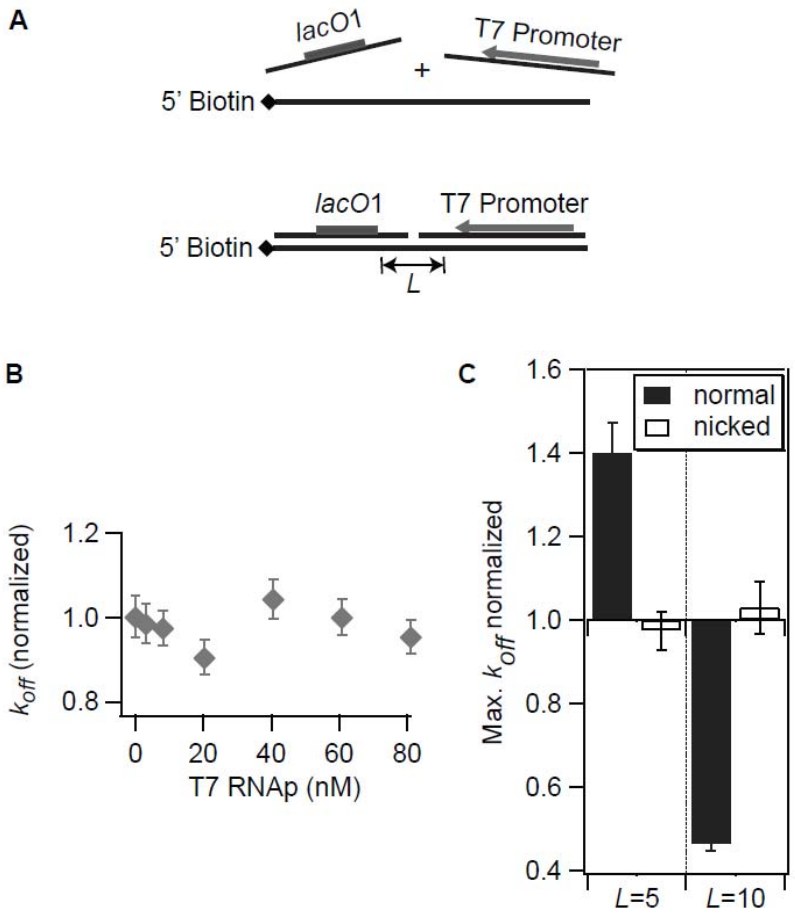

Fig. S8. Effect of a nick in the linker DNA. **(A)** Nicked DNA template. 3 pieces of the DNA oligonucleotides were annealed to produce a nicked template. A nick was located in the middle of the spacer region on the nontemplate strand. **(B-C)** No cooperativity on the nicked DNA template. In the presence of a nick between two protein binding sites, the cooperativity between LacR and T7 RNAp was no longer observed. **(B)** Titration curve on L = 10 with a nick. **(C)** Comparison between normal and nicked DNA for L = 5 and 10 (black bars: normal dsDNA template, white bars: nicked dsDNA template).

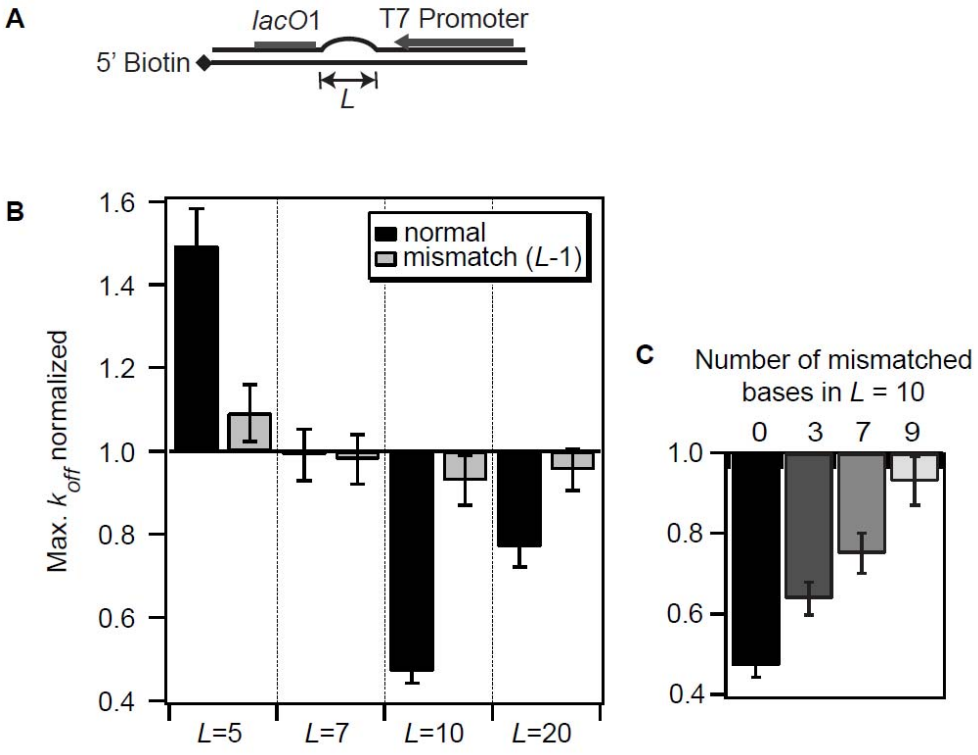

Fig. S9. Effect of base pair mismatches in the linker DNA. **(A)** Illustration of a mismatched template. **(B)** The effect of mismatched bases in the linker region on the cooperativity between LacR and T7 RNAp. Cooperativity was removed when most bases in the linker DNA, *i.e.*, (*L*-1) bases, were mismatched (black bars = normal DNA, gray bars = mismatched DNA). **(C)** The effect of the number of mismatches in the linker region. For $L = 10$ template, the number of mismatched bases was varied from 3, 7 to 9. Gradual decrease in cooperativity was observed with increasing size of the mismatched region in $L = 10$.

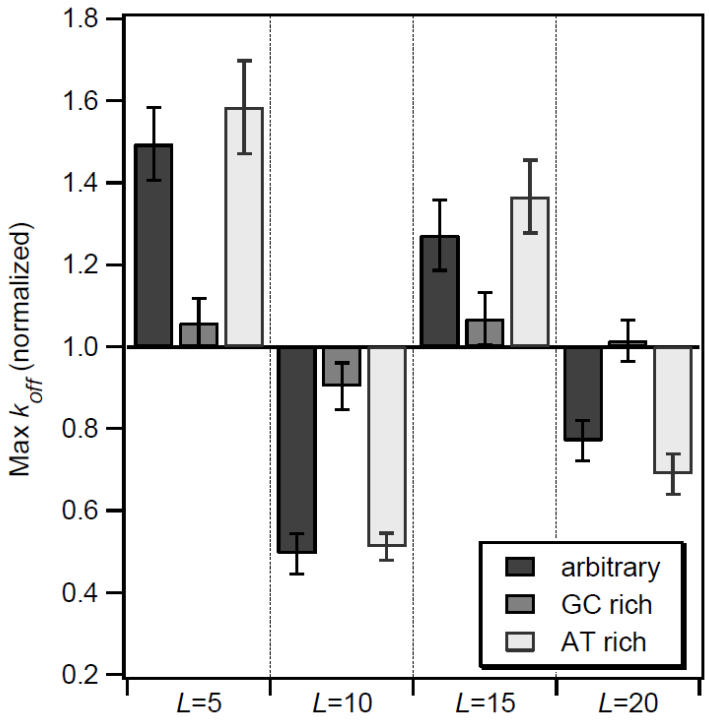

Fig. S10. Effect of base composition in the linker DNA. We measured the cooperativity between LacR and T7 RNAp using new DNA templates, in which the original random sequence in the linker region (*L*) was replaced by GC- or AT-rich sequences. Surprisingly, replacing the sequence to complete G/C bases (gray bar) removes the cooperativity observed previously (black bar; same data values as in Fig. 3D top). On the contrary, AT rich template (light gray bar) reproduced a similar cooperativity pattern observed in a random sequence.

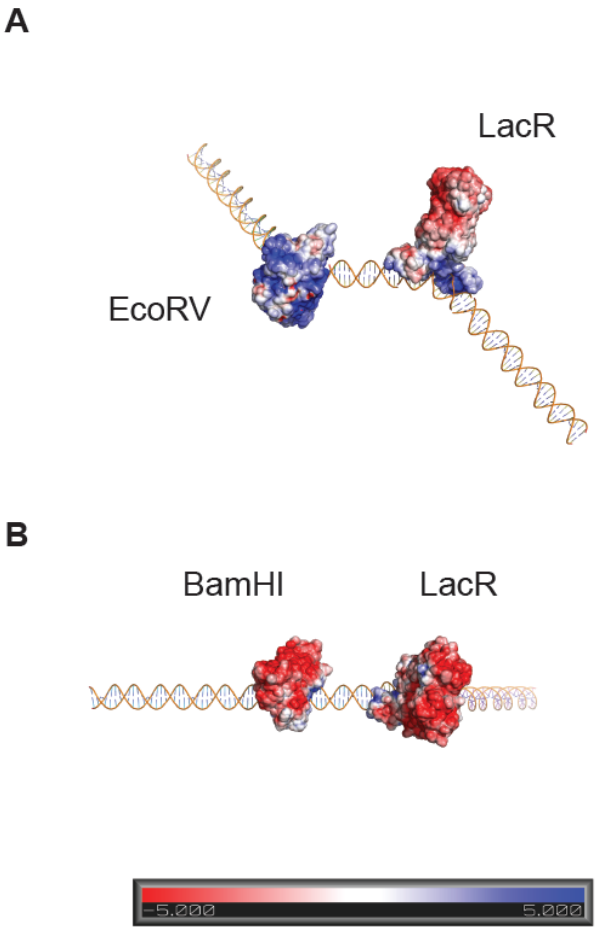

Fig. S11. Structural models for two DNA-binding proteins on a DNA template. **(A)** LacR-EcoRV complexed with DNA of $L = 10$. **(B)**Top view of LacR–BamHI complexed with DNA of $L = 10$. Blue and red surfaces represent areas of positive and negative potential, respectively (1 unit in the scale bar corresponds to 26.7 mV).

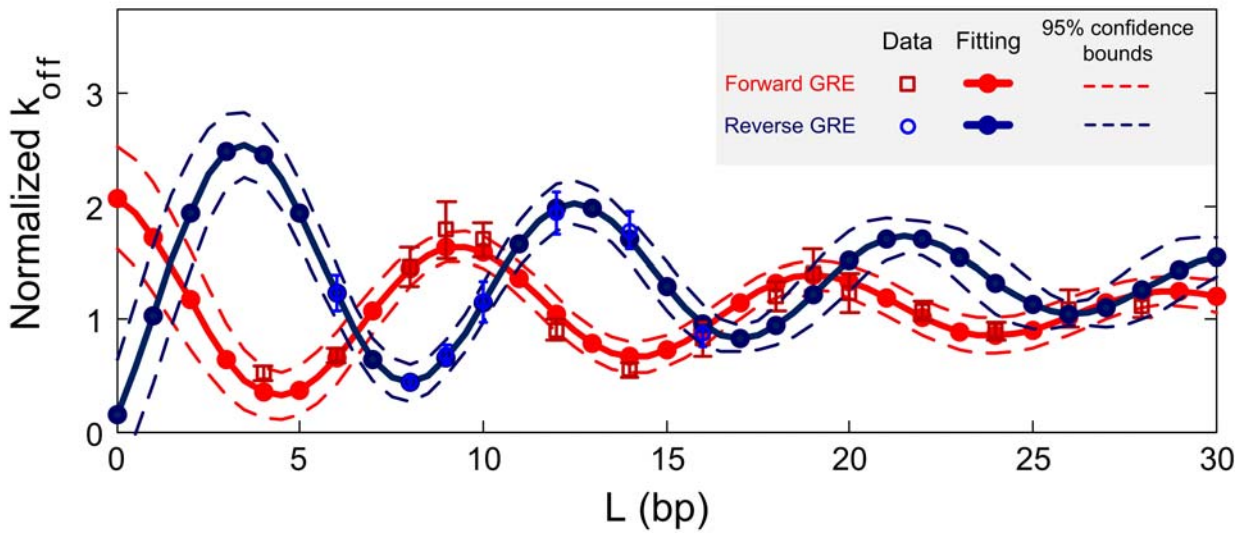

Fig. S12. Data fit to characterize oscillation of k_{off} of GRDBD as a function of its distance to 3-bp hairpin loop (Fig. 1D).

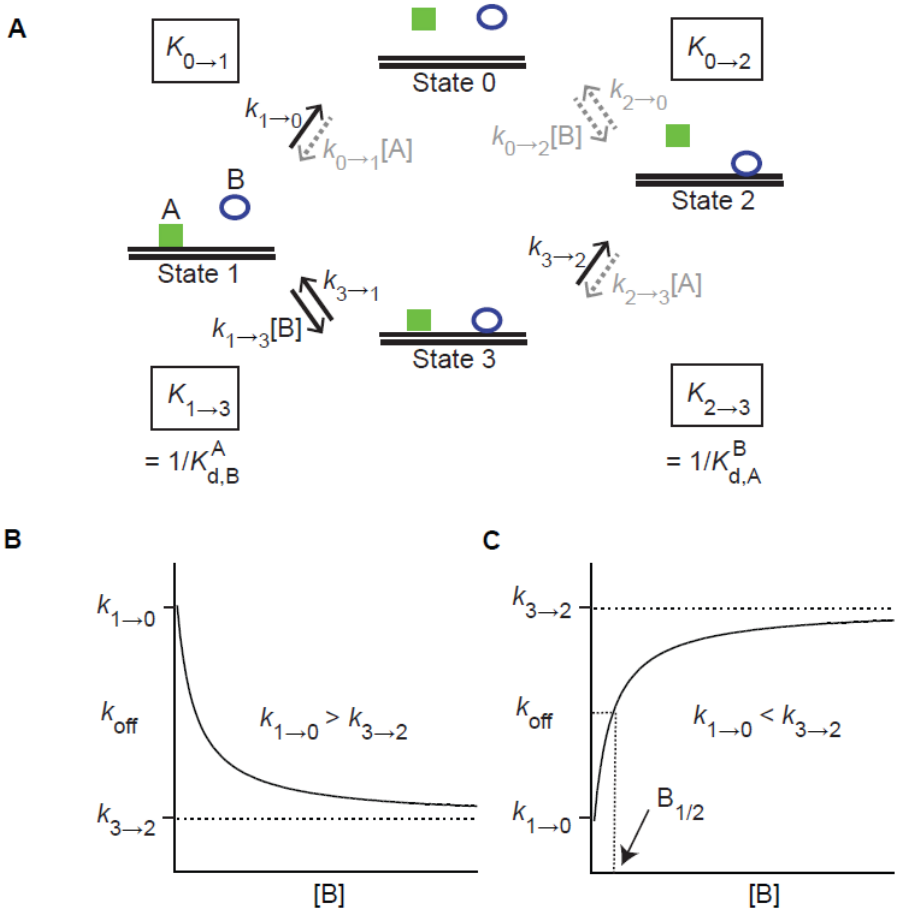

Fig. S13. Thermodyamic model for dissociation of LacR (indicated as A) in the presence of a second protein (indicated as B). **(A)** In our *in vitro* experiment, we start with State 1 and observe dissociation of LacR to State 0 or State 2 (via State 3). Reassociation of LacR is not counted in the observation, and any reaction pathways for the reassociation are dashed in light gray. [A] = concentration of LacR, [B] = concentration of T7 RNAp, $K_{i\to j}$ equilibrium constant for reaction from State i to j, and $k_{i\to j}$ rate constant from State i to j. **(B-C)** Hypothetical titration curves. Depending on the relative magnitudes of $k_{1\to0}$ and $k_{3\to2}$, titration curves may result in three different forms. In the case of $k_{1\to0} \approx k_{3\to2}$, the titration curve would be flat, *i.e.*, k_{off} will not change with [*B*] (fig. S4B). If $k_{1\to0} > k_{3\to2}$, the titration curve would be of decreasing hyperbolic shape. The opposite would be the case when $k_{1\to0} > k_{3\to2}$. [*B*] that gives k_{off} of LacR half way to the saturation value is $B_{1/2}$.

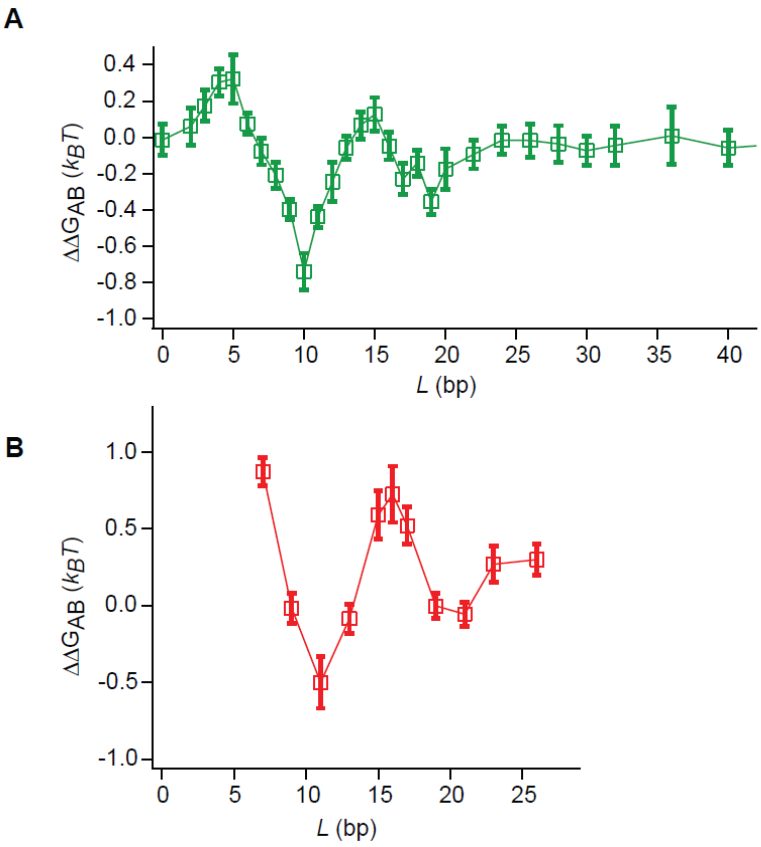

Fig. S14. Allostery coupling energetics between LacR and T7 RNAp **(A)** and between GRDBD and BamHI **(B)**.

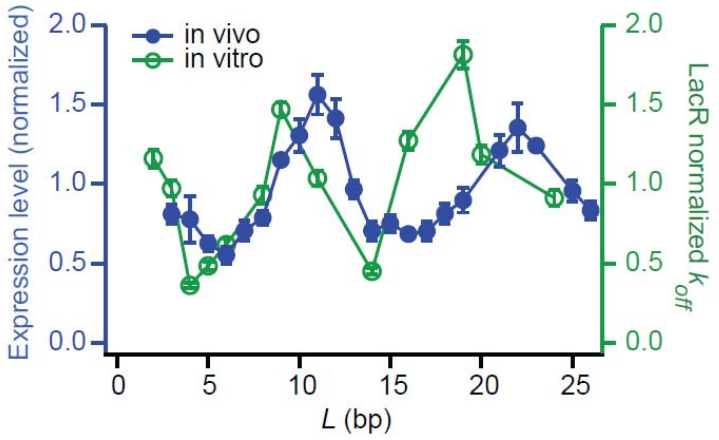

Fig. S15. Allostery through DNA *in vivo*. The expression level of *lacZ* (solid blue) oscillates as a function of *L* with a periodicity of 10 bp, similar to the corresponding *in vitro* data (open green). The data were normalized with the average expression levels of all *L*'s. This *in vivo* result suggests the relevance of allostery through DNA to gene regulation.

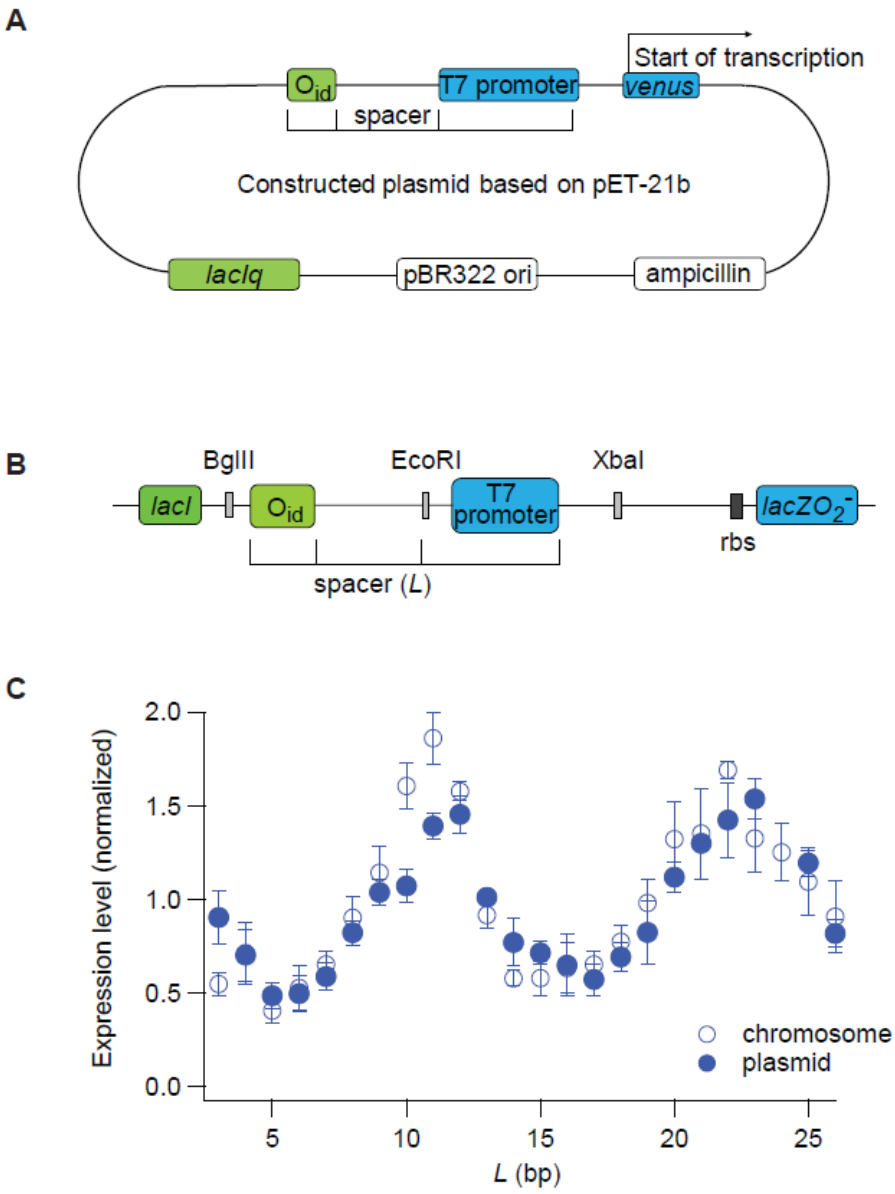

Fig. S16. DNA allostery experiment using symmetric *lacO* (O_{id}) *in vivo*. **(A)** Schematic for the plasmid constructed to investigate LacR-T7 RNAp allosteric interactions in live *E. coli*. **(B)** Schematic for re-engineered *lac* operon locus in *E. coli* to study DNA allostery on the chromosome. **(C)** Expression levels measured by fluorescence intensities of Venus or β-galactosidase assay for allostery on the plasmid or chromosome, respectively. For plasmid data, Venus fluorescence was corrected by OD600 of each cell culture and normalized by the average readings of all samples. Both on the plasmid and chromosome, the expression level of a gene under T7 promoter oscillates as a function of the separation between O_{id} and T7 promoter, with a periodicity of ~10 bp and comparable phase.

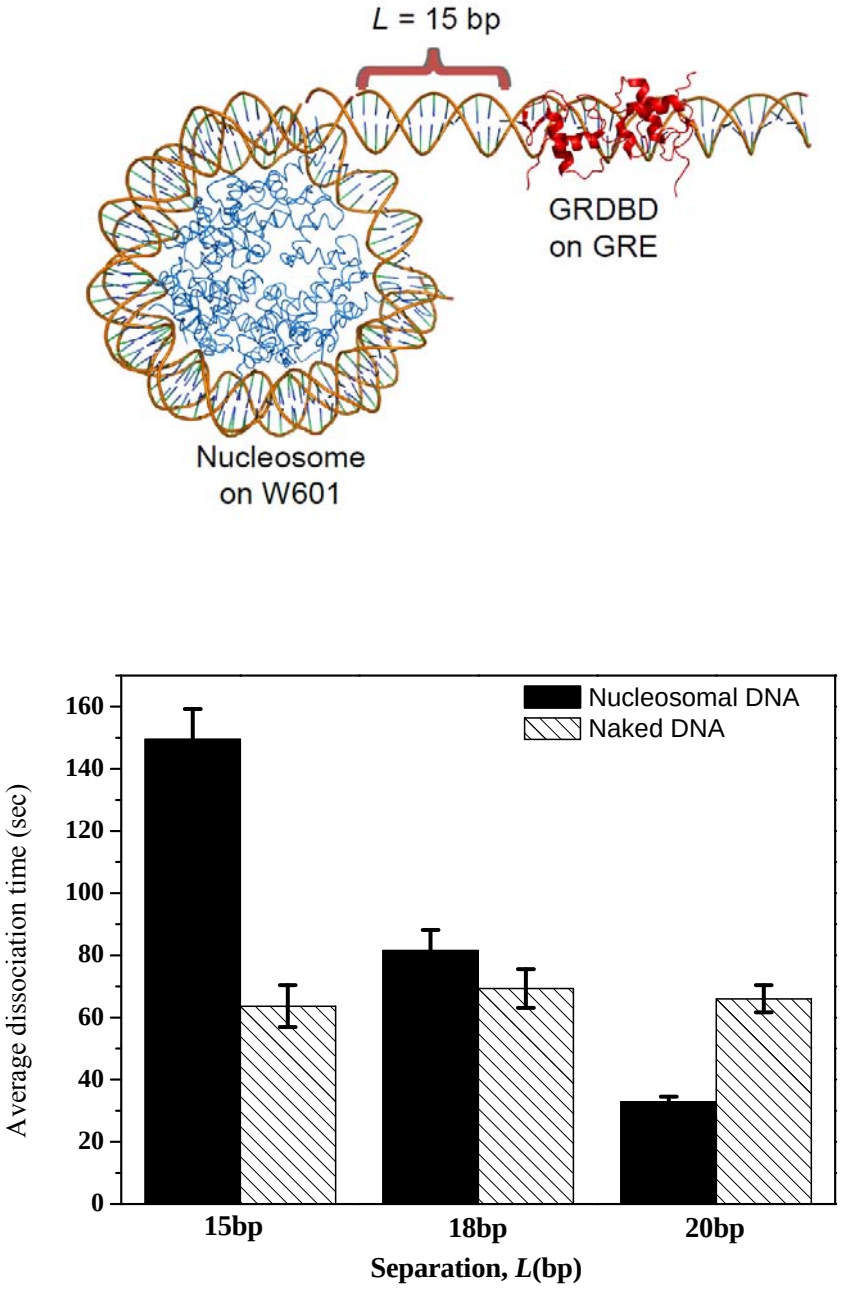

Fig. S17. GRDBD-nucleosome experiments. (top) Structural model of the ternary complex formed by GRDBD and nucleosome with linker DNA of L = 15 bp. Nucleosome was positioned on W601, the Widom 601 nucleosome positioning sequence (PDB ID for nucleosome = 3LZ0 (*75*)) and for GRDBD = 1R4R (*13*)). (bottom) Average dissociation time ($1/k_{off}$) of Cy3b-labeled GRDBD on DNA in the presence (black solid bar) and absence (striped bar) of the nucleosome complex. In the absence of the nucleosome, the dissociation time of GRDBD does not change much at different seperation between GRE and W601.

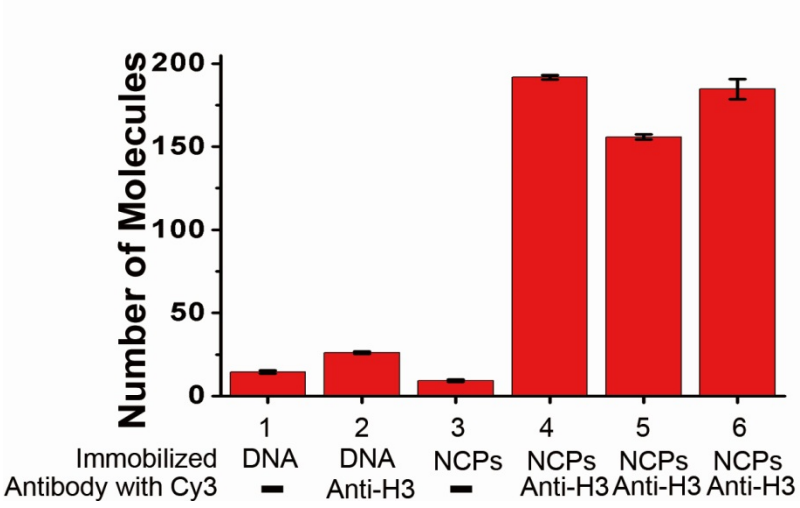

Fig. S18. Histone stability assay**.** Here we present the number of anti-H3 bound to histone (*i.e.*, the number of Cy3 spots) under different experimental conditions. In condition 1-2, naked DNAs were immobilized on the surface to check false positive (1) and background binding of Anti-H3 (2). In condition 3-6, nucleosome complexes were immobilized on the surface (NCP), and Anti-H3 was added (4). To check the effect of laser illumination and continuous flow, we measured the number of spots after 15 minutes of continuous laser illumination (5) and after washing with imaging buffer for 2 hours (6), respectively. Error bars denote standard deviation (n = 10). This result indicates that laser- or flow-induced dissociation of histone is negligible in the time scale of GRDBD binding.

SUPPORTING TABLES

Table S1. **DNA template sequences used in the GRDBD-BamHI experiments**
'a' strand is composed of (1) 5'end biotin, (2) 28-bp random sequence, (3) 6-bp BamHI recognition sequence, (4) linker (*L*), (5) forward GRE or reverse GRE, and (6) 14-bp random sequence (from 5' to 3').
Biotinylated 28-bp random sequence = 5'Biotin-TTTTTTCATGTATAGCTGTATGTGTGA
BamHI binding site = GGATCC
Linker (*L*) = in the table ((A) for forward GRE and (B) for reverse GRE)
Forward GRE= AGAACATCATGTTCT
Reverse GRE= AGAACATGATGTTCT
14-bp random sequence = GAGGTCAGTAGTTC
The complementary strand ('b' strand) is not biotinylated.

(A) BamHI and forward GRE template

Template (*L*)	Sequence for *L*
7	ATAGACC
9	ACTGAGACC
11	ACTGAACGACC
13	ACTGACTACGACC
15	ACTGATACTACGACC
16	ACTGACTACTACGACC
17	ACTGATCTACTACGACC
19	ACTGATGTCTACTACGACC
21	ACTGACATGTCTACTACGACC
23	ACTGACTAATGTCTACTACGACC
26	ACTGACTACTCATGTCTACTACGACC

(B) BamHI and reverse GRE template

Template (*L*)	Sequence for *L*
9	ATAGAGACC
11	ACTGAACGACC
12	ACTGATACGACC
13	ACTGACTACGACC
14	ACTGAACTACGACC
15	ACTGATACTACGACC
17	ACTGATCTACTACGACC
19	ACTGATGTCTACTACGACC

Table S2. **DNA template sequences used in the GRDBD-hairpin experiments**

DNA templates for GRDBD-hairpin were constructed using 3 single-stranded DNA oligonucleotides. As shown in fig. S1, first two ssDNA oligonucleotides are annealed together to form biotinylated dsDNA handle.

'a' strand:

Biotin-TTTTTTTCATGTATAGCTGTATGTGTGACCTGATTACTCGCATCAATCGG

'b' strand:

TGCCGATTGATGCGAGTAATCAGGTCACACATACAGCTATACATG

The third piece of ssDNA is designed to form a hairpin. It has sticky ends at the end of the stem for ligation with the dsDNA handle. Some exemplary sequences for this third piece of ssDNA are shown below. These sequences contain GRE and the linker sequence in the stem region and GGG in the loop region. For the longer, 15-bp loop, we had GGGATGTGTGAAGGG instead of GGG.

Template (*L*)	Sequence
4	CATCCAGAACATCATGTTCTGAGAGGGTCTCAGAACATGATG TTCTGGA
6	CATCCAGAACATCATGTTCTGAGATAGGGTATCTCAGAACAT GATGTTCTGGA
8	CATCCAGAACATCATGTTCTGAGATAGAGGGTCTATCTCAGA ACATGATGTTCTGGA
10	CATCCAGAACATCATGTTCTGAGACTGAGAGGGTCTCAGTCT CAGAACATGATGTTCTGGA
12	CATCCAGAACATCATGTTCTGAGACTGAACGAGGGTCGTTCA GTCTCAGAACATGATGTTCTGGA
14	CATCCAGAACATCATGTTCTGAGACTGACTACGAGGGTCGTA GTCAGTCTCAGAACATGATGTTCTGGA

Table S3. DNA template sequences used in the LacR-T7 RNAp experiments (fig. S2).
(A) Our 'a' strand is biotinylated at the 5' end and is complimentary to the 'b' strand, part of which is provided in the table. 'b' strand is composed of (1) T7 promoter from -40, (2) +1~+20 common sequence, (3) linker (*L*), (4) *lacO*1, and (5) ~30 random sequence (from 5' to 3').
T7 promoter = TGGCGACGGCAGCGAGGCTAAATTAATACGACTCACTATA
+1 ~ + 20 common sequence = GGGAGATGTTGTTGATCCCC
Linker (*L*) = in the table
*lacO*1 = AATTGTGAGCGGATAACAATT
Random sequence = TCACACAAACAGCTATGACATCTGTGTGG (an example)

(B) In our reverse *lacO*1 experiment (fig. S5), we constructed 'b' strand the same way as above, except using reverse complementary *lacO*1 sequence. The linker DNA sequence was slightly different at certain templates to avoid secondary structures. In $L = 26^{*}$, we replaced last C of the common sequence with A.

(A) Forward *lacO*1 templates

Template (*L*)	Sequence for *L*
0	
2	TG
3	AAT
4	AATG
5	AATGG
6	AATGTG
7	AATTGTG
8	AATTGTGT
9	AATGGTGTG
10	AATTGTGTGG
11	AATTCTCTGGC
12	AATTGTGTGGCG
13	AATTGTGTGGCTG
14	ATATGTGTGTCGTG
15	AATTGTGTGGCGTAT
16	AATTGTGCTGTATGTG
17	AATTGTGCTCGTATGTG
18	AATTGCGCTCGTATGTGG
19	AATTGTGCTCGTCTGAGTG
20	AATTGCGCTCGTATGTGTGG
22	ATTGCACTCCAATTGCTCGTGG
24	ATTGCACTCCAATTGCCCTCGTGG
26	ATTGCACTCCAATTGCTCTCGTGTGG
28	ATTGCACTCCAATTGCTCTCGTATGTGG
30	ATTGCACTCCAATTGCGCTCGTATGTGTGG
32	ATTGCACTCCAATTCGCTTCGTTCATGTGTGG
36	ATTGCACTCCAATTCGCTTCGTTCTCGTATGTGTGG
40	ATTGCACTCCAATTCGCTTCGTTCGCGCTCGTATGTGTGG

(B) Reverse *lacO*1 templates

Template (*L*)	Sequence for *L*
0	
1	T
5	AATGG
6	AATGTG
7	AGTTGTG
8	AATTGTGT
9	AATGGTGTG
10	AATTGTGTGG
11	AATTCTCTGGC
12	AATTGTGTGGCG
15	AATTGTGTGGCGTAT
18	AATTGCGCTCGTATGTGG
20	AATTGCGCTCGTATGTGTGG
26*	(A)ATTGCACTCCAATTGCTCTCGTGTGG

Table S4. Additional DNA templates used in the LacR-T7 RNAp experiments.
As noted in Table S3, all the sequences in this table are parts of 'b' strands.
(A) Bubble template: The 'a' strand used in the main experiment (Table S3) was annealed with a newly designed 'b' strand template. The new 'b' strand is the same as the original 'b' strand except the mismatched sequences (underlined). To create 19 mismatches in $L = 20$, we used 'a' strand of the AT-rich template and 'b' strand of the GC-rich template.

(B) GC or AT rich template: Sequence for *L* is provided. The other regions of the 'b' strand were the same as those in the original 'b' strand. See the legend for Table S3.

(A) Bubble template

Template	Sequence for *L*
4 mismatch in $L = 5$	GGCCG
6 mismatch in $L = 7$	GCCCCCG
9 mismatch in $L = 10$	GCCCCCCCCG
7 mismatch in $L = 10$	CCCCCCCGG
3 mismatch in $L = 10$	CCCGTGG
19 mismatch in $L = 20$	See legend

(B) GC or AT rich template

Template	Sequence for *L*
GC rich $L = 5$	GGCCG
GC rich $L = 10$	GCCCCCCCCG
GC rich $L = 15$	GCCCCCCCCCGCCCG
GC rich $L = 20$	GGGGGGGCGGGGGGGGGGGG
AT rich $L = 5$	AATAG
AT rich $L = 10$	AATAAAATAG
AT rich $L = 15$	AATAAAATAAAATAG
AT rich $L = 20$	AATTAAAATAAAATATAAAG

Table S5. DNA templates used in the LacR (upstream)-T7 RNAp experiments (*in vitro* and *in vivo* shown in Fig. 4B and fig. S15).

For convenience, we show the sequences used in the *in vitro* experiments. However, the same sequence between *lacO*1 and T7 promoter was employed in the *E.coli* strains for the chromosome experiments, shown in Fig. 4B.

All sequences are for the biotinylated 'a' strands, which are composed of (1) 5'biotin-common sequence, (2) *lacO*1, (3) linker (*L*), (4) -25 ~ -22 common sequence, (5) T7 promoter from -21, (6) +1 ~ +23 common sequence (from 5' to 3'; see fig. S2).

Common sequence = CAGATGTAATAGCTGTATGTGTGAGG
*lacO*1 = AATTGTGAGCGGATAACAATT
Linker (*L*) = in the table
-27~-18 common sequence of T7 promoter = GAATTCAAAT
T7 promoter = TAATACGACTCACTATA
+1 ~ +23 common sequence = GGGAGATGTTGTTGATCTCGACA

Template (*L*)	Sequence for *L*
0	
2	CG
3	CGA
4	CGTA
5	CGTTA
6	CGGTTA
8	CGCAGTTA
9	CGTCAGTTA
11	CGCGTCAGTTA
14	CGATTCGTCAGTTA
16	CGATTCTCGTCAGTTA
19	CGATTCTAATCGTCAGTTA
20	CGATTCTAAATCGTCAGTTA
24	CGATTCTGACAAGATCGTCAGTTA

Table S6. DNA templates used in the LacR-endonuclease experiments.
All the sequences in this table are parts of 'b' strands.
(A) EcoRV template: 'b' strand is composed of: ~40 bp random sequence + EcoRV recognition sequence (6 bp) + CCCC + spacer (*L*) + *lacO*1 + ~30 bp random sequence (ordered from 5' to 3'). Base contents in the spacer (*L*) are provided in the table.
(B) BamHI template: 'b' strand is composed of: ~40 bp random sequence + BamHI recognition sequence (6 bp) + CTCC + spacer (*L*) + *lacO*3+ ~30 bp random sequence (ordered from 5' to 3'). Base contents in the spacer (*L*) are provided in the table. Some templates were also made with *lacO*1, and those are indicated by *.
*lacO*3 = GGCAGTGAGCGCAACGCAATT

(A) EcoRV template

Template (*L*)	Sequence for *L*
3	TGG
5	AACGC
8	AACGCTTG
10	TGTAGTGAGG
12	AACGCTTGTATG
15	AACGCTCGTATGTTG
16	AACACACCAACGAAGG
18	AACGCTCGTATGTTGTGT
20	AACGCTCGTATGTTGTGTGG
22	AACGCTCGTATGTTCTGTGTGG
25	AGGAGTAGATGACGCTCGTCTGTGG
28	AGGAGTAGATGACGCTCGTCTTCTGTGG
30	AGGAGTAGATGACGCTCGTATGTTCTGTGG

(B) BamHI template

Template (*L*)	Sequence for *L*
0	
3	GTG
5	TCGTG
7	TCGTGTG
8	GTGTGG
9	AACGTGTAGA
10*	TCGTGTGTGG
11	AACGTCGTAGA
12	AACGTCGTATGA
13	AACTCTCTATGAT
14	AACTCTCGTATGAT
15	AACGCTCGTATGAGT
18	AACGCTCGTATGTAGTGG
19	AACGCTCGTATGTAGTTGG
20*	AACGCTCGTATGTTGAGTGG

22	AACGCTCGTATGTTGAGTGGAA
23	AACGCTCGTATGTTGATGTGGAA
25	AACGCTCGTATGTTGATGAGTGGAA
28	AGCTCGTATGTTGATGACTGACGTGGAA
30	AGCTCGTATGTTGATGGACTAGACGTGGAA
35	AGCTCGTATGTAGAAGTGATGGACTAGACGTGGAA

Table S7. DNA primer sequences used to make GRDBD-nucleosome templates.
Full length primers are GCTTAGGTCGAATTGTTATCCGCTCACAATT + sequences in the table, representing GR-8 to GR-27.

Template (L)	Sequence for L
8	CCTGCAGGTGGAGAATCCCGGTG
9	GCCTGCAGGTGGAGAATCCCGGTG
10	GTCCTGCAGGTGGAGAATCCCGGTG
11	GTGCCTGCAGGTGGAGAATCCCGGTG
12	GTGACCTGCAGGTGGAGAATCCCGGTG
13	GTGACCCTGCAGGTGGAGAATCCCGGTG
14	GTGACGCCTGCAGGTGGAGAATCCCGGTG
15	GTGACGCCCTGCAGGTGGAGAATCCCGGTG
16	GTGACGCCCCTGCAGGTGGAGAATCCCGGTG
17	GTGACGCCTCCTGCAGGTGGAGAATCCCGGTG
18	GTGACGCCTGCCTGCAGGTGGAGAATCCCGGTG
19	GTGACGCCTGCCCTGCAGGTGGAGAATCCCGGTG
20	GTGACGCCTGCGCCTGCAGGTGGAGAATCCCGGTG
21	GTGACGCCTGCGCCCTGCAGGTGGAGAATCCCGGTG
22	GTGACGCCTGCGCACCTGCAGGTGGAGAATCCCGGTG
23	GTGACGCCTGCGCATCCTGCAGGTGGAGAATCCCGGTG
24	GTGACGCCTGCGCATGCCTGCAGGTGGAGAATCCCGGTG
25	GTGACGCCTGCGCATGGCCTGCAGGTGGAGAATCCCGGTG
26	GTGACGCCTGCGCATGGACCTGCAGGTGGAGAATCCCGGTG
27	GTGACGCCTGCGCATGGAGCCTGCAGGTGGAGAATCCCGGTG

Table S8. DNA sequences used in the live cell *lacO$_{id}$* experiments.
Constructs for the *in vivo* experiments contain the DNA sequence for O_{id}– spacer –T7 promoter (from 5' to 3') (fig. S16). The base sequence used for the spacer is provided in the table.
O_{id} = AATTGTGAGCGCTCACAATT
-27~-18 common sequence of T7 promoter = GAATTCAAAT
T7 promoter = TAATACGACTCACTATA

Template (*L*)	Sequence for the spacer
3	CGA
4	CGTA
5	CGTTA
6	CGGTTA
7	CGAGTTA
8	CGCAGTTA
9	CGTCAGTTA
10	CGGTCAGTTA
11	CGCGTCAGTTA
12	CGTCGTCAGTTA
13	CGTTCGTCAGTTA
14	CGATTCGTCAGTTA
15	CGATTTCGTCAGTTA
16	CGATTCTCGTCAGTTA
17	CGATTCTTCGTCAGTTA
18	CGATTCTATCGTCAGTTA
19	CGATTCTAATCGTCAGTTA
20	CGATTCTAAATCGTCAGTTA
21	CGATTCTAAGATCGTCAGTTA
22	CGATTCTAAAGATCGTCAGTTA
23	CGATTCTACAAGATCGTCAGTTA
24	CGATTCTAACAAGATCGTCAGTTA
25	CGATTCTACACAAGATCGTCAGTTA
26	CGATTCTAGCACAAGATCGTCAGTTA

Table S9. Primers used to construct *E.coli* strains.

A1	GCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCTGTGACGGAAGATCACTTCG
A2	CAGATGAAAACGGTGTAAAAAAGATAGATACATCAGAGCTTTTACGAGTTATCAAAGGGAAAACTGTCCATAT
B1	GCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAAGGAGATGGCGCCCAACAG
B2	GTTAACGCCATCAAAAATAATTCGCGTCTGGC

Table S10. Parameters for the three simulation systems.

System	Constraint equilibrium position(Å)	Force constant (kcal/mol · Å)	Simulation length (ns)
Normal DNA			156
Constrained DNA(0.5 Å)	3.7	20	18

Table S11. Parameters for the decaying oscillation, obtained by fitting GRDBD-hairpin data. Also, see fig. S12.

	Forward (with 95% confidence bounds)	Reverse (with 95% confidence bounds)
a (amplitude)	1.01 (0.60, 1.42)	1.49 (1.21, 1.78)
b (decay constant)	16.19 (6.76, 25.62)	16.19 (fixed)
c (period)	0.64 (0.60, 0.69)	0.69 (0.64, 0.75)
	9.82 (10.47, 9.11) bp	**9.11 (9.82 8.38) bp**
d (phase)	0.18 (-0.36 0.73)	-2.49 (-2.98, 2.00)
	Phase shift = 152.98º	
e (offset)	1.08 (1.01 1.15)	1.34 (1.24, 1.44)

References and Notes

1. J. Monod, J. Wyman, J. P. Changeux, On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12**, 88 (1965). doi:10.1016/S0022-2836(65)80285-6 Medline

2. D. E. Koshland, Jr., G. Némethy, D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* **5**, 365 (1966). doi:10.1021/bi00865a047 Medline

3. F. M. Pohl, T. M. Jovin, W. Baehr, J. J. Holbrook, Ethidium bromide as a cooperative effector of a DNA structure. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3805 (1972). doi:10.1073/pnas.69.12.3805 Medline

4. M. Hogan, N. Dattagupta, D. M. Crothers, Transmission of allosteric effects in DNA. *Nature* **278**, 521 (1979). doi:10.1038/278521a0 Medline

5. B. S. Parekh, G. W. Hatfield, Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1173 (1996). doi:10.1073/pnas.93.3.1173 Medline

6. J. Rudnick, R. Bruinsma, DNA-protein cooperative binding through variable-range elastic coupling. *Biophys. J.* **76**, 1725 (1999). doi:10.1016/S0006-3495(99)77334-0 Medline

7. D. Panne, T. Maniatis, S. C. Harrison, An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111 (2007). doi:10.1016/j.cell.2007.05.019 Medline

8. R. Moretti *et al*., Targeted chemical wedges reveal the role of allosteric DNA modulation in protein-DNA assembly. *ACS Chem. Biol.* **3**, 220 (2008). doi:10.1021/cb700258r Medline

9. E. F. Koslover, A. J. Spakowitz, Twist- and tension-mediated elastic coupling between DNA-binding proteins. *Phys. Rev. Lett.* **102**, 178102 (2009). doi:10.1103/PhysRevLett.102.178102 Medline

10. H. G. Garcia *et al*., Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. *Cell Reports* **2**, 150 (2012). doi:10.1016/j.celrep.2012.06.004 Medline

11. S. Kim, P. C. Blainey, C. M. Schroeder, X. S. Xie, Multiplexed single-molecule assay for enzymatic activity on flow-stretched DNA. *Nat. Methods* **4**, 397 (2007). Medline

12. Materials and methods are available as supplementary materials on *Science* Online.

13. B. F. Luisi *et al*., Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* **352**, 497 (1991). doi:10.1038/352497a0 Medline

14. M. Newman, T. Strzelecka, L. F. Dorner, I. Schildkraut, A. K. Aggarwal, Structure of Bam HI endonuclease bound to DNA: Partial folding and unfolding on DNA binding. *Science* **269**, 656 (1995). doi:10.1126/science.7624794 Medline

15. F. K. Winkler *et al*., The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12**, 1781 (1993). Medline

16. M. Lewis *et al*., Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247 (1996). doi:10.1126/science.271.5253.1247 Medline

17. C. G. Kalodimos *et al*., Plasticity in protein-DNA recognition: lac repressor interacts with its natural operator 01 through alternative conformations of its DNA-binding domain. *EMBO J.* **21**, 2866 (2002). doi:10.1093/emboj/cdf318 Medline

18. K. J. Durniak, S. Bailey, T. A. Steitz, The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. *Science* **322**, 553 (2008). doi:10.1126/science.1163433 Medline

19. S. B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* **258**, 1122 (1992). doi:10.1126/science.1439819 Medline

20. Z. Bryant *et al*., Structural transitions and elasticity from torque measurements on DNA. *Nature* **424**, 338 (2003). doi:10.1038/nature01810 Medline

21. A. A. Travers, DNA conformation and protein binding. *Annu. Rev. Biochem.* **58**, 427 (1989). doi:10.1146/annurev.bi.58.070189.002235 Medline

22. R. Rohs *et al*., Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **79**, 233 (2010). doi:10.1146/annurev-biochem-060408-091030 Medline

23. A. Újvári, C. T. Martin, Evidence for DNA bending at the T7 RNA polymerase promoter. *J. Mol. Biol.* **295**, 1173 (2000). doi:10.1006/jmbi.1999.3418 Medline

24. G.-Q. Tang, S. S. Patel, T7 RNA polymerase-induced bending of promoter DNA is coupled to DNA opening. *Biochemistry* **45**, 4936 (2006). doi:10.1021/bi0522910 Medline

25. Y. Jia, A. Kumar, S. S. Patel, Equilibrium and stopped-flow kinetic studies of interaction between T7 RNA polymerase and its promoters measured by protein and 2-aminopurine fluorescence changes. *J. Biol. Chem.* **271**, 30451 (1996). doi:10.1074/jbc.271.48.30451 Medline

26. J. H. Miller, *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, 1972).

27. J. Müller, S. Oehler, B. Müller-Hill, Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J. Mol. Biol.* **257**, 21 (1996). doi:10.1006/jmbi.1996.0143 Medline

28. W. Lee *et al*., A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **39**, 1235 (2007). doi:10.1038/ng2117 Medline

29. E. Sharon *et al*., Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* **30**, 521 (2012). doi:10.1038/nbt.2205 Medline

30. S. John *et al*., Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat. Genet.* **43**, 264 (2011). doi:10.1038/ng.759 Medline

31. S. H. Meijsing *et al*., DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407 (2009). doi:10.1126/science.1164265 Medline

32. H. Viadiu, A. K. Aggarwal, The role of metals in catalysis by the restriction endonuclease BamHI. *Nat. Struct. Biol.* **5**, 910 (1998). doi:10.1038/2352 Medline

33. R. Daber, S. Stayrook, A. Rosenberg, M. Lewis, Structural analysis of lac repressor bound to allosteric effectors. *J. Mol. Biol.* **370**, 609 (2007). doi:10.1016/j.jmb.2007.04.028 Medline

34. P. T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19 (1998). doi:10.1006/jmbi.1997.1494 Medline

35. U. Strähle, G. Klock, G. Schütz, A DNA sequence of 15 base pairs is sufficient to mediate both glucocorticoid and progesterone induction of gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7871 (1987). doi:10.1073/pnas.84.22.7871 Medline

36. G.-Q. Tang, R. P. Bandwar, S. S. Patel, Extended upstream A-T sequence increases T7 promoter strength. *J. Biol. Chem.* **280**, 40707 (2005). doi:10.1074/jbc.M508013200 Medline

37. C. Joo, T. Ha, in *Single Molecule Techniques: A Laboratory Manual*, P. R. Selvin, T. Ha, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, 2008), pp. 3–35.

38. E. R. Eismann, B. Müller-Hill, lac repressor forms stable loops in vitro with supercoiled wild-type lac DNA containing all three natural lac operators. *J. Mol. Biol.* **213**, 763 (1990). doi:10.1016/S0022-2836(05)80262-1 Medline

39. J. Elf, G.-W. Li, X. S. Xie, Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* **316**, 1191 (2007). doi:10.1126/science.1141967 Medline

40. J. Chen, K. S. Matthews, Deletion of lactose repressor carboxyl-terminal domain affects tetramer formation. *J. Biol. Chem.* **267**, 13843 (1992). Medline

41. J. M. Gottesfeld *et al*., Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. *J. Mol. Biol.* **309**, 615 (2001). doi:10.1006/jmbi.2001.4694 Medline

42. C. M. Schroeder, P. C. Blainey, S. Kim, X. S. Xie, in *Single Molecule Techniques: A Laboratory Manual*, P. R. Selvin, T. Ha, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, 2008), pp. 461-492.

43. P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, X. S. Xie, A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5752 (2006). doi:10.1073/pnas.0509723103 Medline

44. G. Luo, M. Wang, W. H. Konigsberg, X. S. Xie, Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7

DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12610 (2007). doi:10.1073/pnas.0700920104 Medline

45. W. Gilbert, B. Müller-Hill, Isolation of the lac repressor. *Proc. Natl. Acad. Sci. U.S.A.* **56**, 1891 (1966). doi:10.1073/pnas.56.6.1891 Medline

46. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640 (2000). doi:10.1073/pnas.120163297 Medline

47. E. C. Lee *et al*., A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56 (2001). doi:10.1006/geno.2000.6451 Medline

48. S. Oehler, E. R. Eismann, H. Krämer, B. Müller-Hill, The three operators of the lac operon cooperate in repression. *EMBO J.* **9**, 973 (1990). Medline

49. M. G. Munteanu, K. Vlahovicek, S. Parthasarathy, I. Simon, S. Pongor, Rod models of DNA: Sequence-dependent anisotropic elastic modelling of local bending phenomena. *Trends Biochem. Sci.* **23**, 341 (1998). doi:10.1016/S0968-0004(98)01265-1 Medline

50. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126 (2004). doi:10.1107/S0907444904019158 Medline

51. G. M. T. Cheetham, D. Jeruzalmi, T. A. Steitz, Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature* **399**, 80 (1999). doi:10.1038/19999 Medline

52. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, (Web Server issue), W665 (2004). doi:10.1093/nar/gkh381 Medline

53. H. Li, A. D. Robertson, J. H. Jensen, Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **61**, 704 (2005). doi:10.1002/prot.20660 Medline

54. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037 (2001). doi:10.1073/pnas.181342398 Medline

55. D. Case *et al*., *AMBER 11* (University of California, San Francisco, 2010).

56. H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269 (1987). doi:10.1021/j100308a038

57. A. Pérez *et al*., Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophys. J.* **92**, 3817 (2007). doi:10.1529/biophysj.106.097782 Medline

58. J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327 (1977). doi:10.1016/0021-9991(77)90098-5

59. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1993). doi:10.1063/1.464397

60. H. J. C. Berendsen, J. P. M. Postma, W. F. Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684 (1984). doi:10.1063/1.448118

61. C. M. Falcon, K. S. Matthews, Operator DNA sequence variation enhances high affinity binding by hinge helix mutants of lactose repressor protein. *Biochemistry* **39**, 11074 (2000). doi:10.1021/bi000924z Medline

62. J. Romanuka *et al.*, Specificity and affinity of Lac repressor for the auxiliary operators O_2 and O_3 are explained by the structures of their protein-DNA complexes. *J. Mol. Biol.* **390**, 478 (2009). doi:10.1016/j.jmb.2009.05.022 Medline

63. J. J. Perona, Type II restriction endonucleases. *Methods* **28**, 353 (2002). doi:10.1016/S1046-2023(02)00242-6 Medline

64. I. B. Vipond, G. S. Baldwin, S. E. Halford, Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases. *Biochemistry* **34**, 697 (1995). doi:10.1021/bi00002a037 Medline

65. N. C. Horton, J. J. Perona, Crystallographic snapshots along a protein-induced DNA-bending pathway. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5729 (2000). doi:10.1073/pnas.090370797 Medline

66. J. Israelachvili, *Intermolecular & Surface Forces* (Academic Press, New York, ed. 2, 1991).

67. C. G. Baumann, S. B. Smith, V. A. Bloomfield, C. Bustamante, Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6185 (1997). doi:10.1073/pnas.94.12.6185 Medline

68. J. R. Wenner, M. C. Williams, I. Rouzina, V. A. Bloomfield, Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys. J.* **82**, 3160 (2002). doi:10.1016/S0006-3495(02)75658-0 Medline

69. J. Gore *et al.*, Mechanochemical analysis of DNA gyrase using rotor bead tracking. *Nature* **439**, 100 (2006). doi:10.1038/nature04319 Medline

70. S. Hormeño *et al.*, Mechanical properties of high-G.C content DNA with a-type base-stacking. *Biophys. J.* **100**, 1996 (2011). doi:10.1016/j.bpj.2011.02.051 Medline

71. Y. W. Yin, T. A. Steitz, Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science* **298**, 1387 (2002). doi:10.1126/science.1077464 Medline

72. V. F. Holmes, N. R. Cozzarelli, Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1322 (2000). doi:10.1073/pnas.040576797 Medline

73. S. C. Dillon, C. J. Dorman, Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**, 185 (2010). doi:10.1038/nrmicro2261 Medline

74. A. Jain *et al*., Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484 (2011). doi:10.1038/nature10016 Medline

75. D. Vasudevan, E. Y. D. Chua, C. A. Davey, Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J. Mol. Biol.* **403**, 1 (2010). doi:10.1016/j.jmb.2010.08.039 Medline