

Supporting Information

to

Direct Substitution and Assisted Dissociation Pathways for Turning off Transcription by a MerR-family Metalloregulator

Chandra P. Joshi^{a,1}, Debashis Panda^{a,1}, Danya J. Martell^{a,1}, Nesha May Andoy^a, Tai-Yen Chen^a, Ahmed Gaballa^b, John D. Helmann^b, and Peng Chen^{a,2}

^a Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA. ^b Department of Microbiology, Cornell University, Ithaca, New York 14853, USA.

Table of Content

| | |
|--|----|
| S1. Materials and Methods..... | 6 |
| S2. Experimental Strategy..... | 9 |
| S3. Location of Fluorescence Probes..... | 9 |
| S4. Measurement of Nanovesicle Diameter by Dynamic Light Scattering..... | 10 |
| S5. Apo-CueR and DNA Interactions..... | 11 |
| S6. Dependence of $\langle \tau_1 \rangle^{-1}$ on Protein Concentration..... | 13 |
| S7. Cy3-DNA Interaction with a Mixture of holo-CueR _{Cy5-C129} and holo-CueR _{Cy5-E96C} | 13 |
| S8. Protein Monomer-Dimer Equilibrium and Surface Effects Are Insignificant..... | 14 |
| S9. Comparison of the Dwell Time Distributions from the Direct Immobilization Scheme and Those from the Nanovesicle Trapping Scheme..... | 15 |
| S10. CueR–DNA Dissociation Constant from Ensemble Fluorescence Anisotropy Titration | 17 |
| S11. Derivation of the Probability Density Function $f(\tau)$ and $\langle \tau \rangle^{-1}$ for the Dwell Time τ_2 and τ_1 | 19 |
| S12. Derivation of Probability Density Function $f_0(\tau)$ and $\langle \tau_0 \rangle^{-1}$ for the Dwell Time τ_0 | 23 |
| S13. Simplified Kinetic Mechanism of CueR–DNA Interaction..... | 24 |
| S14. Inter-conversion Rate Constants (k_3 and k_{-3}) of Apo and Holo-CueR–DNA Interaction | 25 |
| S15. The Equilibrium Ratio between Protein-bound Population and Unbound Population...26 | |
| S16. Additional Details about Procedure of Extracting Kinetic Parameters..... | 27 |
| S17. Determination of the Corresponding CueR–DNA Ensemble Dissociation Constants Using Single Molecule Dissociation Constants..... | 27 |
| S18. Analysis of in Vitro Transcription Run-off Assay Results..... | 28 |
| S19. CueR Behaves Similarly in Its Interactions with A Long, 121 Base-pair, DNA Containing the Complete Promoter..... | 28 |
| S20. CueR Can Still Flip Its Binding Orientation on DNA in the Presence of RNAP..... | 30 |
| S21. Apo-CueR Can Directly Substitute Holo-CueR and Vice Versa on DNA..... | 31 |
| S22. Supplementary references..... | 32 |

List of Supporting Figures

Fig. S1. (A) Experimental scheme of surface immobilization of DNA; CueR is supplied in a continuously flowing solution. Upon CueR binding to DNA, FRET occurs from the donor Cy3 (green sphere) to the acceptor (red sphere). The Cy5 location indicated here corresponds approximately to that in CueR_{Cy5-C129}. (B) Gel image of RNA production from the *in vitro* transcription assay in the presence of RNAP only, and in the presence of both RNAP and increasing amounts of holo wild-type (wt) CueR, CueR_{E96C}, or CueR_{C129}. (C) Amount of RNA produced as a function of holo-CueR concentration from B. Solid lines are the fits with Equation S[43]. Error bars are s.d. of five trials in quantifying the RNA amounts from the image in B.9

Fig. S2. (A) The locations of fluorescent probes on a model structure of a CueR–DNA complex. Cy3 is attached to the 5'-end of one of the DNA strands. A CueR monomer can be divided into a DNA binding domain, a dimerization helix and a metal binding domain shown in green, blue, and red, respectively. The other monomer is shown in grey. The Cy3–Cy5 anchor-to-anchor distances are denoted. (B) The cartoon representation of A. Orange color on DNA marks the positions of the dyad symmetric sequence recognized by CueR. 10

Fig. S3. DSL-measured size distribution of the 100-nm lipid nanovesicles prepared via membrane extrusion. Error bars are s.d. The solid line is the fit of the size distribution by a Gaussian function centered at 90 nm with FWHM of ~20 nm. Refractive index for lipid-bilayer used is 1.47 and for buffer 1.33. Measured viscosity of buffer is 0.99. 11

Fig. S4. (A) An E_{FRET} trajectory of an immobilized Cy3-DNA in interaction with 7 nM apo-CueR_{Cy5-C129} in solution. (B) Histogram of E_{FRET} trajectories of apo-CueR_{Cy5-C129}–DNA interactions. Data from >500 E_{FRET} trajectories were combined here. The solid lines are the fits of the E_{FRET} distribution by Voigt functions centered at ~0.07, 0.25, and 0.92, with percentage peak areas of $89.7 \pm 0.1\%$, $5.5 \pm 0.3\%$, and $4.8 \pm 0.2\%$, respectively. [apo-CueR_{Cy5-C129}] = 7 nM. Bin size = 0.005. (C) A comparison between the distributions of τ_2 and τ_1 at [apo-CueR] = 2 nM. These two distributions as well as those in E and F were fitted globally by a sum of two exponentials: $N[A\gamma_1 \exp(-\gamma_1 \tau) + (1-A)\gamma_2 \exp(-\gamma_2 \tau)]$ with $\gamma_1 = 0.35 \pm 0.02 \text{ s}^{-1}$, $\gamma_2 = 1.6 \pm 0.06 \text{ s}^{-1}$, and $A = 0.57 \pm 0.02$. Data compiled from ~700 E_{FRET} trajectories with ~1100 events for τ_2 and ~1300 events for τ_1 ; bin size = 0.15 s. (D) Distributions of τ_1 for holo- and apo-CueR_{Cy5-C129} interactions with DNA, both at 2 nM protein concentrations. Data were collected from ~500 (holo) and ~700 (apo) E_{FRET} trajectories. Bin size: 0.15 s. The two distributions were normalized to the first time point. Solid lines are the fits of the data with a sum of two exponentials: $N[A\gamma_1 \exp(-\gamma_1 \tau) + (1-A)\gamma_2 \exp(-\gamma_2 \tau)]$. For holo: fit parameters are the same as in Fig. 3A; for apo: fit parameters are the same as in C. (E, F) Distributions of $\tau_{2 \rightarrow 0}$ and $\tau_{2 \rightarrow 1}$, and $\tau_{1 \rightarrow 0}$ and $\tau_{1 \rightarrow 2}$ at [apo-CueR_{Cy5-C129}] = 2 nM. Solid lines are global fits with a sum of two exponentials for the apo data in C, E and F. Bin size: 0.15 s. 12

Fig. S5. [CueR] dependence of $\langle \tau_1 \rangle^{-1}$ for CueR_{Cy5-C129}–DNA interactions. Solid lines are fits with Equation 6 (holo) and 7 (apo). 13

Fig. S6. (A) An E_{FRET} trajectory of an immobilized Cy3-DNA in interaction with a mixture of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} of 5 nM each in solution. (B) Histogram of E_{FRET} trajectories of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} interactions with DNA. Data from >500 E_{FRET} trajectories were combined here. The solid lines are the fit of the E_{FRET} histogram by a sum of five Voigt functions centered at ~0.07 (E_0), 0.25 (E_1), 0.92 (E_2), 0.42 (E_1'), and 0.62 (E_2') with percentage peak areas of $88.0 \pm 0.1\%$, $3.3 \pm 0.2\%$, and $2.8 \pm 0.3\%$, $3.2 \pm 0.3\%$, and $2.7 \pm 0.1\%$, respectively. [holo-CueR_{Cy5-C129}] = 5 nM and [holo-CueR_{Cy5-E96C}] = 5 nM. Bin size = 0.005. 14

Fig. S7. CueR–DNA interactions measured inside ~80-nm inner diameter lipid vesicles. (A) Schematic of vesicle trapping of a single CueR–DNA pair. (B) E_{FRET} trajectory of a holo-CueR_{Cy5-C129} molecule interacting with a specific DNA inside a vesicle showing fluctuations between two states. (C) Compiled histogram of ~70 E_{FRET} trajectories of holo-CueR_{Cy5-C129}–DNA interactions inside vesicles. Solid lines are Gaussian resolved peaks centered at ~0.92 and 0.24. Bin size: 0.02. (D) Distribution of τ_2 from (B); bin size: 0.15 s. Solid line is a fit with a sum of two exponentials: $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$; $\gamma_1 = 0.60 \pm 0.02 \text{ s}^{-1}$, $\gamma_2 = 5.0 \pm 0.3 \text{ s}^{-1}$, and $A = 0.80 \pm 0.01$ 15

Fig. S8. (A) The distributions of dwell time τ_2 and τ_1 from holo-CueR_{Cy5-C129}–DNA interactions using the direct immobilization scheme as in Fig. S1A at a protein concentration of 2 nM. (B) Same as in (A) but using the nanovesicle trapping scheme. (C) The comparison between the τ_2 distribution of holo-CueR_{Cy5-C129}–DNA interactions in the direct immobilization experiment (protein concentration ~ 0.5 nM) and that in the nanovesicle experiment. The two distributions were normalized to the first data point. (D) Distributions of $\tau_{1 \rightarrow 0}$ and $\tau_{1 \rightarrow 2}$ at [holo-CueR_{Cy5-C129}] = 2 nM using the direct immobilization scheme. Solid lines are global fits with a sum of two exponentials as for the holo data in A. Bin size: 0.15 s. The two distributions are both double-exponentials and have identical exponents, indicating that the transitions leaving the E_1 state start from the same one kinetic species of the two binding modes within the E_1 state. The distributions in panels of A & D as well as those of panel of Fig 3B (main text) were fitted simultaneously by a sum of two exponentials (solid lines): $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$ with fit parameters: $\gamma_1 = 0.55 \pm 0.01 \text{ s}^{-1}$, $\gamma_2 = 6.4 \pm 0.2 \text{ s}^{-1}$, and $A = 0.68 \pm 0.01$. The distributions in B & C were fitted globally using a sum of two exponentials (solid lines): $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$ with parameters $\gamma_1 = 0.60 \pm 0.02 \text{ s}^{-1}$, $\gamma_2 = 5.0 \pm 0.3 \text{ s}^{-1}$, and $A = 0.80 \pm 0.01$. Data were compiled from 518 E_{FRET} trajectories for A and D, 75 molecules for B, 485 molecules for C. All bin size = 0.15 s. 16

Fig. S9. (A, B) Fluorescence anisotropy titration of Cy3-labelled specific-DNA with holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}. (C) Same as in A but using Cy3-labeled nonspecific-DNA. The experiment was performed at 50 nM Cy3-DNA in 50 mM pH 7.35 Tris buffer with 10 mM NaCl and 2 mM MgCl₂. The solid line is a fit using Equation S[1], with $K_D \sim 3.5 \pm 0.9 \text{ nM}$ for holo-CueR_{Cy5-C129}, $2.6 \pm 0.5 \text{ nM}$ for holo-CueR_{Cy5-E96C} in interactions with the specific DNA, and $K_D \sim 356 \pm 34 \text{ nM}$ for holo-CueR_{Cy5-C129} in interactions with the nonspecific DNA. Note that the x -axes of A, B, and C are the total concentration of the homodimeric CueR in solution, not the concentration of free CueR. (D) [holo-CueR_{Cy5-C129}] dependence of $\langle\tau_{\text{off}}\rangle^{-1}$ for CueR_{Cy5-C129} interactions with the nonspecific DNA. Each data point is an average of the dwell times from ~250 E_{FRET} trajectories. The solid line is linear fit whose slope gives k_1 , the binding rate constant for CueR binding to nonspecific DNA. 18

Fig. S10. The kinetic processes associated with the dwell time τ_2 on the FRET state E_2 . k 's are corresponding kinetic rate constants. The two binding modes of CueR in complex with the DNA in the E_2 state are designated as \mathbf{I}_2 and \mathbf{I}_2' . The free DNA is designated as \mathbf{D} . \mathbf{I}_1 is the equivalent to \mathbf{I}_2 , except that the protein binds in an opposite orientation; and it is kinetically connected from the E_2 state by direct flipping (k_4) and direct substitution (k_{2a}). $[\mathbf{P}]$ stands for the protein concentration. Note there is a factor of $1/2$ in front of $k_{2a}[\mathbf{P}]$ for the direct substitution because the incoming protein that replaces the incumbent protein can land in two different orientations and only one of them has the other orientation leading to a transition to the E_1 state. k_{2b} is the rate constant for the protein assisted dissociation. 20

Fig. S11. The kinetic processes related to the dwell time τ_0 on the FRET state E_0 . Note there is a $1/2$ factor in front of $k_1[\mathbf{P}]$ because for each protein binding, only 50% of time it will bind in one of the two orientations. 23

Fig. S12. Simplified kinetic mechanism of CueR interacting with a specific DNA, where the fluorescent probes on CueR and DNA are removed. As such, the two protein binding orientations are identical (i.e., $\mathbf{I}_2 = \mathbf{I}_1 \equiv \mathbf{I}$, $\mathbf{I}_2' = \mathbf{I}_1' \equiv \mathbf{I}'$; reference Fig. 5 in the main text).25

Fig. S13. The numerical values of k_3 and k_{-3} determined at various protein concentrations for (A) holo-CueR and (B) apo-CueR interactions with DNA. The solid lines are fits with horizontal lines. The error bars are s.d.26

Fig. S14. Protein concentration dependence of ratio of histogram area ratio to $N_{2 \rightarrow 1}/N_{2 \rightarrow 0}$ (i.e., “Branching ratio” in the plot) for apo-CueR_{Cy5-C129}-DNA interactions. Solid line is a fit with Equation S[41].27

Fig. S15. Holo-CueR_{Cy5-C129} interactions with the 121 base-pair DNA. (A) The section of 121 base-pair DNA around the promoter region of the *copA* gene spanning from -60 to +61. The -35 and -10 sequences are shown as boxes. The dyad symmetric sequence that CueR recognizes is shown in pink. The Cy5 dye is attached at -41 position (red vertical arrow) via an amine modified T. The Biotin-TEG is attached to the 5'-end of +61 position for immobilizing the DNA. The experimentally determined CueR footprint (1) is shaded in grey. The 25 base-pair specific DNA used in this paper is underlined, and it overlaps almost entirely with the CueR footprint. (B) An E_{FRET} trajectory at 2 nM holo-CueR_{Cy5-C129}. Buffer: 50 mM pH 7.35 Tris, 100 mM Potassium Glutamate, 0.1 mg/mL BSA, 2 mM MgCl₂, 1 mM CaCl₂, 2 mM DTT, 0.1 mM EDTA, 5% glycerol, and 1 mM Trolox. (C) Histogram of the E_{FRET} trajectories. Data from 85 E_{FRET} trajectories were combined here. The solid lines are the fits of the E_{FRET} histogram by Voigt functions centered at ~0.07, 0.23, and 0.81, with percentage peak areas of $81.4 \pm 0.3\%$, $9.9 \pm 0.2\%$, and $8.7 \pm 0.3\%$, respectively. [holo-CueR_{Cy5-C129}] = 2 nM. Bin size = 0.002. (D) Distribution of τ_2 , at 2 nM protein concentration, in comparison with that for 25 base-pair DNA from Figure 3A. Data compiled from ~1300 E_{FRET} trajectories. Bin size: 0.15 s. The two distributions were normalized to the first data point for comparison. Solid lines are fits with a sum of two exponentials: $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$. For 25 base-pair: $\gamma_1 = 0.55 \pm 0.01 \text{ s}^{-1}$, $\gamma_2 = 6.4 \pm 0.2 \text{ s}^{-1}$, $A = 0.68 \pm 0.01$; for 121 base-pair: $\gamma_1 = 0.92 \pm 0.09 \text{ s}^{-1}$, $\gamma_2 = 8.9 \pm 0.8 \text{ s}^{-1}$, $A = 0.40 \pm 0.02$. (E) [holo-CueR_{Cy5-C129}] dependence of $\langle \tau_2 \rangle^{-1}$ in comparison with that for the 25 base-pair DNA from Figure 4A. Solid lines are fits with Equation 6 as in Figure 4A.29

Fig. S16. (A) An E_{FRET} trajectory of an immobilized 121 base-pair Cy3-DNA in interaction with 2 nM holo-CueR_{Cy5-C129} in the presence of 5 nM RNAP in solution. (B) Histogram of the E_{FRET} trajectories of holo-CueR_{Cy5-C129} interactions as in A. Data from 80 E_{FRET} trajectories were combined here. The solid lines are the fits of the E_{FRET} histogram by Voigt functions centered at ~0.07, 0.18, and 0.77, with percentage peak areas of $77.4 \pm 0.5\%$, $12.4 \pm 0.3\%$, and $10.2 \pm 0.4\%$, respectively. Bin size = 0.002. (C) Histogram of E_{FRET} trajectories of holo-CueR_{Cy5-C129}-DNA interactions in the absence and presence of 5 nM RNAP. The histogram in the absence of RNAP is from Fig. S15C. The arrows indicate the E_1 and E_2 peaks. Buffer: 50 mM pH 7.35 Tris, 100 mM Potassium Glutamate, 0.1 mg/mL BSA, 2 mM MgCl₂, 1 mM CaCl₂, 2 mM DTT, 0.1 mM EDTA, 5% glycerol and 1 mM Trolox.30

Fig. S17. (A) An E_{FRET} trajectory of an immobilized 25 base-pair Cy3-DNA in interaction with a mixture of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} of about 5 nM each in solution. The blue arrows denote the transitions from the holo-protein bound states to the apo-protein bound states, and the black arrows denote the reverse transitions; these transitions report the direct substitution of a DNA-bound holo-protein by an apo-protein or the reverse. This figure is the same as Fig. 1E in the main text. (B) Histogram of the E_{FRET} trajectories of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} interactions with the specific DNA. Data from ~90 E_{FRET} trajectories were combined here. The solid lines are the fit of the E_{FRET} histogram by a sum of five Voigt functions centered at ~0.07 (E_0), 0.25 (E_1), 0.92 (E_2), 0.42 (E_1'), and 0.62 (E_2')

with percentage peak areas of $79.0 \pm 0.2\%$, $5.5 \pm 0.8\%$, and $3.8 \pm 0.6\%$, $6.4 \pm 0.5\%$, and $5.3 \pm 0.6\%$, respectively. $[\text{apo-CueR}_{\text{Cy5-C129}}] \approx 5 \text{ nM}$ and $[\text{holo-CueR}_{\text{Cy5-E96C}}] \approx 5 \text{ nM}$. Bin size = 0.005.31

List of Supporting Tables

Table S1. Dissociation constants for CueR–DNA interactions measured at various conditions and using different methods. 19

S1. Materials and Methods

A. Protein mutagenesis, expression, purification, fluorescence labeling, and copper removal

To label CueR with the FRET acceptor Cy5 specifically, we used site-directed mutagenesis to make two CueR variants that contain a unique cysteine in its monomer, apart from the Cu⁺ binding cysteines. Each monomer of the wild-type CueR has four cysteines: two of them (C112 and C120) bind Cu⁺ and can be protected from labeling via binding to Cu⁺, and the other two cysteines are C129 and C130. In one CueR variant (CueR-C130S; referred to as CueR_{C129} hereafter), the natural solvent-accessible C129 is used for labeling and it is located at the metal-binding domain near the C-terminus (Fig. S2A). In the second CueR variant (CueR-C129S/C130S/E96C; referred to as CueR_{E96C} hereafter), the solvent-accessible E96C is used for labeling and it is located in the middle of the dimerization helix of CueR (Fig. S2A).

Both CueR variants were cloned in a pET30a vector and the sequence confirmed. The proteins were expressed in BL21(DE3) strain and purified as previously described (1, 2). The cells were grown until OD₆₀₀ ~ 0.6 before 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added. After an additional 4 h growth at 37°C, cells were harvested by centrifugation and then disrupted by French press in lysis buffer (10 mM Tris, 300 mM NaCl, 10 mM beta-mercaptoethanol (BME), and 10% glycerol at pH 7.3). The protein in the supernatant was purified first by precipitating with 45% (NH₄)₂SO₄ and then by gel filtration (HiPrep 26/10, GE Healthcare). The collected fractions were further purified through a Heparin affinity column (16/10 Heparin FF, GE Healthcare), a gel filtration column (HILOAD 26/60 Superdex 200 PR, GE Healthcare), and an anion exchange column (Mono Q 5/50 GL, GE Healthcare). Protein purity was confirmed by SDS-PAGE, quantified using bicinchoninic acid (BCA) assay (Pierce), and stored at -80°C in 50 mM pH 7.0 Tris Buffer with ~250 mM NaCl and 30% glycerol. Protein identity was confirmed by mass spectrometry (CueR_{C129}: predicted monomer mass, 15220 Da, observed, 15203 Da; CueR_{E96C}: predicted monomer mass, 15180 Da, observed, 15158 Da). The purified CueR is in its apo-form (2).

The protein was labeled with Cy5 at the targeted cysteine via maleimide chemistry. CueR was initially converted to the holo-form by adding CuSO₄ solution ([CueR monomer]:[Cu] = 1:1.5) in the presence of excess TCEP. TCEP reduces Cu²⁺ to Cu⁺ that CueR binds; it also reduces potential disulfide bonds. The Cu⁺ binding protects the metal-binding cysteines of CueR from dye labeling. Cy5-maleimide (Invitrogen) was added to the holo-CueR solution ([dye]:[CueR monomer] = 3:1) in 100 mM phosphate buffer solution at pH 7. The reaction mixture was kept on shaker at 4°C for ~16 hours and then quenched by adding excess BME. After incubating for additional 2 hours, the excess dye was removed through gel filtration (Superdex peptide 10/300 GL, GE Healthcare). Since CueR is a homodimer, the labeling reaction generates a mixture of unlabeled, mono-labeled and bi-labeled species. The mono-labeled fraction was purified using anion exchange column (Mono Q 5/50 GL) and has a dye:protein ratio of ~0.8. The extinction coefficient of 250,000 M⁻¹cm⁻¹ at 650 nm was used for determining the Cy5 concentration. Similarly, the extinction coefficient of 10,361 M⁻¹cm⁻¹ at 280 nm was used for determining the CueR concentration; this extinction coefficient was calibrated using the BCA protein quantification assay.

The labeled holo-CueR could be converted to the apo form by removing the bound copper using KCN (at 1000× excess) (3). The KCN-protein solution was incubated for 4 hours at room temperature and overnight at 4°C, before removing extracted copper and excess KCN through a desalting column (HiTrap, GE Healthcare). The dye:protein ratio of the resulting apo protein was ~0.9. The copper content of the resulted apo protein was confirmed to be <4% via BCA copper quantification assay (4).

B. DNA preparation, purification and fluorescence labeling

The Cy3 and biotin tagged DNA oligomeric strands were purchased from Integrated DNA Technologies (Coralville, IA) and dissolved in 10 mM Tris buffer solution with 100 mM NaCl at pH 7.3. Two types of 25 base-pair double-strand DNA (dsDNA) constructs were used. One construct is from the *copA* gene promoter and contains the specific dyad sequence recognized by CueR: 5'/Cy3/TGACCTTCCCCTTGCTGGAAGGTTT-3' and the complementary 5'-/BiotTEG/AAACCTTCCAGCAAGGGGAAGGTCA-3'. Underlined is the specific sequence. The other construct is a control DNA that does not contain the CueR-recognition sequence: 5'-/Cy3/TTGACTCTATAGTAACTAGAGGGTG-3', and the BioTEG-tagged complementary strand. Non-biotinylated dsDNA was used in the nanovesicle trapping experiment. The DNA strands were annealed together, purified through Mono Q anion exchange column, and stored in 50 mM pH 7.35 Tris buffer with ~250 mM NaCl. The purity of dsDNA was confirmed by gel electrophoresis.

C. Single-molecule fluorescence experiments and data analysis

The single-molecule fluorescence experiments were performed as previously described (2). Briefly, a prism-type total internal reflection microscope based on an Olympus IX71 inverted microscope was used. The Cy3 probe on DNA was directly excited by a continuous-wave circularly polarized 532-nm laser of ~6 mW focused onto an area of ~150 × 75 μm² on the sample. The fluorescence of both Cy3 and Cy5 was collected by a 60× NA 1.2 water-immersion objective and split by a dichroic mirror into two channels using a Dual-View system (Optical Insights). The HQ550LP filter was used to reject the excitation laser light and each channel of fluorescence was further filtered (HQ580-60m or HQ660LP) and projected onto one half of the imaging area of an EMCCD camera (Andor Ixon) controlled by Andor IQ software. The time resolution for all the single-molecule experiments was 50 ms. A custom IDL program was then used to extract individual fluorescence trajectories of Cy3 and Cy5 for individual immobilized DNA molecules interacting with CueR from the fluorescence movie recorded by the camera. The FRET efficiency was computed using the relationship: $I_{Cy5}/(I_{Cy5}+I_{Cy3})$, where I_{Cy3} and I_{Cy5} are the fluorescence intensities. FRETulator, a home-written Visual basic program, was used to obtain the FRET histograms and the individual waiting times. A forward-backward nonlinear filter implemented in MatLab was used to reduce the noise in the fluorescence trajectories (5-7).

The samples were contained in a microfluidic channel, formed by double-sided tape sandwiched between a quartz slide and a borosilicate cover slip. Quartz slides were first amine-functionalized with Vectabond (Vector Laboratories) and then coated with polyethylene glycol (PEG) polymers (Nanocs, 100 mg/mL m-PEG-SPA-5000 and 1 mg/mL biotin-PEG-NHS-3400) to minimize nonspecific protein and DNA adsorption on quartz surfaces. One percent of the PEG polymers contain a biotinylated terminal group to form biotin-neutravidin linkages for immobilizing biotinylated DNA molecules (Fig. S1A). The neutravidin (Invitrogen) was introduced as 500 μL of 0.2 mg/mL solution and incubated for

15 min. Any possible remaining bare quartz surface patches were further blocked using 3 mL BSA (0.1 mg/ml). ~500 μ L of 10 pM Cy3-labeled biotinylated DNA solution in 50 mM Tris, 2 mM MgCl₂ buffer at pH 7.35 was flowed through the channel for immobilization. The CueR solution of 0.5 to 10 nM containing an oxygen scavenging system (0.1 mg/mL glucose oxidase, 0.025 mg/mL catalase, 4% glucose and 1 mM Trolox) in the same buffer was flowed in continuously at a rate of 10 μ L/min for fluorescence imaging.

D. Nanovesicle trapping of protein and DNA

The nanovesicle trapping experiment was performed as described previously (8-10). Briefly, a mixture of L-phosphatidylcholine (eggPC) and 2% 1,2-dipalmitoyl sn-glycero-3-phosphoethanolamine- N-(cap biotinyl) (16:0 biotinyl cap PE) (Avanti Lipids) in chloroform was dried under nitrogen. Protein-loaded vesicles were prepared by hydrating the lipid film with a 300 μ L solution containing 1 μ M Cy5 labeled CueR, 1 μ M Cy3 labeled DNA, 1 mM Trolox, and 5 mM TCEP in 50 mM Tris, 10 mM NaCl, 2mM MgCl₂ at pH 7.35. Conditions for non-specific DNA were the same as above, except the DNA and protein concentrations were 2 μ M and 3 μ M, respectively. The solution was then repeatedly extruded through a polycarbonate membrane with 100 nm pores (Avanti) to form approximately 100-nm diameter unilamellar vesicles encapsulating the protein and DNA. The inner diameter of the nanovesicles is 80 ± 20 nm, calculated from the outer diameter (90 ± 20 nm) measured by dynamic light scattering (Fig. S3; Malvern Zetasizer Nano-ZS) and a membrane thickness of 5 nm (11) . The corresponding effective concentration for a single molecule trapped inside the nanovesicle is 6 ± 5 μ M. Loaded vesicles were used for experiments immediately or within 48 h of preparation. The quartz surface inside the flow cell was first coated by biotinylated BSA (1 mL of 2 mg/mL solution flow through and 30 min incubation), neutravidin (500 μ L of 0.2 mg/mL solution flow through and 10 min incubation), and then 300 μ L of a 100-200 pM nanovesicle solution was flowed in buffer that contains the oxygen scavenging system.

E. *In vitro* run-off transcription assay

Following reported procedures (3, 12, 13) we performed an *in vitro* transcription run-off assay to confirm that the two CueR variants, CueR_{C129} and CueR_{E96C}, are still transcriptionally active. A 315-bp DNA fragment (referred as *PcopA*), spanning the promoter and part of the *E. coli copA* gene, was copied out of the *E. coli* genome through PCR using the following primers: 5'-TCTTTACGGACTTTTACCCGCCTGG-3' and 5'-CCTTTGGGTGGCTTACAGATGCGTC-3'. This fragment was used as the template for *in vitro* transcription. Run-off transcription reactions were performed by incubating 20 μ g/mL *PcopA* DNA template, with 0, 5, 10, 30, 50, 100 or 200 nM concentrations of apo CueR mutants, and 15 nM *E. coli* RNA polymerase (Epicenter) in a buffer solution (100 mM potassium Glutamate, 10 mM pH 8 Tris-base, 1 mM MgCl₂, 0.1 mM EDTA, 100 μ g/mL of acetylated-BSA, 1 mM CaCl₂, 5% glycerol, 2 mM dithiothreitol (DTT), 10 μ M BCA, and 1 μ M [Cu(CH₃CN)₄]PF₆) for 10 minutes at 37°C in a total volume of 20 μ L. 2.5 μ L of an NTP mixture containing 500 μ M of ATP, GTP, CTP and UTP, 0.5 μ L of [α -³²P]UTP and 0.25 μ L of RNase inhibitor were added to the reaction mixture and incubated at 37°C for another 20 minutes. Following incubation, a mixture of 70 μ L of water, 10 μ L of 3M NaOAc (pH 5.2, adjusted with acetic acid), and 2 μ L of 0.5 M EDTA was added to the reaction mixture. The RNA product was precipitated overnight at -20°C after the addition of 2 μ L of glycogen blue and 330 μ L of ethanol. RNA was re-suspended in 10 μ L of 80% formamide, 1X TBE buffer,

10 mM EDTA, 0.05% SDS and 0.025% bromophenol blue and kept at 100°C for 2 minutes. RNA samples were loaded onto a 6% urea polyacrylamide gel and electrophoresed at 1000 V. The results show that both CueR mutants are active transcriptional factors and have quantitatively similar activity to the wild type CueR (Fig. S1B, C).

S2. Experimental Strategy

Fig. S1A shows our experimental design of using smFRET and surface immobilization to probe CueR–DNA interactions, as described in the text. One end of the DNA had a FRET donor Cy3 attached and the other end on a different strand had a biotin for its surface immobilization via an neutravidin-biotin linkage. The homodimeric CueR was labeled with a single FRET acceptor Cy5 and supplied in a flowing solution across the immobilized DNA. Upon CueR binding to DNA, FRET occurs, causing changes in the fluorescence intensities of Cy3 and Cy5. By monitoring the fluorescence intensities of Cy3 and Cy5 simultaneously, we studied single-pair CueR–DNA interactions in real time.

We studied two CueR variants, with the Cy5 located at two different locations (see Section S3). Fig. S1B and C show that the two CueR variants are both as active as the wild type in activating the transcription controlled by the *copA* promoter (Section S1E).

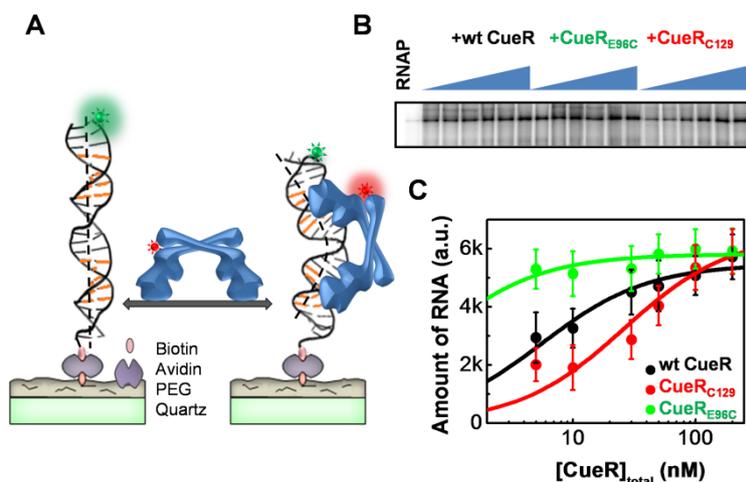


Fig. S1. (A) Experimental scheme of surface immobilization of DNA; CueR is supplied in a continuously flowing solution. Upon CueR binding to DNA, FRET occurs from the donor Cy3 (green sphere) to the acceptor (red sphere). The Cy5 location indicated here corresponds approximately to that in CueR_{Cy5-C129}. (B) Gel image of RNA production from the *in vitro* transcription assay in the presence of RNAP only, and in the presence of both RNAP and increasing amounts of holo wild-type (wt) CueR, CueR_{E96C}, or CueR_{C129}. (C) Amount of RNA produced as a function of holo-CueR concentration from B. Solid lines are the fits with Equation S[43]. Error bars are s.d. of five trials in quantifying the RNA amounts from the image in B.

S3. Location of Fluorescence Probes

Fig. S2A shows the locations of the two fluorescent probes Cy5 and Cy3 on a model structure of a CueR–DNA complex. The model was generated by overlaying in PyMol the

holo-CueR structure (3) on top of the structure of MtaN–DNA complex (14, 15), where MtaN is a homologue of CueR as well as a MerR-family regulator that responds to organic drugs instead of metal ions.

Two CueR constructs were made for the single molecule experiment. In one CueR construct, CueR_{C129}, the Cy5 is attached to the surface-exposed C129; this labeling position is located within the metal-binding domain of CueR near the C-terminal end (Fig. S2A). In the second CueR construct, CueR_{E96C}, the Cy5 is attached to the surface-exposed non-conserved E96C; this labeling position is located at the center of the dimerization helix (Fig. S2A).

The probe in a singly labeled CueR breaks the symmetry of the CueR homodimer, which can thus bind to a terminally-labeled DNA with two different orientations. Based on the model structure (Fig. S2A), the Cy3–Cy5 anchor-to-anchor distances in a CueR_{Cy5-C129}–DNA complex are about 68 Å and 45 Å for the two binding orientations. The experimentally measured E_{FRET} values corresponding to these two binding orientations are 0.25 and 0.92, respectively (Fig. 1F). Similarly, the Cy3–Cy5 distances in a CueR_{Cy5-E96C}–DNA complex are about 61 Å and 53 Å for the two orientations, respectively. The corresponding experimentally measured E_{FRET} values are 0.42 and 0.62, respectively (Fig. 1G). The cartoon representations such as that in Fig. S2B are used throughout the text to depict CueR–DNA complexes.

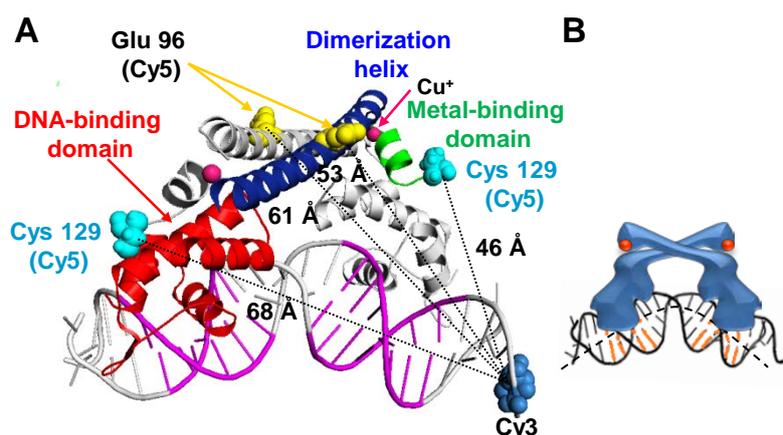


Fig. S2. (A) The locations of fluorescent probes on a model structure of a CueR–DNA complex. Cy3 is attached to the 5'-end of one of the DNA strands. A CueR monomer can be divided into a DNA binding domain, a dimerization helix and a metal binding domain shown in green, blue, and red, respectively. The other monomer is shown in grey. The Cy3–Cy5 anchor-to-anchor distances are denoted. (B) The cartoon representation of A. Orange color on DNA marks the positions of the dyad symmetric sequence recognized by CueR.

S4. Measurement of Nanovesicle Diameter by Dynamic Light Scattering

The 100-nm lipid nanovesicles were made through membrane extrusion (see Section S1D). We further measured the nanovesicle size distribution using dynamic light scattering (DSL, Malvern Zetasizer Nano-ZS) (Fig. S3). The measured hydrodynamic size of the nanovesicles is 90 ± 20 nm. Using a membrane thickness of ~ 5 nm (11), the inner diameter of

nanovesicles is 80 ± 20 nm. Correspondingly, the effective concentration for a single molecule trapped inside is 6 ± 5 μ M.

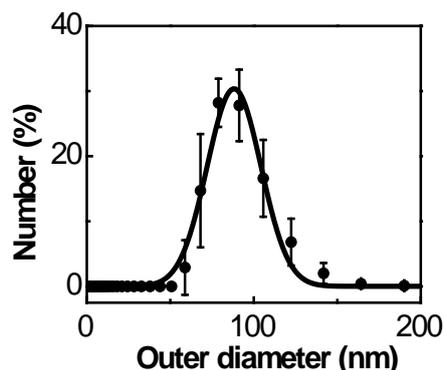


Fig. S3. DSL-measured size distribution of the 100-nm lipid nanovesicles prepared via membrane extrusion. Error bars are s.d. The solid line is the fit of the size distribution by a Gaussian function centered at 90 nm with FWHM of ~ 20 nm. Refractive index for lipid-bilayer used is 1.47 and for buffer 1.33. Measured viscosity of buffer is 0.99.

S5. Apo-CueR and DNA Interactions

Fig. S4A presents a representative E_{FRET} trajectory of apo-CueR_{Cy5-C129}–DNA interactions. Fig. S4B presents the E_{FRET} histogram of ~ 500 such trajectories. Both the trajectory and the histogram clearly resolve three E_{FRET} states: E_0 , E_1 , and E_2 , corresponding to the unbound state (i.e., free DNA form) and two orientations of the CueR–DNA complex, as we discussed in the main text. The E_{FRET} values (Fig. S4A and B) are similar to those obtained in holo-CueR_{Cy5-C129}–DNA interactions (Fig. 1A and F).

Fig. S4C presents the distributions of τ_2 and τ_1 in the presence of 2 nM apo-CueR_{Cy5-C129}. The two distributions are identical within experimental error, consistent with that the E_2 and E_1 states are equivalent, corresponding to the two CueR binding orientations. Both distributions follow double exponential decay behavior, reflecting the two different binding modes of apo-CueR on DNA, similar to holo-CueR–DNA interactions as discussed in the main text (Fig. 3A). Fig. S4D compares the τ_1 distributions of apo- and holo-CueR_{Cy5-C129} interacting with DNA, parallel to the comparison of τ_2 distributions in Fig. 3A.

Moreover, both τ_2 and τ_1 can be divided into two sub-types. For τ_2 : $\tau_{2 \rightarrow 1}$ and $\tau_{2 \rightarrow 0}$, depending on whether an $E_2 \rightarrow E_1$ or $E_2 \rightarrow E_0$ transition concludes a τ_2 period. For τ_1 : $\tau_{1 \rightarrow 2}$ and $\tau_{1 \rightarrow 0}$, depending on whether an $E_1 \rightarrow E_2$ or $E_1 \rightarrow E_0$ transition concludes a τ_1 period. For apo-CueR_{Cy5-C129}–DNA interactions, $\tau_{2 \rightarrow 1}$ and $\tau_{2 \rightarrow 0}$ follow the same double-exponential distribution with identical exponents (Fig. S4E), and $\tau_{1 \rightarrow 2}$ and $\tau_{1 \rightarrow 0}$ follow the same double-exponential distribution with identical exponents (Fig. S4F). This indicates that upon leaving the E_2 (or E_1) state, transitions to the E_1 (or E_2) or E_0 state must start from the same one kinetic species of the two binding modes within the E_2 (or E_1) state (see Section S11 for the derivation of the probability density functions of $\tau_{2 \rightarrow 1}$ and $\tau_{2 \rightarrow 0}$, as well as $\tau_{1 \rightarrow 2}$ and $\tau_{1 \rightarrow 0}$).

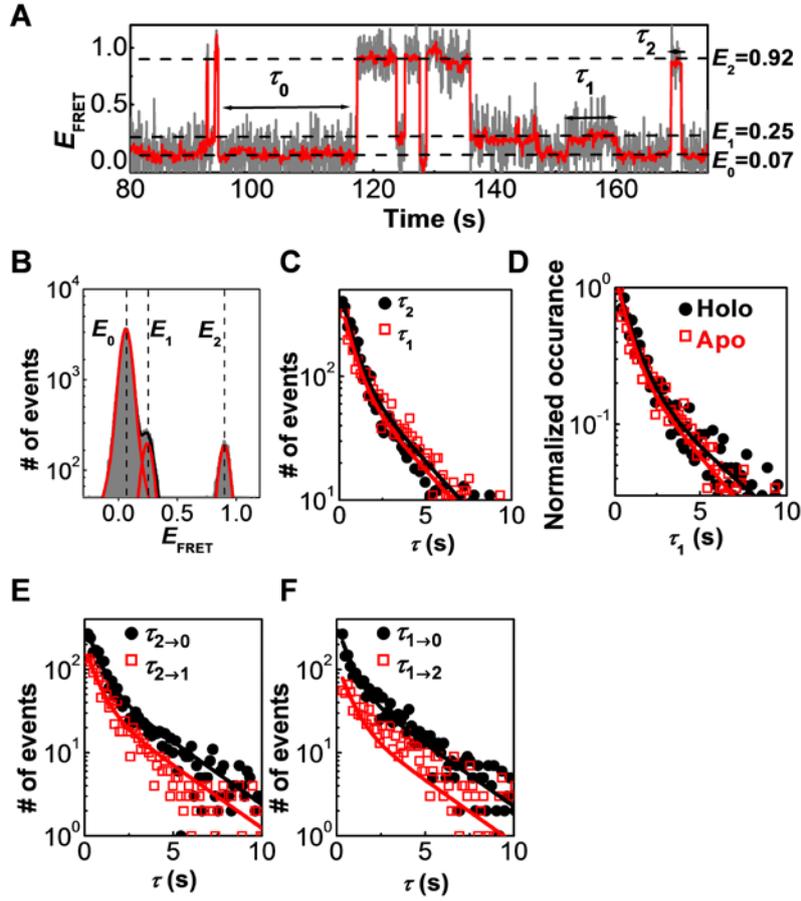


Fig. S4. (A) An E_{FRET} trajectory of an immobilized Cy3-DNA in interaction with 7 nM apo-CueR_{Cy5-C129} in solution. (B) Histogram of E_{FRET} trajectories of apo-CueR_{Cy5-C129}-DNA interactions. Data from >500 E_{FRET} trajectories were combined here. The solid lines are the fits of the E_{FRET} distribution by Voigt functions centered at ~ 0.07 , 0.25 , and 0.92 , with percentage peak areas of $89.7 \pm 0.1\%$, $5.5 \pm 0.3\%$, and $4.8 \pm 0.2\%$, respectively. [apo-CueR_{Cy5-C129}] = 7 nM. Bin size = 0.005. (C) A comparison between the distributions of τ_2 and τ_1 at [apo-CueR] = 2 nM. These two distributions as well as those in E and F were fitted globally by a sum of two exponentials: $N[A\gamma_1 \exp(-\gamma_1 \tau) + (1-A)\gamma_2 \exp(-\gamma_2 \tau)]$ with $\gamma_1 = 0.35 \pm 0.02 \text{ s}^{-1}$, $\gamma_2 = 1.6 \pm 0.06 \text{ s}^{-1}$, and $A = 0.57 \pm 0.02$. Data compiled from ~ 700 E_{FRET} trajectories with ~ 1100 events for τ_2 and ~ 1300 events for τ_1 ; bin size = 0.15 s. (D) Distributions of τ_1 for holo- and apo-CueR_{Cy5-C129} interactions with DNA, both at 2 nM protein concentrations. Data were collected from ~ 500 (holo) and ~ 700 (apo) E_{FRET} trajectories. Bin size: 0.15 s. The two distributions were normalized to the first time point. Solid lines are the fits of the data with a sum of two exponentials: $N[A\gamma_1 \exp(-\gamma_1 \tau) + (1-A)\gamma_2 \exp(-\gamma_2 \tau)]$. For holo: fit parameters are the same as in Fig. 3A; for apo: fit parameters are the same as in C. (E, F) Distributions of $\tau_{2 \rightarrow 0}$ and $\tau_{2 \rightarrow 1}$, and $\tau_{1 \rightarrow 0}$ and $\tau_{1 \rightarrow 2}$ at [apo-CueR_{Cy5-C129}] = 2 nM. Solid lines are global fits with a sum of two exponentials for the apo data in C, E and F. Bin size: 0.15 s.

S6. Dependence of $\langle \tau_1 \rangle^{-1}$ on Protein Concentration

For the microscopic dwell time τ_1 , $\langle \tau_1 \rangle^{-1}$ represents the rate of leaving the protein-bound state E_1 . $\langle \tau_1 \rangle^{-1}$ was observed to be dependent on the concentration of protein in the solution (Fig. S5): it increases linearly with increasing concentration of either holo- or apo-CueR. This behavior is similar to that of $\langle \tau_2 \rangle^{-1}$ (Fig. 4A), as expected, and it indicates that a protein molecule coming from the surrounding solution must interact with the protein molecule in complex with DNA to cause it to leave the E_1 state, leading to either the state E_2 or the state E_0 . The former corresponds to a direct protein substitution and the latter to a protein-assisted dissociation, as discussed in the main text on the protein concentration dependence of $\langle \tau_2 \rangle^{-1}$ (reference Fig. 4A).

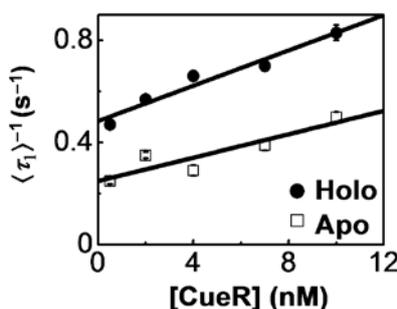


Fig. S5. [CueR] dependence of $\langle \tau_1 \rangle^{-1}$ for CueR_{Cy5-C129}-DNA interactions. Solid lines are fits with Equation 6 (holo) and 7 (apo).

S7. Cy3-DNA Interaction with a Mixture of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}

To verify the direct protein substitution process in holo-CueR-DNA interactions, we performed an experiment using a mixture of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}, whose DNA-bound complexes have distinct E_{FRET} values (Fig. 1A and F vs. Fig. 1B and G). Fig. S6A (same as Fig. 1D) presents an E_{FRET} trajectory of Cy3-DNA interaction with a mixture of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}. Fig. S6B presents the E_{FRET} histogram of ~ 500 such trajectories. Both the trajectory and the histogram clearly resolve five E_{FRET} states: E_0 , E_1 , E_1' , E_2' , and E_2 , corresponding to the unbound state (i.e., free DNA form), the two binding orientations of the holo-CueR_{Cy5-C129}-DNA complexes (E_1 and E_2), and the two binding orientations of the holo-CueR_{Cy5-E96C}-DNA complexes (E_1' and E_2'), as we discussed in the main text (Fig. 1F and G). Most importantly, in the E_{FRET} trajectory (Fig. S6A), there are direct transitions between the bound states of holo-CueR_{Cy5-C129}-DNA complexes (i.e., E_1 and E_2) and those of holo-CueR_{Cy5-E96C}-DNA complexes (i.e., E_1' and E_2'). These transitions directly show the exchange of a holo-CueR on DNA, i.e., direct protein substitution.

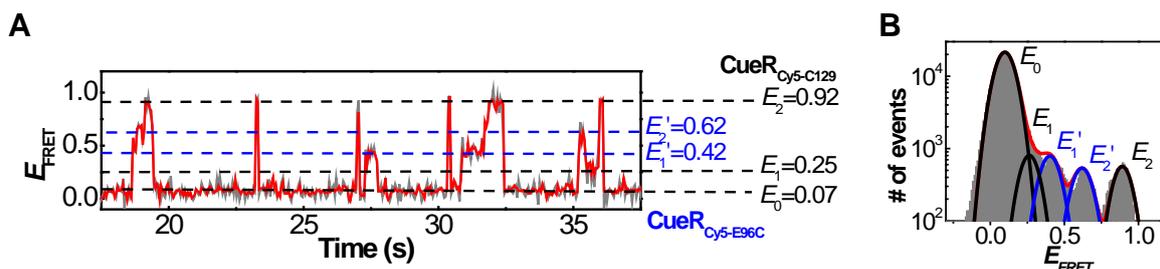


Fig. S6. (A) An E_{FRET} trajectory of an immobilized Cy3-DNA in interaction with a mixture of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} of 5 nM each in solution. (B) Histogram of E_{FRET} trajectories of holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} interactions with DNA. Data from >500 E_{FRET} trajectories were combined here. The solid lines are the fit of the E_{FRET} histogram by a sum of five Voigt functions centered at ~ 0.07 (E_0), 0.25 (E_1), 0.92 (E_2), 0.42 (E_1'), and 0.62 (E_2') with percentage peak areas of $88.0 \pm 0.1\%$, $3.3 \pm 0.2\%$, and $2.8 \pm 0.3\%$, $3.2 \pm 0.3\%$, and $2.7 \pm 0.1\%$, respectively. $[\text{holo-CueR}_{\text{Cy5-C129}}] = 5 \text{ nM}$ and $[\text{holo-CueR}_{\text{Cy5-E96C}}] = 5 \text{ nM}$. Bin size = 0.005.

S8. Protein Monomer-Dimer Equilibrium and Surface Effects Are Insignificant

CueR is a homodimer. The dissociation constant for this dimerization is not known, although we expect it to be in the nanomolar regime because the physiological concentration of CueR in *E. coli* is about $\sim 400 \text{ nM}$ (see Discussion section in the main text). In our single-molecule experiments we flowed in nanomolar concentrations of CueR. To ensure that a dynamic monomer-dimer equilibrium does not play a significant role in the CueR–DNA interaction kinetics observed here, we further used a vesicle trapping scheme to study holo-CueR_{Cy5-C129}–DNA interactions, a scheme we previously applied to study dynamic protein–protein interactions (8-10).

In this scheme we trapped a single pair of Cy3-labeled DNA and holo-CueR_{Cy5-C129} inside a $\sim 80\text{-nm}$ inner diameter lipid vesicle, which is immobilized on the surface (Fig. S7A, Section S1D). Because of the confined volume, the effective concentration of a molecule inside is high, about a few μM , and CueR dissociation into monomers is negligible. Fig. S7B shows a typical E_{FRET} trajectory of holo-CueR_{Cy5-C129}–DNA interactions inside a vesicle. Two states are clear at $E_1 \sim 0.24$ and $E_2 \sim 0.92$, corresponding to the two protein-bound states with opposite orientations (reference Fig. 1A, cartoons to the right). The state where CueR and DNA are not in complex is negligible, as their K_D is in the nanomolar regime (Table 1 and Table S1), much smaller than the effective concentration inside a vesicle. The transitions between the two states come from direct protein flipping or protein unbinding followed by rapid rebinding where the unbound state is not detected (the binding rate here is $\sim 50 \text{ s}^{-1}$ using the binding rate constant ($k_1 \sim 9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$, Table 1) and the effective protein concentration ($\sim 6 \mu\text{M}$). The direct protein substitution and assisted protein dissociation cannot occur here because there is *no* a second protein molecule inside the same vesicle.

The distribution of τ_2 here again follows a double-exponential decay (Fig. S7D, and Fig. S8B for the distribution of τ_1), similar to those where direct DNA immobilization scheme is used (in Fig. 3A). Within experimental error, the two exponential decay constants here are the same as those from the τ_2 distribution of holo-CueR_{Cy5-C129}–DNA interactions using the direct immobilization scheme at the lowest protein concentration (0.5 nM) where both the direct protein substitution and assisted protein dissociation are insignificant (Fig. S8C). Therefore, CueR–DNA interactions follow similar kinetics using either the vesicle trapping scheme or the direct immobilization. This similarity indicates that (1) the monomer-dimer equilibrium, if exists, is insignificant in considering their interaction kinetics, and (2) the surface effect is also negligible in the direct immobilization scheme because the vesicle trapping eliminates the interactions of the protein and DNA with the quartz surface.

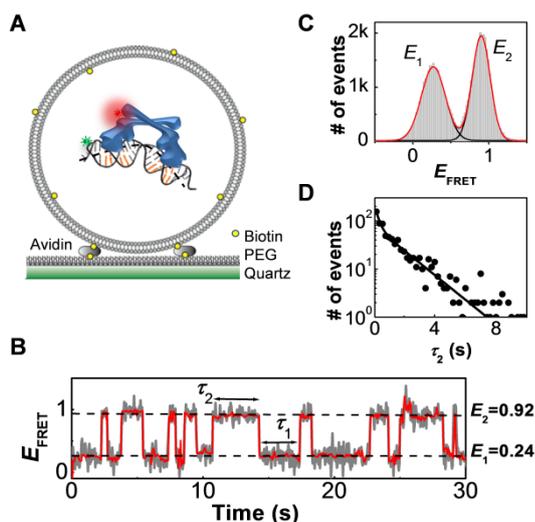


Fig. S7. CueR–DNA interactions measured inside ~ 80 -nm inner diameter lipid vesicles. (A) Schematic of vesicle trapping of a single CueR–DNA pair. (B) E_{FRET} trajectory of a holo-CueR_{Cy5-C129} molecule interacting with a specific DNA inside a vesicle showing fluctuations between two states. (C) Compiled histogram of ~ 70 E_{FRET} trajectories of holo-CueR_{Cy5-C129}–DNA interactions inside vesicles. Solid lines are Gaussian resolved peaks centered at ~ 0.92 and 0.24 . Bin size: 0.02 . (D) Distribution of τ_2 from (B); bin size: 0.15 s. Solid line is a fit with a sum of two exponentials: $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$; $\gamma_1 = 0.60 \pm 0.02$ s⁻¹, $\gamma_2 = 5.0 \pm 0.3$ s⁻¹, and $A = 0.80 \pm 0.01$.

S9. Comparison of the Dwell Time Distributions from the Direct Immobilization Scheme and Those from the Nanovesicle Trapping Scheme

Fig. S8A presents the distributions of τ_2 and τ_1 for holo-CueR_{Cy5-C129}–DNA interactions obtained using the direct immobilizations scheme (Fig. S1A). The two distributions both follow double exponential decay and are identical to each other within experimental error, consistent with that the E_2 and the E_1 state are the protein–DNA complexes with the two opposite protein binding orientations. Fig. S8B presents the similar data, but obtained using the nanovesicle trapping scheme (Fig. S7A); again the two distributions are identical, as expected.

Fig. S8C compares the τ_2 distributions of holo-CueR_{Cy5-C129}–DNA interactions using the direct immobilization scheme at the lowest protein concentration (0.5 nM) and using the nanovesicle trapping scheme. The two distributions are identical within experimental error. Therefore, CueR–DNA interactions follow similar kinetics using either the nanovesicle trapping or the direct immobilization scheme. This similarity also rules out the possibility that CueR monomer-dimer equilibrium plays a significant role in the interaction kinetics, as this equilibrium is insignificant in the nanovesicle trapping scheme, and the possibility that nonspecific interactions may play a role, as any interactions with the surface is removed in the trapping scheme, as discussed in Section S8 earlier.

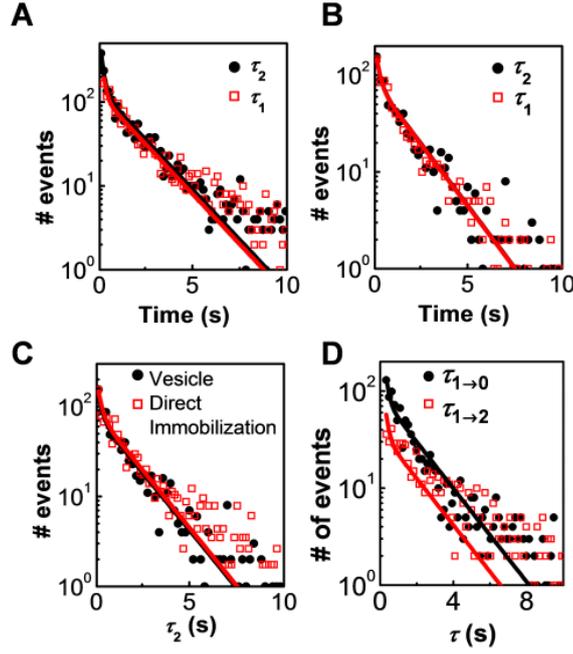


Fig. S8. (A) The distributions of dwell time τ_2 and τ_1 from holo-CueR_{Cy5-C129}-DNA interactions using the direct immobilization scheme as in Fig. S1A at a protein concentration of 2 nM. (B) Same as in (A) but using the nanovesicle trapping scheme. (C) The comparison between the τ_2 distribution of holo-CueR_{Cy5-C129}-DNA interactions in the direct immobilization experiment (protein concentration ~ 0.5 nM) and that in the nanovesicle experiment. The two distributions were normalized to the first data point. (D) Distributions of $\tau_{1 \rightarrow 0}$ and $\tau_{1 \rightarrow 2}$ at [holo-CueR_{Cy5-C129}] = 2 nM using the direct immobilization scheme. Solid lines are global fits with a sum of two exponentials as for the holo data in A. Bin size: 0.15 s. The two distributions are both double-exponentials and have identical exponents, indicating that the transitions leaving the E_1 state start from the same one kinetic species of the two binding modes within the E_1 state. The distributions in panels of A & D as well as those of panel of Fig. 3B (main text) were fitted simultaneously by a sum of two exponentials (solid lines): $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$ with fit parameters: $\gamma_1 = 0.55 \pm 0.01 \text{ s}^{-1}$, $\gamma_2 = 6.4 \pm 0.2 \text{ s}^{-1}$, and $A = 0.68 \pm 0.01$. The distributions in B & C were fitted globally using a sum of two exponentials (solid lines): $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$ with parameters $\gamma_1 = 0.60 \pm 0.02 \text{ s}^{-1}$, $\gamma_2 = 5.0 \pm 0.3 \text{ s}^{-1}$, and $A = 0.80 \pm 0.01$. Data were compiled from 518 E_{FRET} trajectories for A and D, 75 molecules for B, 485 molecules for C. All bin size = 0.15 s.

S10. CueR–DNA Dissociation Constant from Ensemble Fluorescence Anisotropy Titration

To complement our single-molecule measurements, we also used ensemble fluorescence anisotropy titration to determine the CueR–DNA binding affinity. In this titration, the Cy3-labeled DNA was titrated with increasing amounts of holo-CueR_{Cy5-C129} or holo-CueR_{Cy5-E96C}, while the fluorescence anisotropy of the Cy3-DNA was monitored, following procedures by Andoy et al. (2). The titration curve shows a normal saturation behavior (Fig. S9A, B and C).

The fluorescence anisotropy (r_s) titrations were fitted with

$$r_s = r_D + (r_{PD} - r_D) \times \frac{[D]_T + [P]_T + K_D - \sqrt{([D]_T + [P]_T + K_D)^2 - 4[D]_T[P]_T}}{2[D]_T} \quad \text{S[1]}$$

where, r_D and r_{PD} are the anisotropy values for free and protein-bound DNA, respectively, $[D]_T$ is the total DNA concentration, $[P]_T$ is the total protein concentration, and K_D is the dissociation constant of protein-DNA complex. Note that the model behind Equation S[1] is a simple binding equilibrium between CueR and DNA, in which no consideration is given to the existence of multiple, different CueR–DNA complexes; thus the ensemble K_D in Equation S[1] is different from the single-molecule K_D in Table 1, which equals k_{-1}/k_1 , the rate constants defined in Fig. 5. Relation between the ensemble K_D and the single-molecule K_D is derived in Section S17.

Fitting the titration data in Fig. S9A and B give the dissociation constants $K_D \sim 3.5 \pm 0.9$ nM and $\sim 2.6 \pm 0.5$ nM for holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} interactions with the specific DNA, both of which are comparable to the previously reported value of 1.9 ± 0.8 nM obtained on wild-type holo-CueR using the same measurements (2). This agreement demonstrates that the mutation and the Cy5-labeling result in minimal perturbation on CueR's DNA binding affinity and it is also consistent with that CueR_{C129} and CueR_{E96C} are as active as the wild-type CueR in activating transcription in the presence of Cu^+ (Fig. S1B and C).

Fig. S9C shows the anisotropy titration of holo-CueR_{Cy5-C129} in interaction with the Cy3-labeled nonspecific DNA. Fitting the data gives the dissociation constant $K_D \sim 356 \pm 34$ nM, much weaker than that of holo-CueR_{Cy5-C129} interacting with specific DNA, as expected. As the CueR binding to nonspecific DNA follows a simple binding-unbinding model without the presence of any intermediate, the dissociation constant could also be determined from the inverse of waiting times $\langle \tau_{\text{on}} \rangle^{-1}$ (Fig. 4B) and $\langle \tau_{\text{off}} \rangle^{-1}$ (Fig. S9D), which equal k_{-1} and $k_1[\text{P}]$, respectively. The values of rate constants k_{-1} and k_1 are 5.9 ± 0.1 s⁻¹ and 0.016 ± 0.001 nM⁻¹s⁻¹, giving the dissociation constant K_D of 369 ± 24 nM, very similar to that (356 ± 34 nM) determined from the ensemble fluorescence anisotropy titration measurement in Fig. S9C.

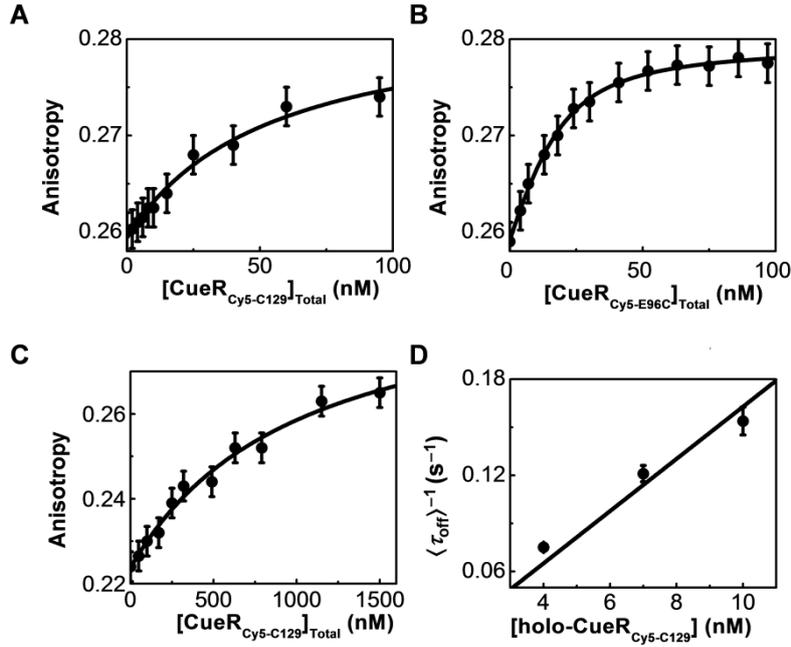


Fig. S9. (A, B) Fluorescence anisotropy titration of Cy3-labelled specific-DNA with holo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}. (C) Same as in A but using Cy3-labeled nonspecific-DNA. The experiment was performed at 50 nM Cy3-DNA in 50 mM pH 7.35 Tris buffer with 10 mM NaCl and 2 mM MgCl₂. The solid line is a fit using Equation S[1], with $K_D \sim 3.5 \pm 0.9$ nM for holo-CueR_{Cy5-C129}, 2.6 ± 0.5 nM for holo-CueR_{Cy5-E96C} in interactions with the specific DNA, and $K_D \sim 356 \pm 34$ nM for holo-CueR_{Cy5-C129} in interactions with the nonspecific DNA. Note that the x -axes of A, B, and C are the total concentration of the homodimeric CueR in solution, not the concentration of free CueR. (D) [holo-CueR_{Cy5-C129}] dependence of $\langle \tau_{\text{off}} \rangle^{-1}$ for CueR_{Cy5-C129} interactions with the nonspecific DNA. Each data point is an average of the dwell times from ~ 250 E_{FRET} trajectories. The solid line is linear fit whose slope gives k_1 , the binding rate constant for CueR binding to nonspecific DNA.

Table S1 below summarizes the K_D 's measured through different methods for wild-type and variant CueR interacting with either specific and nonspecific DNA. Many of them are mentioned in the text and later in this SI.

Table S1. Dissociation constants for CueR–DNA interactions measured at various conditions and using different methods.

| | Experiments | Apo-CueR (nM) | Holo-CueR (nM) | Reference | |
|------------------|-------------------------------|---------------|-----------------|--------------------------|-----------------|
| Specific DNA | In vitro transcription assay | | 5.6 ± 1.1^a | This study (Fig. S1C) | |
| | | | | | 26 ± 9^b |
| | | | | | 1.0 ± 0.5^c |
| | Single molecule imaging | 39 ± 16^d | 52 ± 36^d | This study (Eq. S[42]) | |
| | Ensemble anisotropy titration | | 3.5 ± 0.9^d | This study (Fig. S9A) | |
| Non-Specific DNA | Ensemble anisotropy titration | | 356 ± 34^d | This study (Fig. S9C) | |
| | Single molecule imaging | | 369 ± 24^d | This study (Section S10) | |

^a wt-CueR; ^b CueR_{C129}; ^c CueR_{E96C}; ^d CueR_{Cy5-C129}; ^e CueR_{Cy5-E96C}

S11. Derivation of the Probability Density Function $f(\tau)$ and $\langle \tau \rangle^{-1}$ for the Dwell Time τ_2 and τ_1

Fig. S10 below is part of the kinetic mechanism presented in Fig. 5 in the main text. It includes all the kinetic processes that occur within the E_2 state or depart from the E_2 state. These processes determine the length of τ_2 , the dwell time on the E_2 state.

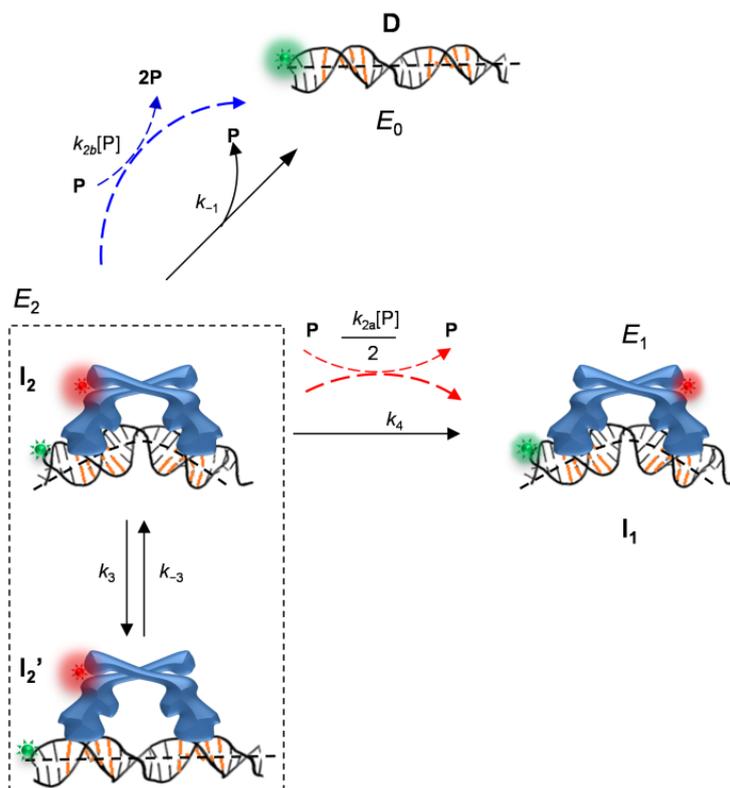


Fig. S10. The kinetic processes associated with the dwell time τ_2 on the FRET state E_2 . k 's are corresponding kinetic rate constants. The two binding modes of CueR in complex with the DNA in the E_2 state are designated as \mathbf{I}_2 and \mathbf{I}_2' . The free DNA is designated as \mathbf{D} . \mathbf{I}_1 is the equivalent to \mathbf{I}_2 , except that the protein binds in an opposite orientation; and it is kinetically connected from the E_2 state by direct flipping (k_4) and direct substitution (k_{2a}). $[\mathbf{P}]$ stands for the protein concentration. Note there is a factor of $1/2$ in front of $k_{2a}[\mathbf{P}]$ for the direct substitution because the incoming protein that replaces the incumbent protein can land in two different orientations and only one of them has the other orientation leading to a transition to the E_1 state. k_{2b} is the rate constant for the protein assisted dissociation.

The single-molecule rate equations from the kinetic scheme in Fig. S10 are:

$$\frac{dP_{\mathbf{I}_2'}(\tau)}{d\tau} = k_3 P_{\mathbf{I}_2}(\tau) - k_{-3} P_{\mathbf{I}_2'}(\tau) \quad \text{S[2]}$$

$$\frac{dP_{\mathbf{I}_2}(\tau)}{d\tau} = -\left(k_3 + k_4 + k_{-1} + \frac{k_{2a}[\mathbf{P}]}{2} + k_{2b}[\mathbf{P}]\right) P_{\mathbf{I}_2}(\tau) + k_{-3} P_{\mathbf{I}_2'}(\tau) \quad \text{S[3]}$$

$$\frac{dP_{\mathbf{I}_1}(\tau)}{d\tau} = \left(k_4 + \frac{k_{2a}[\mathbf{P}]}{2}\right) P_{\mathbf{I}_2}(\tau) \quad \text{S[4]}$$

$$\frac{dP_{\mathbf{D}}(\tau)}{d\tau} = (k_{-1} + k_{2b}[\mathbf{P}]) P_{\mathbf{I}_2}(\tau) \quad \text{S[5]}$$

where $P(\tau)$'s are the probabilities of finding CueR and DNA either in bound (\mathbf{I}_2 , \mathbf{I}_2' , \mathbf{I}_1) or free (\mathbf{D}) states at time τ , k 's are the rate constants for the transitions denoted in Fig. S10, and $[\mathbf{P}]$ is the protein concentration. The initial conditions for solving the above differential equations are: $P_{\mathbf{I}_2}(0) = 1$, $P_{\mathbf{I}_2'}(0) = 0$, $P_{\mathbf{I}_1}(0) = 0$, and $P_{\mathbf{D}}(0) = 0$ (at $\tau = 0$, CueR and DNA are in the form of \mathbf{I}_2 because this is the first species to be formed after protein binding, flipping or direct substitution). And at any time, $P_{\mathbf{I}_2}(\tau) + P_{\mathbf{I}_2'}(\tau) + P_{\mathbf{I}_1}(\tau) + P_{\mathbf{D}}(\tau) = 1$. We can then evaluate the probability density function of τ_2 , $f_2(\tau)$. The probability of finding a particular $\tau_2 = \tau$ is $f_2(\tau)\Delta\tau$, which is equal to the probability for the species \mathbf{I}_2 to switch to \mathbf{I}_1 or \mathbf{D} between τ and $\tau + \Delta\tau$, i.e., $\Delta P_{\mathbf{I}_1}(\tau) + \Delta P_{\mathbf{D}}(\tau)$ (2, 17, 18). Therefore, $f_2(\tau)\Delta\tau = \Delta P_{\mathbf{I}_1}(\tau) + \Delta P_{\mathbf{D}}(\tau)$. In the limit of infinitesimal $\Delta\tau$, $f_2(\tau)$ is equal to $\frac{d}{d\tau}(P_{\mathbf{I}_1}(\tau) + P_{\mathbf{D}}(\tau))$. Solving Equations S[2]-[5] for $P_{\mathbf{I}_2}(\tau)$, $P_{\mathbf{I}_2'}(\tau)$, $P_{\mathbf{I}_1}(\tau)$ and $P_{\mathbf{D}}(\tau)$ using the initial conditions, the probability density function $f_2(\tau)$ is:

$$f_2(\tau) = \frac{D}{4M} \left[(M+C) \exp\left\{-\frac{(N-2M)\tau}{4}\right\} + (M-C) \exp\left\{-\frac{N\tau}{4}\right\} \right] \quad \text{S[6]}$$

Here

$$M = \sqrt{\left[2k_{-3} + 2k_{-1} + 2k_3 + 2k_4 + (k_{2a} + 2k_{2b})[\mathbf{P}]\right]^2 - 8k_{-3} \left[2k_{-1} + 2k_4 + (k_{2a} + 2k_{2b})[\mathbf{P}]\right]},$$

$$N = M + 2(k_{-3} + k_{-1} + k_3 + k_4) + (k_{2a} + 2k_{2b})[\mathbf{P}],$$

$$C = \left[2(k_{-3} - k_{-1} - k_3 - k_4) - (k_{2a} + 2k_{2b})[\mathbf{P}]\right], \quad \text{and} \quad D = \left[2(k_{-1} + k_4) + (k_{2a} + 2k_{2b})[\mathbf{P}]\right]. \quad \text{And}$$

$$\int_0^\infty f_2(\tau) d\tau = 1.$$

Note depending on whether holo-CueR or apo-CueR is concerned, either $k_{2b} = 0$ (for holo) or $k_{2a} = 0$ (for apo).

Equation S[6] is given as Equation 2 in the main text. Equation S[6] predicts that the distribution of τ_2 follows a double-exponential decay function, consistent with the experimental observation (Fig. 3A). Furthermore, τ_2 can be divided into two sub-types: $\tau_{2 \rightarrow 0}$, the dwell time that precedes each $E_2 \rightarrow E_0$ transition, and $\tau_{2 \rightarrow 1}$, the dwell time that precedes each $E_2 \rightarrow E_1$ transition. Similarly, the probability density functions for $\tau_{2 \rightarrow 0}$ and $\tau_{2 \rightarrow 1}$ are:

$$f_{2 \rightarrow 0}(\tau) = \frac{dP_D(\tau)}{d\tau} = \frac{k_{-1}}{2M} \left[(M+C) \exp\left\{-\frac{(N-2M)\tau}{4}\right\} + (M-C) \exp\left\{-\frac{N\tau}{4}\right\} \right] \quad \text{S[7]}$$

$$f_{2 \rightarrow 1}(\tau) = \frac{dP_{I_1}(\tau)}{d\tau} = \frac{2k_4 + (k_{2a} + k_{2b})[\mathbf{P}]}{4M} \left[(M+C) \exp\left\{-\frac{(N-2M)\tau}{4}\right\} + (M-C) \exp\left\{-\frac{N\tau}{4}\right\} \right] \quad \text{S[8]}$$

And $f_{2 \rightarrow 0}(\tau) + f_{2 \rightarrow 1}(\tau) = f_2(\tau)$. Equations S[7] and S[8] predict that the distributions of $\tau_{2 \rightarrow 0}$ and $\tau_{2 \rightarrow 1}$ should both follow double-exponential decay functions with the same exponents, consistent with the experimental data (Fig. 3B and Fig. S4E).

We can then obtain $N_{2 \rightarrow 1}/N_{2 \rightarrow 0}$, the ratio between the numbers of $E_2 \rightarrow E_1$ and $E_2 \rightarrow E_0$ transitions, i.e., their relative probabilities, in the E_{FRET} trajectories

$$\frac{N_{2 \rightarrow 1}}{N_{2 \rightarrow 0}} = \frac{\int_0^{\infty} f_{2 \rightarrow 1}(\tau) d\tau}{\int_0^{\infty} f_{2 \rightarrow 0}(\tau) d\tau} = \frac{k_4 + \frac{1}{2}k_{2a}[\mathbf{P}]}{k_{-1} + k_{2b}[\mathbf{P}]} \quad \text{S[9]}$$

When $k_{2a} \gg k_{2b} = 0$, i.e., the direct protein substitution dominates over assisted protein dissociation,

$$\frac{N_{2 \rightarrow 1}}{N_{2 \rightarrow 0}} = \frac{k_4 + \frac{1}{2}k_{2a}[\mathbf{P}]}{k_{-1}} \quad \text{S[10]}$$

Equation [10] predicts that $N_{2 \rightarrow 1}/N_{2 \rightarrow 0}$ will increase linearly with increasing $[\mathbf{P}]$, consistent with what was observed for holo-CueR_{Cy5-C129}-DNA interactions (Fig. 4C). When $k_{2b} \gg k_{2a} = 0$, i.e., the assisted protein dissociation dominates over direct protein substitution:

$$\frac{N_{2 \rightarrow 1}}{N_{2 \rightarrow 0}} = \frac{k_4}{k_{-1} + k_{2b}[\mathbf{P}]} \quad \text{S[11]}$$

Equation [11] predicts that $N_{2 \rightarrow 1}/N_{2 \rightarrow 0}$ will decrease with increasing $[\mathbf{P}]$, while its inverse $N_{2 \rightarrow 0}/N_{2 \rightarrow 1}$ will increase linearly with increasing $[\mathbf{P}]$, consistent with what was observed for apo-CueR_{Cy5-C129}-DNA interactions (Fig. 4D). Equations S[10] and [11] are given as Equation 3 and 4 in the main text.

As the E_1 state is chemically equivalent to the E_2 state, the probability density functions of τ_1 are the same as those of τ_2 :

$$f_1(\tau) = f_2(\tau) \quad \text{S[12]}$$

$$f_{1 \rightarrow 2}(\tau) = f_{2 \rightarrow 1}(\tau) \quad \text{S[13]}$$

$$f_{1 \rightarrow 0}(\tau) = f_{2 \rightarrow 0}(\tau) \quad \text{S[14]}$$

To use the above equations to extract kinetic parameters from our experimental results, we did the following. The data such as those Fig. 3A were fitted empirically with a sum of two exponentials:

$$y_2(\tau) = N_2[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)] \quad \text{S[15]}$$

where N_2 is a normalization factor. Correlating Equations S[15] and S[6], we have

$$\frac{N}{4} = \gamma_2 \quad \text{S[16]}$$

$$\frac{N-2M}{4} = \gamma_1 \quad \text{S[17]}$$

$$\frac{(M+C)D}{4M} = A\gamma_1 \quad \text{S[18]}$$

$$\frac{(M-C)D}{4M} = (1-A)\gamma_2 \quad \text{S[19]}$$

Solving Equations S[16] to S[19]:

$$C = \frac{2(\gamma_2 - \gamma_1)[A\gamma_1 - (1-A)\gamma_2]}{[A\gamma_1 + (1-A)\gamma_2]} \quad \text{S[20]}$$

$$D = 2[A\gamma_1 + (1-A)\gamma_2] \quad \text{S[21]}$$

$$M = 2(\gamma_2 - \gamma_1) \quad \text{S[22]}$$

$$N = 4\gamma_2 \quad \text{S[23]}$$

Also,

$$C+D=2(k_{-3}-k_3) \quad \text{S[24]}$$

$$N-M+C=4k_{-3} \quad \text{S[25]}$$

Substituting Equations S[20]-[23] to S[24]-[25], we get,

$$k_3 = \frac{(\gamma_1 - \gamma_2)}{2} \left((1-2A) + \frac{[A\gamma_1 - (1-A)\gamma_2]}{[A\gamma_1 + (1-A)\gamma_2]} \right) \quad \text{S[26]}$$

$$k_{-3} = \frac{(\gamma_2 + \gamma_1)}{2} + \left(\frac{(\gamma_2 - \gamma_1)[A\gamma_1 - (1-A)\gamma_2]}{2[A\gamma_1 + (1-A)\gamma_2]} \right) \quad \text{S[27]}$$

The parameters A , γ_1 , and γ_2 were determined from fitting the experimental distribution of dwell time at all the protein concentrations, at each of which the rate constants, k_3 and k_{-3} , were determined. k_3 and k_{-3} are plotted against CueR concentration in Fig. S13 for holo and apo CueR. The k_3 and k_{-3} are independent of CueR concentration, as expected. So, the final values of k_3 and k_{-3} for apo and holo were determined by averaging the data across the different protein concentration in Fig. S13 (Table 1). K_{3D} was then calculated as k_3/k_{-3} (Table 1).

Additionally, the inverse of the average dwell time is,

$$\langle \tau_2 \rangle^{-1} = \langle \tau_1 \rangle^{-1} = 1 \int_0^{\infty} \tau f_2(\tau) d\tau = \frac{k_{-1} + k_4}{1 + K_{3D}} + \frac{k_{2a} + 2k_{2b}}{2(1 + K_{3D})} [\mathbf{P}] \quad \text{S[28]}$$

We can consider two limiting conditions: (1) When the direct substitution pathway dominates over the assisted dissociation, i.e., $k_{2a} \gg k_{2b} = 0$. Then, equation S[28] reduces to:

$$\langle \tau_2 \rangle^{-1} = \frac{k_{-1} + k_4}{1 + K_{3D}} + \frac{k_{2a}}{2(1 + K_{3D})} [\mathbf{P}] \quad \text{S[29]}$$

which predicts that $\langle \tau_2 \rangle^{-1}$ increases linearly with increasing $[\mathbf{P}]$, as observed for holo-CueR–DNA interactions (Fig. 4A). (2) When the assisted dissociation dominates, i.e., $k_{2b} \gg k_{2a} = 0$ Then the equation S[28] reduces to:

$$\langle \tau_2 \rangle^{-1} = \frac{k_{-1} + k_4}{1 + K_{3D}} + \frac{k_{2b}}{1 + K_{3D}} [\mathbf{P}] \quad \text{S[30]}$$

which also predicts that $\langle \tau_2 \rangle^{-1}$ increases linearly with increasing $[\mathbf{P}]$, as observed for apo-CueR–DNA interactions (Fig. 4A).

Equations S[29] and S[30] are given as Equations 6 and 7 in the main text. They were used to fit the data in Fig. 4A to determine the kinetic parameters k_4 and k_{2a} (for holo) and k_{-1} and k_{2b} (for apo) (See Section S16 for more details for obtaining all kinetic parameters).

S12. Derivation of Probability Density Function $f_0(\tau)$ and $\langle \tau_0 \rangle^{-1}$ for the Dwell Time τ_0

Fig. S11 shows the kinetic processes that are associated with the dwell time τ_0 . These are the ones that depart from the E_0 state:

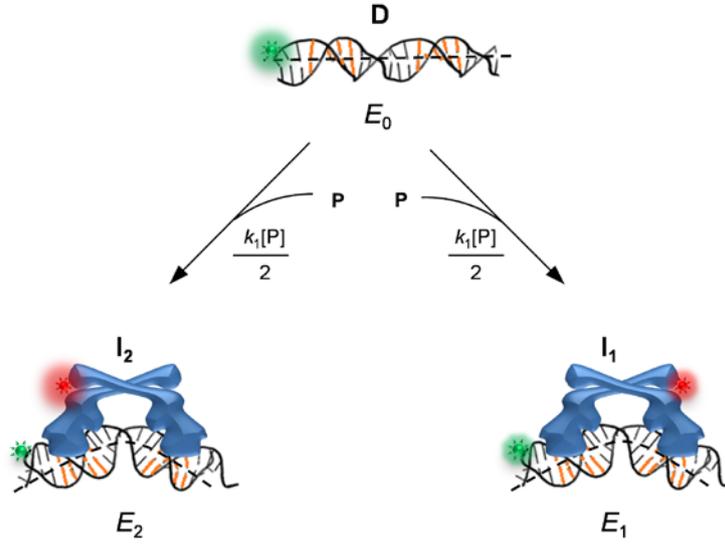


Fig. S11. The kinetic processes related to the dwell time τ_0 on the FRET state E_0 . Note there is a $\frac{1}{2}$ factor in front of $k_1[\mathbf{P}]$ because for each protein binding, only 50% of time it will bind in one of the two orientations.

The single-molecule rate equations from the above kinetic scheme are:

$$\frac{d P_{I_2}(\tau)}{d\tau} = \frac{k_1[\mathbf{P}]}{2} P_D(\tau) \quad \text{S[31]}$$

$$\frac{d P_{I_1}(\tau)}{d\tau} = \frac{k_1[\mathbf{P}]}{2} P_D(\tau) \quad \text{S[32]}$$

$$\frac{d P_D(\tau)}{d\tau} = -k_1[\mathbf{P}] P_D(\tau) \quad \text{S[33]}$$

where $P(\tau)$'s are the probabilities of finding CueR and DNA either in bound (I_2 and I_1) or unbound (D) states at time τ , k_1 is the binding rate constant, and $[\mathbf{P}]$ is the protein concentration. The initial conditions for solving the above differential equations are: $P_{I_2}(0) = 0$, $P_{I_1}(0) = 0$, $P_D(0) = 1$, with $\tau = 0$ being CueR and DNA in the free state D . And at any time $P_{I_2}(\tau) + P_{I_1}(\tau) + P_D(\tau) = 1$. We can then evaluate the probability density function of τ_0 , $f_0(\tau)$. The probability of finding a particular τ_0 is $f_0(\tau)\Delta\tau$, which is equal to the

probability for the free DNA (**D**) switching to **I**₂ or **I**₁ between τ and $\tau + \Delta\tau$, i.e. $\Delta P_{I_2}(\tau) + \Delta P_{I_1}(\tau)$. Therefore, $f_0(\tau)\Delta\tau = \Delta P_{I_2}(\tau) + \Delta P_{I_1}(\tau)$. In the limit of infinitesimal $\Delta\tau$, $f_0(\tau)$ is equal to $\frac{d}{d\tau}(P_{I_2}(\tau) + P_{I_1}(\tau))$. Solving equations S[31]-[33] for $P_{I_2}(\tau)$, $P_{I_1}(\tau)$ and $P_D(\tau)$ by using above initial conditions, the probability density function of $f_0(\tau)$ is:

$$f_0(\tau) = k_1[\mathbf{P}] \exp(-k_1[\mathbf{P}]\tau) \quad \text{S[34]}$$

Equation S[34] predicts the distribution $f_0(\tau)$ to be a single exponential, consistent with the experimental observation (Fig. 2B). Equation S[34] is presented as Equation (1) in the main text. Then, the inverse of average dwell time,

$$\langle \tau_0 \rangle^{-1} = 1 / \int_0^{\infty} \tau f_0(\tau) d\tau = k_1[\mathbf{P}] \quad \text{S[35]}$$

Equation S[35] predicts a linear relationship between $\langle \tau_0 \rangle^{-1}$ and $[\mathbf{P}]$, as observed experimentally (Fig. 2A). Equation S[35] is presented as Equation (5) in the main text. Note that the experimental $\langle \tau_0 \rangle^{-1}$ do not go through zero exactly when extrapolated to $[\mathbf{P}] = 0$ (Fig. 2A); this is likely because that the longest $\langle \tau_0 \rangle$ (which corresponds to the smallest $\langle \tau_0 \rangle^{-1}$) we could ever measure is the average observation time (typically tens of seconds to about 2 min) before the FRET probes bleach, which results in a finite $\langle \tau_0 \rangle^{-1}$ even at the lowest $[\mathbf{P}]$. To partially address this issue, a (0, 0) point was added in both the holo and apo data of $\langle \tau_0 \rangle^{-1}$ versus $[\mathbf{P}]$ results, and the data were fitted with a straight line with a non-zero y-intercept (Fig. 2A); the slope obtained from the fitting was used for the value of k_1 as in Equation S[35].

S13. Simplified Kinetic Mechanism of CueR–DNA Interaction

The chemical species involved in the kinetic mechanism in Fig. 5 take into account the fluorescent labels on CueR and DNA, which differentiate the two orientations of CueR upon binding DNA. Removing the fluorescent labels, the two orientations become identical, (i.e., $\mathbf{I}_2 = \mathbf{I}_1 \equiv \mathbf{I}$ and $\mathbf{I}_2' = \mathbf{I}_1' \equiv \mathbf{I}'$), the E_2 and E_1 states merge, and the kinetic mechanism in Fig. 5 simplifies to that in Fig. S12 below. This simplified mechanism only involves three different chemical species: the free DNA (**D**), the CueR–DNA complex in which the CueR recognizes the targeting sequence and distorts the DNA structure (**I**), and the complex in which the CueR–DNA interaction mimics the case for interacting with a nonspecific DNA (**I'**). Moreover, because the two protein binding orientations are now un-differentiated, the $\frac{1}{2}$ factor in front of $k_1[\mathbf{P}]$ and $k_{2a}[\mathbf{P}]$ in Fig. 5 vanishes in Fig. S12.

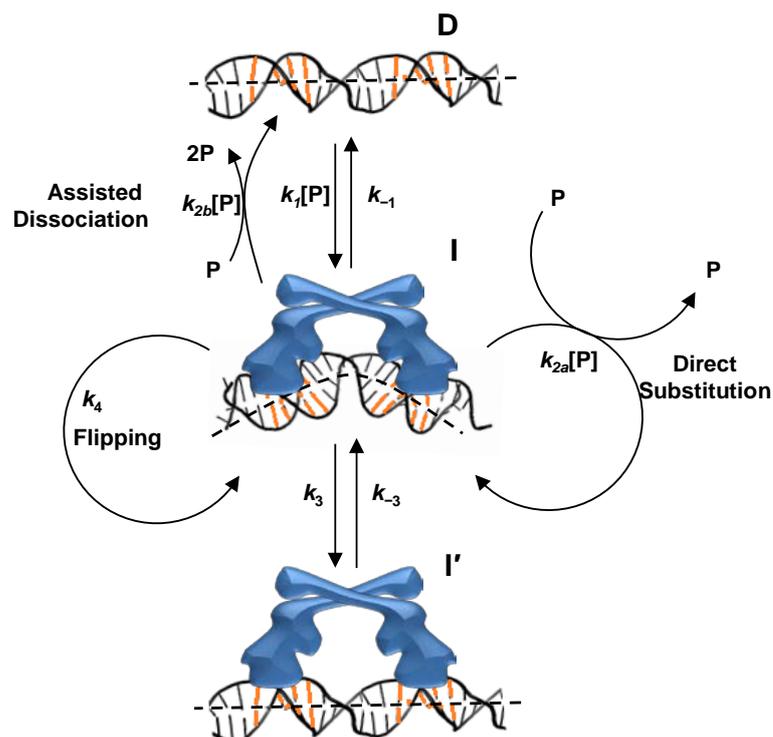


Fig. S12. Simplified kinetic mechanism of CueR interacting with a specific DNA, where the fluorescent probes on CueR and DNA are removed. As such, the two protein binding orientations are identical (i.e., $\mathbf{I}_2 = \mathbf{I}_1 \equiv \mathbf{I}$, $\mathbf{I}_2' = \mathbf{I}_1' \equiv \mathbf{I}'$; reference Fig. 5 in the main text).

S14. Inter-conversion Rate Constants (k_3 and k_{-3}) of Apo and Holo-CueR–DNA Interaction

Fig. S13 plots the rate constants k_3 and k_{-3} determined at various protein concentrations for holo-CueR and apo-CueR interactions with DNA, using Equations S[26] and S[27]. Both k_3 and k_{-3} are independent of CueR concentration, as expected. The variations of their values at different protein concentrations reflect experimental uncertainties. Therefore, their final values were taken as the averages of those determined at various protein concentrations and their final values were also used to calculate K_{3D} ($= k_3/k_{-3}$, Table 1).

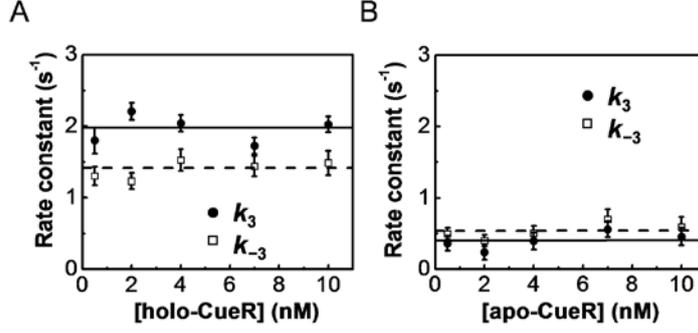


Fig. S13. The numerical values of k_3 and k_{-3} determined at various protein concentrations for (A) holo-CueR and (B) apo-CueR interactions with DNA. The solid lines are fits with horizontal lines. The error bars are s.d.

S15. The Equilibrium Ratio between Protein-bound Population and Unbound Population

The equilibrium ratio between the protein-bound DNA population and the unbound free DNA population can be determined from the area ratio of the resolved peaks in the E_{FRET} histogram (e.g., Fig. 1F):

$$\frac{\text{Combined area of } E_1 \text{ and } E_2 \text{ peaks}}{\text{area of } E_0 \text{ peak}} = \frac{\text{Population of all complexes}}{\text{Population of free DNA}} = \frac{N_I + N_I}{N_D} \quad \text{S[36]}$$

The population of all complexes includes the population of species **I** (N_I) and that of species **I'** ($N_{I'}$) (reference Fig. S12). The population of free DNA denoted as N_D . With the kinetic mechanism in Fig. S12, we can write:

$$\frac{N_I}{N_D} = \frac{k_1[\mathbf{P}]}{k_{-1} + k_{2b}[\mathbf{P}]} \quad \text{S[37]}$$

$$\frac{N_{I'}}{N_I} = \frac{k_3}{k_{-3}} \equiv K_{3D} \quad \text{S[38]}$$

Combining Equations S[36] to S[38], we have:

$$\frac{(E_1 + E_2) \text{ peak area}}{E_0 \text{ peak area}} = (1 + K_{3D}) \frac{k_1[\mathbf{P}]}{k_{-1} + k_{2b}[\mathbf{P}]} \quad \text{S[39]}$$

This equation is linear in $[\mathbf{P}]$ when $[\mathbf{P}]$ is very small and will saturate at higher $[\mathbf{P}]$ (Fig. 1H, apo). For holo-CueR–DNA interaction, the assisted protein dissociation is negligible (i.e., $k_{2b} = 0$), and Equation S[39] reduces to:

$$\frac{(E_1 + E_2) \text{ peak area}}{E_0 \text{ peak area}} = (1 + K_{3D}) \frac{k_1[\mathbf{P}]}{k_{-1}} \quad \text{S[40]}$$

This equation is linear in $[\mathbf{P}]$, as observed experimentally (Fig. 1H, holo). Equations S[39] and S[40] are presented as Equations 9 and 8, respectively, in the main text. Note that the experimental $(E_1 + E_2)/E_0$ area ratio for holo-CueR in Fig. 1H does not go through zero exactly when extrapolated to $[\mathbf{P}] = 0$. This is likely because we only analyzed E_{FRET} trajectories that had at least 2 transitions from the E_0 to E_1 or E_2 states, as the trajectories without transitions are difficult to differentiate from nonspecifically adsorbed dye molecules. This would result in a finite $(E_1 + E_2)/E_0$ area ratio even at lowest possible $[\mathbf{P}]$. To partially address this issue, we added a y -offset in fitting Fig. 1H with a linear function.

S16. Additional Details about Procedure of Extracting Kinetic Parameters

For holo-CueR: the protein binding rate constant k_1 was determined by fitting the data in Fig. 2A with Equation 5 (i.e., Equation S[35]; see more details in Section S12). Then, the inter-conversion rate constants k_3 and k_{-3} , and thus K_{3D} , were determined as described in Section S11 (and Fig. S13). This then allowed for the determination of k_{-1} by fitting the data in Fig. 1H to Equation 8 (i.e., Equation S[40]). Finally, k_{2a} and k_4 rates were determined by fitting the data in Fig. 4A to Equation 6 (i.e. Equation S[29]).

For apo-CueR: the values of k_1 , k_3 , k_{-3} , and K_{3D} were determined with the same procedure as that used for holo-CueR. Then, the ratio between Equations 9 (i.e., Equation S[39]) and Equation 4 (i.e., Equation S[11]) was calculated to determine the value of k_4 , which is:

$$\frac{(E_1 + E_2) \text{ peak area} / E_0 \text{ peak area}}{N_{2 \rightarrow 1} / N_{2 \rightarrow 0}} = \frac{(1 + K_{3D})k_1[\mathbf{P}]}{k_4} \quad \text{S[41]}$$

The histogram peak area ratio and the $N_{2 \rightarrow 1} / N_{2 \rightarrow 0}$ are presented in Fig. 1H and Fig. 4D. The ratio of these two data is plotted in Fig. S14. Then, the value of k_4 rate was determined by fitting the data of Fig. S14 to Equation S[41]. Finally, k_{2b} and k_{-1} rates were determined by fitting the data in Fig. 4A to Equation 7 (i.e., Equation S[30]). The determined values of all kinetic parameters are presented in Table 1.

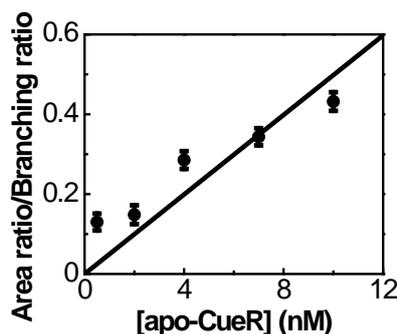


Fig. S14. Protein concentration dependence of ratio of histogram area ratio to $N_{2 \rightarrow 1} / N_{2 \rightarrow 0}$ (i.e., “Branching ratio” in the plot) for apo-CueR_{Cy5-C129}-DNA interactions. Solid line is a fit with Equation S[41].

S17. Determination of the Corresponding CueR–DNA Ensemble Dissociation Constants Using Single Molecule Dissociation Constants

From the K_D and K_{3D} determined from the single-molecule measurements (Table 1), we can obtain the corresponding ensemble dissociation constant K_D' , which does not take into account the presence of multiple protein–DNA complexes. Using the kinetic scheme in Fig. 5, we have:

$$\frac{1}{K_D'} = \frac{[\mathbf{I}_2] + [\mathbf{I}_2'] + [\mathbf{I}_1] + [\mathbf{I}_1']}{[\mathbf{D}][\mathbf{P}]} = \frac{2[\mathbf{I}_2] + 2[\mathbf{I}_2']}{[\mathbf{D}][\mathbf{P}]} = \frac{1}{K_D} + \frac{K_{3D}}{K_D} \quad \text{S[42]}$$

where, $\frac{2[\mathbf{I}_2]}{[\mathbf{D}][\mathbf{P}]} = \frac{1}{K_D}$ and $\frac{2[\mathbf{I}_2']}{[\mathbf{D}][\mathbf{P}]} = \frac{[\mathbf{I}_2']}{[\mathbf{I}_2]} \frac{2[\mathbf{I}_2]}{[\mathbf{D}][\mathbf{P}]} = \frac{K_{3D}}{K_D}$. The calculated ensemble dissociation constant K_D' for holo-CueR_{Cy5-C129}-DNA interaction by using the values of K_D and K_{3D} (Table 1) is 52 ± 36 nM. Similarly, the calculated ensemble dissociation constant K_D' for apo-CueR_{Cy5-C129}-DNA interaction is 4.6 ± 2.9 nM. All these values are comparable to previously reported ensemble dissociation constants (Table S1).

S18. Analysis of in Vitro Transcription Run-off Assay Results

To model the in vitro transcriptional assay data quantitatively (Fig. S1B and C), we assumed that the amount of mRNA transcript generated is linearly proportional to the concentration of holo-CueR- P_{copA} complex in the presence of RNAP and that CueR- P_{copA} interactions follow simple binding equilibrium. The measured amount of mRNA transcript as a function of CueR concentration (Fig. S1C) would then follow:

$$\text{Amount of mRNA} = C \left[P_{copA} \right]_0 Y = \frac{C \left[P_{copA} \right]_0 [\text{holo-CueR}]}{K_D + [\text{holo-CueR}]} \quad \text{S[43]}$$

where C is a proportionality constant, $[P_{copA}]_0$ is the total P_{copA} concentration, Y is the fraction of P_{copA} that has a holo-CueR bound, $[\text{holo-CueR}]$ is the free holo-CueR concentration (taken approximately as the total holo-CueR concentration when P_{copA} concentration is small), and K_D is the dissociation constant of the CueR- P_{copA} complex. The Equation S[43] was used to fit the data of Fig. S1C. The dissociation constants obtained from the fits are 5.6 ± 1.1 nM for wild type CueR, 26 ± 9 nM for CueR_{C129} and 1.0 ± 0.5 nM for CueR_{E96C}. Although these values are comparable to reported values (Table S1), note the fitting of the in vitro transcription data here are crude and the values here should not be taken literally.

S19. CueR Behaves Similarly in Its Interactions with A Long, 121 Base-pair, DNA Containing the Complete Promoter

We have used a 25 base-pair DNA that contains the specific dyad symmetric sequence recognized by CueR (Fig. S1). This 25 base-pair sequence covers approximately the entire CueR footprint on DNA (Fig. S15A), determined by Outten et al (1), and provides a sufficiently tight binding by CueR ($K_D \sim$ nM, Table S1), comparable to CueR binding to longer DNA that contains the complete -10 and -35 sequences ($K_D \sim$ nM, from gel-shift assay by Stoyanov et al (16), Table S1).

To probe if a complete promoter region that contains the -10 and -35 sequences would make a difference to CueR-DNA interactions, we further studied holo-CueR_{Cy5-C129} interaction with a 121 base-pair DNA that contains the complete promoter region of the $copA$ gene regulated by CueR (Fig. S15A). This DNA spans from the -60 to $+61$ position around the promoter, including the -10 and -35 sequences and the CueR-specific dyad symmetric sequence. The Cy3 is labeled to an amine modified T at position -41 (Fig. S15A). This 121 base-pair DNA was made via PCR off purified *E. coli* chromosome using Cy3 and biotin-TEG labeled primers.

$c_{129}] = 2 \text{ nM}$. Bin size = 0.002. (D) Distribution of τ_2 , at 2 nM protein concentration, in comparison with that for 25 base-pair DNA from Figure 3A. Data compiled from $\sim 1300 E_{\text{FRET}}$ trajectories. Bin size: 0.15 s. The two distributions were normalized to the first data point for comparison. Solid lines are fits with a sum of two exponentials: $N[A\gamma_1\exp(-\gamma_1\tau) + (1-A)\gamma_2\exp(-\gamma_2\tau)]$. For 25 base-pair: $\gamma_1 = 0.55 \pm 0.01 \text{ s}^{-1}$, $\gamma_2 = 6.4 \pm 0.2 \text{ s}^{-1}$, $A = 0.68 \pm 0.01$; for 121 base-pair: $\gamma_1 = 0.92 \pm 0.09 \text{ s}^{-1}$, $\gamma_2 = 8.9 \pm 0.8 \text{ s}^{-1}$, $A = 0.40 \pm 0.02$. (E) [holo-CueR $_{\text{Cy}5\text{-C}129}$] dependence of $\langle \tau_2 \rangle^{-1}$ in comparison with that for the 25 base-pair DNA from Figure 4A. Solid lines are fits with Equation 6 as in Figure 4A.

S20. CueR Can Still Flip Its Binding Orientation on DNA in the Presence of RNAP

To probe if RNAP would alter the behavior of CueR–DNA interactions, we performed preliminary experiments on holo-CueR $_{\text{Cy}5\text{-C}129}$ interactions with the 121 base-pair DNA in the presence of RNAP. This 121 base-pair DNA contains the -10 and -35 sequences for interaction with RNAP (Fig. S15A). Moreover, on the basis of the RNAP footprints on the MerR and ZntR specific promoters (12, 13) (ZntR is also a MerR-family metalloregulator), the Cy3-labeling at position -41 on the DNA should not interfere significantly with RNAP binding.

The E_{FRET} trajectories show direct $E_1 \leftrightarrow E_2$ transitions (Fig. S16A). Therefore, holo-CueR can still flip its binding orientations on DNA in the presence of RNAP. Moreover, the E_1 and E_2 values in the presence of RNAP are slightly shifted from those in the absence of RNAP (Fig. S16B and C); this shift confirms that RNAP does bind to the CueR–DNA complex using the 121 base-pair DNA with the Cy3 labeled at position -41 .

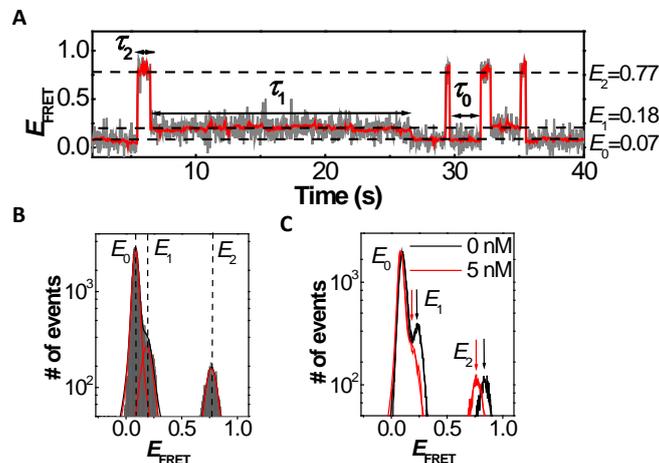


Fig. S16. (A) An E_{FRET} trajectory of an immobilized 121 base-pair Cy3-DNA in interaction with 2 nM holo-CueR $_{\text{Cy}5\text{-C}129}$ in the presence of 5 nM RNAP in solution. (B) Histogram of the E_{FRET} trajectories of holo-CueR $_{\text{Cy}5\text{-C}129}$ interactions as in A. Data from 80 E_{FRET} trajectories were combined here. The solid lines are the fits of the E_{FRET} histogram by Voigt functions centered at ~ 0.07 , 0.18 , and 0.77 , with percentage peak areas of $77.4 \pm 0.5\%$, $12.4 \pm 0.3\%$, and $10.2 \pm 0.4\%$, respectively. Bin size = 0.002. (C) Histogram of E_{FRET} trajectories of holo-CueR $_{\text{Cy}5\text{-C}129}$ –DNA interactions in the absence and presence of 5 nM RNAP. The histogram in the absence of RNAP is from Fig. S15C. The arrows indicate the E_1 and E_2 peaks. Buffer: 50 mM pH 7.35

Tris, 100 mM Potassium Glutamate, 0.1 mg/mL BSA, 2 mM MgCl₂, 1 mM CaCl₂, 2 mM DTT, 0.1 mM EDTA, 5% glycerol and 1 mM Trolox.

S21. Apo-CueR Can Directly Substitute Holo-CueR and Vice Versa on DNA

Our results presented so far were all from experiments using either all holo- or all apo-proteins, from which we could extract quantitative kinetic parameters. To probe if apo-CueR can directly substitute holo-CueR on DNA, we performed another experiment using a mixture of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} in interaction with the 25 base-pair specific DNA. This experiment is based on the assumption that there is no facile Cu⁺ exchange between apo- and holo-CueR within the time frame of our single-molecule imaging experiments (~30 min), so that the separately prepared apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} maintain their respective metallation states after mixing them. This assumption is likely valid considering the following: (1) CueR binds Cu⁺ tightly with an affinity of ~10⁻²¹ M (3). (2) Cu⁺ removal from holo-CueR by CN⁻ chelation takes many hours' incubation (see Section S1A), which suggests that the kinetics of Cu⁺ removal is slow.

Fig. S17A shows an E_{FRET} trajectory using a mixture of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C}. There are direct transitions from the holo-CueR_{Cy5-E96C} bound states ($E_1' \sim 0.42$ and $E_2' \sim 0.62$) to the apo-CueR_{Cy5-C129} bound states ($E_1 \sim 0.25$ and $E_2 \sim 0.92$), and vice versa. These transitions indicate that an apo-protein can directly substitute a DNA-bound holo-protein and vice versa, supporting our hypothesis that the direct protein substitution is a possible pathway for turning off transcription after transcription activation.

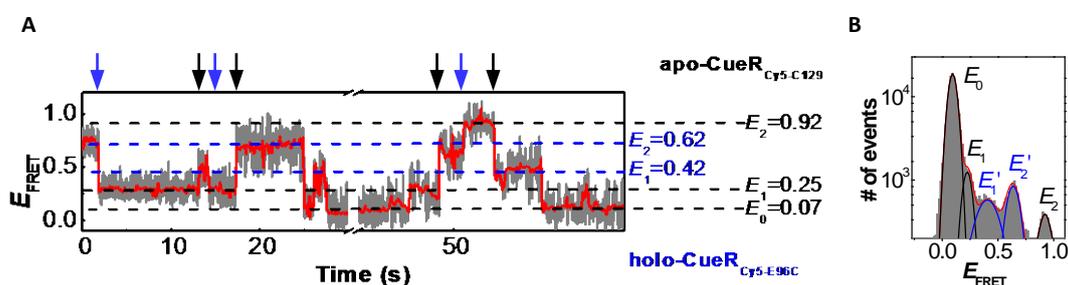


Fig. S17. (A) An E_{FRET} trajectory of an immobilized 25 base-pair Cy3-DNA in interaction with a mixture of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} of about 5 nM each in solution. The blue arrows denote the transitions from the holo-protein bound states to the apo-protein bound states, and the black arrows denote the reverse transitions; these transitions report the direct substitution of a DNA-bound holo-protein by an apo-protein or the reverse. This figure is the same as Fig. 1E in the main text. (B) Histogram of the E_{FRET} trajectories of apo-CueR_{Cy5-C129} and holo-CueR_{Cy5-E96C} interactions with the specific DNA. Data from ~90 E_{FRET} trajectories were combined here. The solid lines are the fit of the E_{FRET} histogram by a sum of five Voigt functions centered at ~0.07 (E_0), 0.25 (E_1), 0.92 (E_2), 0.42 (E_1'), and 0.62 (E_2') with percentage peak areas of $79.0 \pm 0.2\%$, $5.5 \pm 0.8\%$, and $3.8 \pm 0.6\%$, $6.4 \pm 0.5\%$, and $5.3 \pm 0.6\%$, respectively. [apo-CueR_{Cy5-C129}] \approx 5 nM and [holo-CueR_{Cy5-E96C}] \approx 5 nM. Bin size = 0.005.

S22. Supplementary references

1. Outten FW, Outten CE, Hale J, O'Halloran TV (2000) Transcriptional activation of an *Escherichia coli* copper efflux regulation by the chromosomal MerR homologue, CueR. *J. Biol. Chem.* 275:31024-31029.
2. Andoy NM, et al. (2009) Single-molecule study of metalloregulator CueR-DNA interactions using engineered Holliday junctions. *Biophys. J.* 97:844-852.
3. Changela A, et al. (2003) Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301:1383-1387.
4. Brenner AJ, Harris ED (1995) A quantitative test for copper using bicinchoninic acid. *Anal. Biochem.* 226:80-84.
5. Chung SH, Kennedy RA (1991) Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. Neurosci. Methods* 40:71-86.
6. Haran G (2004) Noise reduction in single-molecule fluorescence trajectories of folding proteins. *Chem. Phys.* 307:137-145.
7. Nan X, Sims PA, Chen P, Xie XS (2005) Observation of individual microtubule motor steps in living cells with endocytosed quantum dots. *J. Phys. Chem. B* 109:24220-24224.
8. Benitez JJ, et al. (2008) Probing real-time transient metallochaperone-target protein interactions at the single-molecule level with nanovesicle trapping. *J. Am. Chem. Soc.* 130:2446-2447.
9. Benitez JJ, Keller AM, Chen P (2010) Nanovesicle trapping for studying weak protein interactions by single-molecule FRET. *Meth. Enzymol.* 472:41-60.
10. Benitez JJ, et al. (2011) Relating dynamic protein interactions of metallochaperones with metal transfer at the single-molecule level. *Faraday Discuss.* 148:71-82.
11. Cremer PS, Boxer SG (1999) Formation and spreading of lipid bilayers on planar glass supports. *J. Phys. Chem. B* 103:2554-2559.
12. Outten CE, Outten FW, O'Halloran TV (1999) DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in *Escherichia coli*. *J. Biol. Chem.* 274:37517-37524.
13. O'Halloran TV, Frantz B, Shin MK, Ralston DM, Wright JG (1989) The MerR heavy metal receptor mediates positive activation in a topologically novel transcription complex. *Cell* 56:119-129.
14. Kliegman JI, Griner SL, Helmann JD, Brennan RG, Glasfeld A (2006) Structural basis for the metal-selective activation of the manganese transport regulator of *Bacillus subtilis*. *Biochemistry* 45:3493-3505.
15. Newberry KJ, Brennan RG (2004) The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N-terminus. *J. Biol. Chem.* 279:20356-20362.
16. Stoyanov JV, Hobman JL, Brown NL (2001) CueR (Ybbi) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. *Mol. Microbiol.* 39:502-511.
17. Xie XS (2001) Single-molecule approach to enzymology. *Single Mol.* 2:229-236.
18. Xu W, Kong JS, Chen P (2009) Single-molecule kinetic theory of heterogeneous and enzyme catalysis. *J. Phys. Chem. C* 113:2393-2404.