Supporting Information

Static Kinks or Flexible Hinges: Multiple Conformations of Bent DNA Bound to Integration Host Factor Revealed by Fluorescence Lifetime Measurements

Mitchell Connolly[†], Aline Arra[†], Viktoriya Zvoda[†], Peter J. Steinbach[‡], Phoebe A. Rice[§], and Anjum Ansari[*,†,⊥]

[†]Department of Physics, University of Illinois at Chicago, Chicago, IL 60607, USA

[‡]Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA

[§]Department of Biochemistry & Molecular Biology, University of Chicago, Chicago, IL 60637, USA

[⊥]Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, USA

*Corresponding author

Telephone: (312) 996-8735, Email: ansari@uic.edu

SI Methods

1.1 Determination of active protein concentration

Initial estimates of the IHF protein concentration were obtained from absorption measurements at 276 nm, with extinction coefficient 5800 $M^{-1}cm^{-1}$. More accurate estimates of the active protein concentration were determined on each day of the experiments, by performing equilibrium titration experiments, in which the DNA concentration was held constant at 1 μM while the protein concentration (as estimated by absorption measurements) was varied from 0–3uM, and the acceptor ratio on double-labeled (DNA_DA) samples was measured as a function of [IHF]. The protein concentration at which we observed the maximum acceptor ratio was determined to represent a 1:1 complex, as illustrated in Figure S6; the "active" protein concentration under these conditions was assumed to be 1 μM, the same concentration as the DNA in these measurements. All protein concentrations in the text are rescaled to refer to this active protein concentration.

1.2 Equilibrium acceptor ratio and anisotropy measurements

For acceptor ratio measurements, fluorescence emission spectra for each sample with donor-acceptor-labeled DNA were collected from 500 nm to 700 nm, with excitation of the fluorescein (F) donor at 485 nm. On the same sample, the emission spectra from direct excitation of the acceptor were also collected, from 565 nm to 700 nm for TAMRA (R), with excitation at 555 nm, or from 560 nm to 700 nm for Atto550 (At), with excitation at 555 nm. The acceptor ratio was determined from the measured spectra as follows. A normalized emission spectrum of a donor-only labeled duplex prepared under identical experimental conditions, and excited at 485 nm, was subtracted from the donor-acceptor-labeled emission spectrum, also excited at 485 nm, to isolate the acceptor emission in the double-labeled samples that results from FRET. The area under this corrected acceptor emission spectrum, from 565 nm to 595 nm was divided by the area under the directly excited emission spectrum, also integrated from 565 nm to 595 nm, to obtain the acceptor ratio (denoted as $ratio_A$). For each sample, two independent measurements were performed at each temperature, and errors were estimated as the standard deviations from the average value.

Steady-state anisotropy measurements were performed on donor-acceptor-labeled DNA constructs, in the presence and absence of IHF, at 20 °C, with excitation of the donor (fluorescein) at 485 nm, with fluorescence emission intensities collected at 520 nm. For each set of measurements, the emission intensities for the parallel ($I_{\parallel}$) and perpendicular ($I_{\perp}$) orientations were collected using an integration time of 0.1 s, and the data were processed using a program provided by Horiba to obtain anisotropy values, defined as: $r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$. The errors in the measured anisotropy were calculated as standard deviations from the average of five independent sets of measurements.

1.3 Binding affinity measurements from acceptor ratio titration experiments

Several samples were prepared with a fixed DNA concentration (50 nM) and varying concentrations of IHF, ranging from 1 nM to 100 μM. At each protein concentration, fluorescence emission spectra were collected and $ratio_A$ was calculated as described above. Binding isotherms were obtained by plotting $ratio_A$ versus IHF concentration for each sample, and the corresponding K_d for the IHF-DNA complex was determined by fitting each $ratio_A$ versus [IHF] profile to a 1:1 binding isotherm, as described below.

Assuming only two states for the protein and DNA, free and bound, the measured or acceptor ratio ($ratio_A$) depends on the fraction of DNA in complex (f_x) as:

$$ratio_A = ratio_A^x f_x + ratio_A^{free}(1 - f_x)$$

where $ratio_A^x$ is the acceptor ratio for the complex, $ratio_A^{free}$ is the acceptor ratio for free DNA, and f_x is obtained from:

$$f_x = \frac{1}{2D_0}\left[(P_0 + D_0 + K_d) - \sqrt{(P_0 + D_0 + K_d)^2 - 4(P_0 D_0)}\right]$$

where P_0 and D_0 are the total IHF and DNA concentrations, respectively.

1.4 FRET measurements from MEM lifetime distributions

The average FRET efficiencies were obtained from the lifetime distributions as follows. We defined a normalized amplitude for the *j*-th component of the discretized distribution $f_j(log\tau_j)$ as $A_j = \frac{f_j}{\sum_j f_j}$ and a corresponding lifetime $\tau_j = 10^{\log\tau_j}$. For each distribution, a single characteristic lifetime was computed from the average of the MEM distribution as $< \tau >= \sum_j A_j \tau_j$ and denoted as $< \tau_D >$ or $< \tau_{DA} >$ for donor-only (DNA_D) or donor-acceptor (DNA_DA) labeled samples, respectively. The average FRET efficiency for each sample was computed as $< E >= 1 - \frac{<\tau_{DA}>}{<\tau_D>}$.

The MEM distributions obtained for donor-acceptor labeled samples were further analyzed as follows. For distributions that revealed two distinct peaks, the lifetime distributions were divided into two parts as dictated by the local minimum between the two peaks. Within each peak, the average lifetime, denoted as $< \tau_1 >$ or $< \tau_2 >$ in Table 2, was computed as well as the corresponding average FRET $< E_{1,2} >= 1 - \frac{<\tau_{1,2}>}{<\tau_D>}$. The fractional population in each component (α_1, α_2 in Table 2) was obtained from the sum of the amplitudes A_j from that part of the distribution. The uncertainties reported in Table 2 are standard deviations from the average of two independent sets of measurements.

SI Results

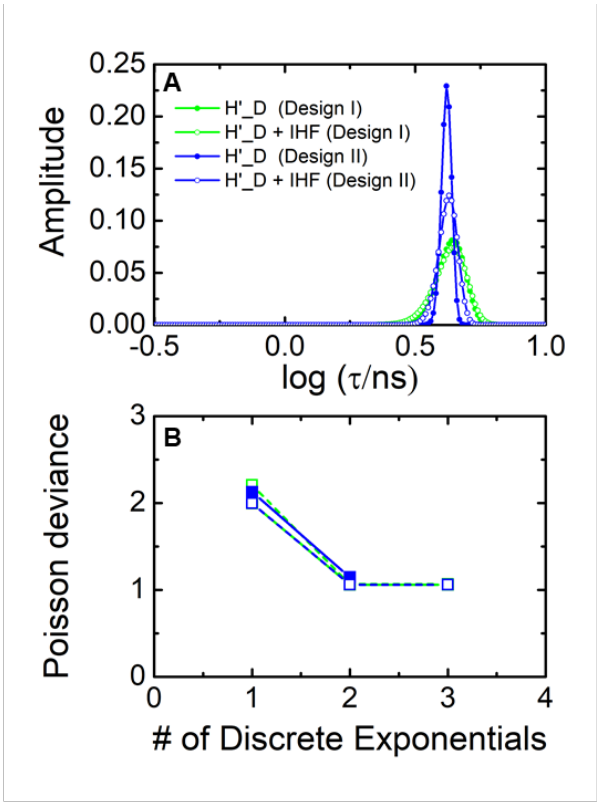
Control experiments to rule out labeling artifacts

We detail here controls that were performed to rule out possible contributions to the two components observed in our lifetime distributions, and in particular, to the low-FRET state, from artifacts such as (i) incompletely labeled (or incompletely annealed) DNA or (ii) partial stacking of the fluorophores at the ends of the DNA that could result in stacked/unstacked populations with different dye orientations.

(i) Incompletely labeled DNA could result in some fraction of donor-only labeled DNA molecules that are missing an acceptor Our measurements of the percent labeling for each of the oligomers indicated >95% for the acceptor-labeled strands, which cannot account for the 20% population we observe in the low-FRET state.

(ii) To examine whether stacking/unstacking of the fluorophores could be contributing to our measurements, we also carried out lifetime measurements on a shorter, 14-bp DNA duplex, with nucleotides adjacent to the dyes identical to that in H' (Figure S3). While the average FRET measured on this construct was larger than in H', as expected, we observed only a single dominant population in the MEM analysis. Indeed, a fit to a single-exponential decay was adequate to describe the decay curves, indicating that different FRET states attributable to different dye orientations are not resolvable, given the noise in these measurements.

These control experiments demonstrate that when two components are observed in the lifetime distributions, they indeed correspond to distinct protein-bound DNA conformations in our ensemble.



C

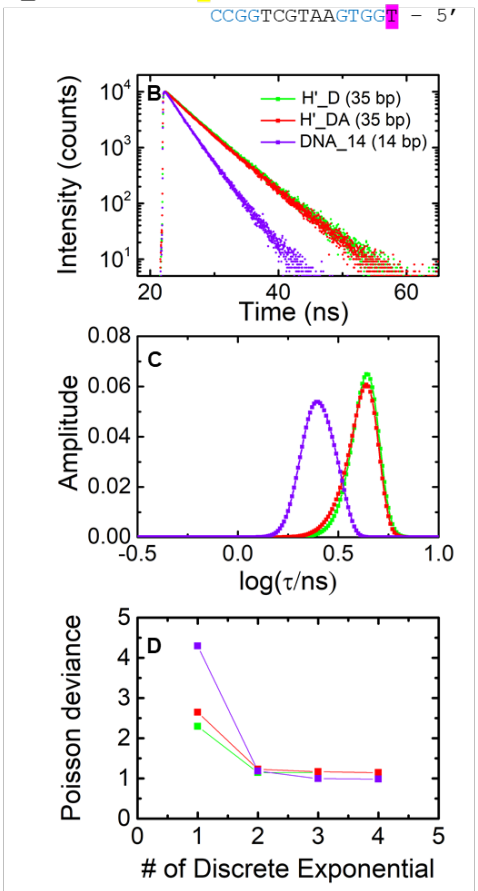
Constructs	$< \tau >$	τ_1	α_1(%)	τ_2	α_2(%)
H'_D (design I)	4.32	4.32	99.0	10.35	1.0
H'_D + IHF (design I)	4.32	4.32	98.8	9.91	1.2
H'_D (design II)	4.17	4.18	99.9	17.95	0.1
H'_D + IHF (design II)	4.22	4.23	99.8	15.35	0.2

Figure S1. MEM and discrete exponential analysis on donor-only labeled H' DNA constructs. (**A**) MEM distributions obtained from fluorescence lifetime decays measured on H'_D in the absence (closed) and presence (open) of IHF are shown for design I (green) and design II (blue) constructs. (**B**) The Poisson deviance obtained from a fit to a sum of discrete exponentials versus the number of exponentials are plotted for each of the decay traces. (**C**) The table summarizes the average lifetimes $< \tau_D >$ from the MEM analysis as well as the lifetimes and amplitudes obtained from a two-exponential fit to the decay trace for each of the samples.

A

H′(design I) 5′ – TGGCCAAAAAAGCATTGCTTATCAATTTGTTGCACC
CCGGTTTTTTCGTAACGAATAGTTAAACAACGTGGT – 5′

DNA_14 5′ – TGGCCAGCATTCACC
CCGGTCGTAAGTGGT – 5′



E

Constructs	⟨τ⟩	τ_1	α_1(%)	τ_2	α_2(%)
H'_D (design I)	4.28	4.27	98.7	9.68	1.3
H'_DA (design I)	4.14	4.12	98.0	8.95	2.0
DNA_14	2.56	2.61	99.8	12.22	0.2

Figure S2. Fluorescence lifetime measurements on 35 bp (H') and 14 bp (DNA_14) DNA constructs in design I. (**A**) The sequences and labeling schemes for the two different DNA constructs are shown. **(B)** Fluorescence intensity decay traces are shown for donor-acceptor-labeled DNA _14_DA (purple), H'_D (green), and H'_DA (red). **(C)** Corresponding MEM distributions are shown for the decay traces of panel B. (**D**) The Poisson deviance obtained from a fit to a sum of discrete exponentials versus the number of exponentials is plotted for each of the decay traces. (**E**) The table summarizes the lifetimes and amplitudes obtained from two-exponential fits to the decay traces, as well as the average lifetimes $< \tau_D >$ and $< \tau_{DA} >$ on donor-only and donor-acceptor-labeled constructs, respectively.

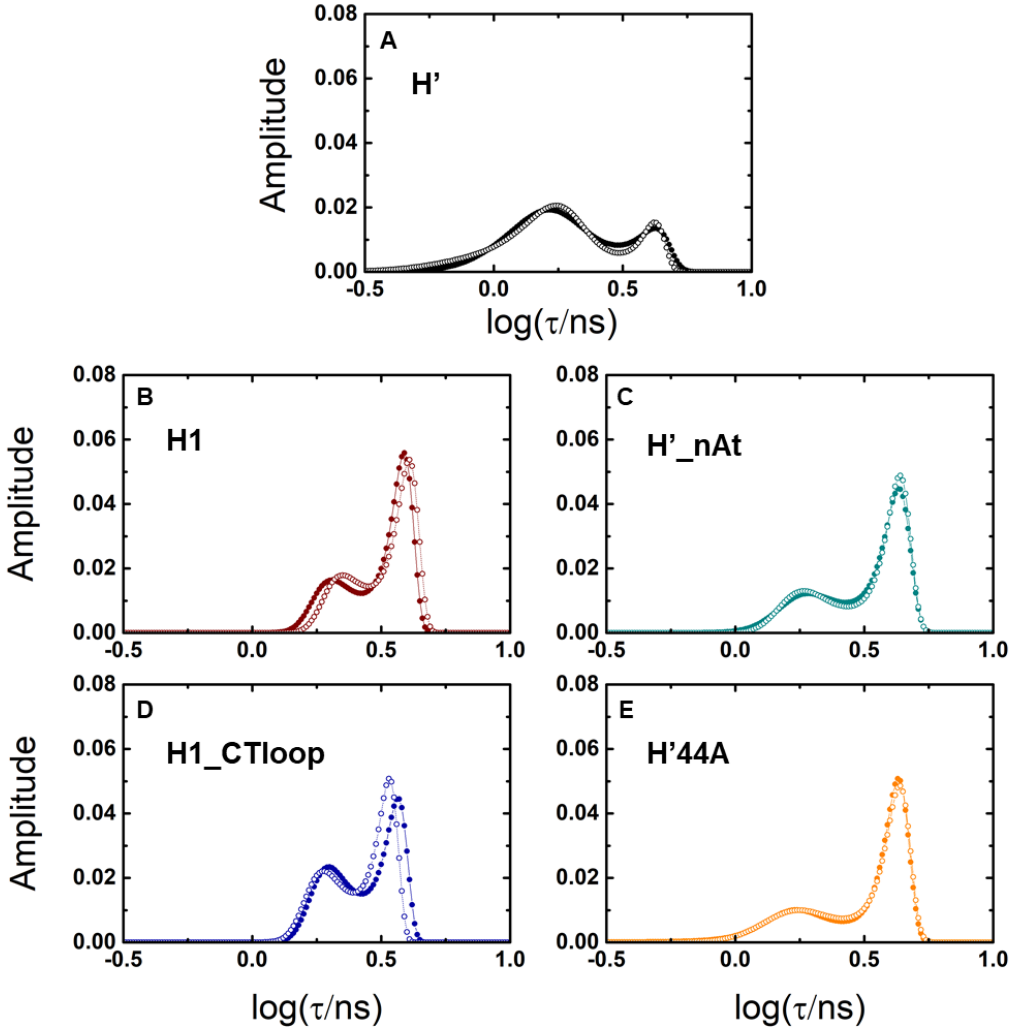


Figure S3. Reproducibility of MEM distributions from measurements on IHF-DNA_DA (design I) complexes. MEM outputs from two sets of measurements (closed versus open symbols) are shown for **(A)** IHF-H', **(B)** IHF-H1, **(C)** IHF-H'_nAt, **(D)** IHF-H1_CTloop, and **(E)** IHF-H'44A.

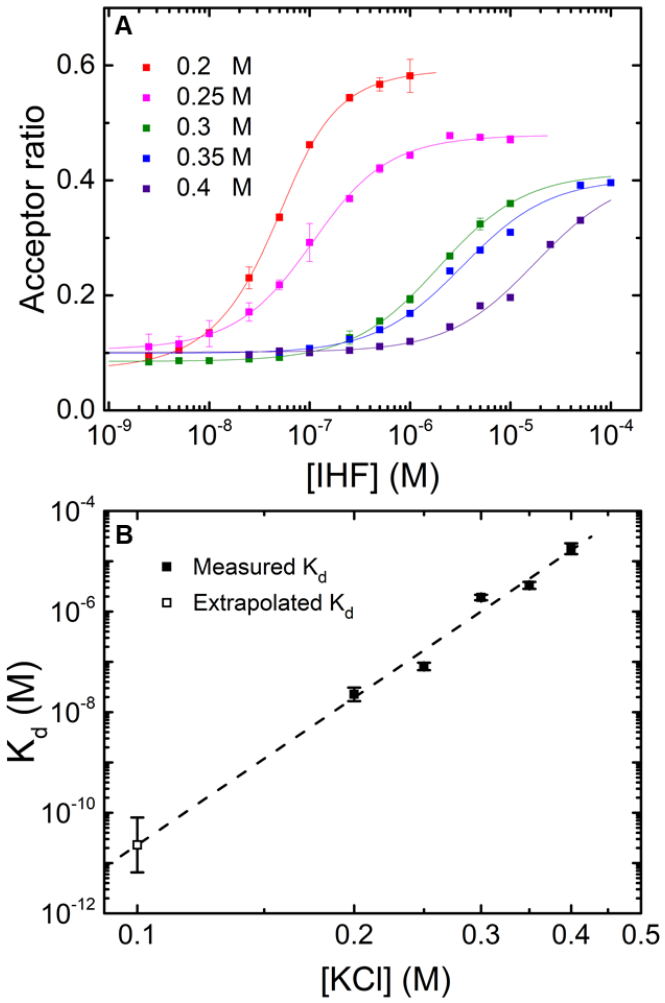


Figure S4. Binding affinity measurements on IHF-H' complex at varying salt concentrations. (**A**) Equilibrium binding isotherms for IHF binding to H' are shown from acceptor ratio measurements on 50 nM H'_DA (design I) as a function of [IHF]. The binding curves are for varying KCL concentrations: 200 mM (red), 250 mM (pink), 300 mM (green), 350 mM (blue), and 400 mM (purple). The lines are the best fit to a 1:1 binding isotherm, yielding the dissociation constants K_d shown in panel (B). (**B**) K_d values (filled squares) obtained from the data in panel (A) are plotted as a function of [KCl]. The dashed line is a linear fit to log(K_d) versus log([KCl]), with a slope of 9.7 ± 1.2. The K_d value at 100 mM KCl, extrapolated from this linear fit, is 23^{+57}_{-14} pM (open square).

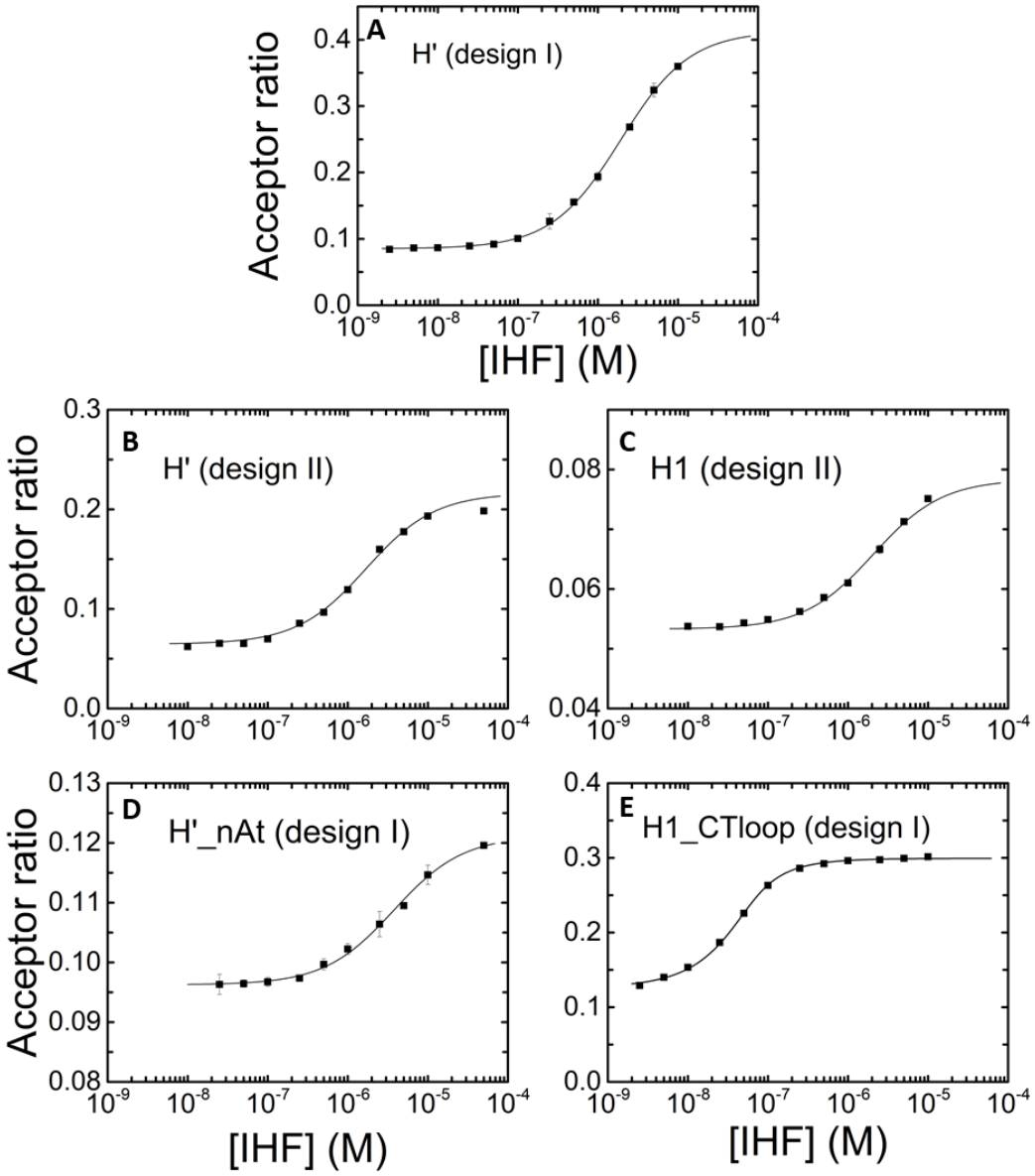


Figure S5. Binding affinity measurements for IHF-DNA complexes in 300 mM KCl. Acceptor ratio values measured on 50 nM DNA are plotted as a function of varying IHF concentration. Data are shown for (**A**) IHF-H' (design I); (**B**) IHF-H' (design II); (**C**) IHF-H1 (design II); **(D)** IHF-H'_nAt (design I); and **(E)** IHF-H1_CTloop (design I). Symbols are the mean of two independent sets of measurements; the error bars are the corresponding standard deviations. The continuous lines represent a fit to the data using a 1:1 binding isotherm.

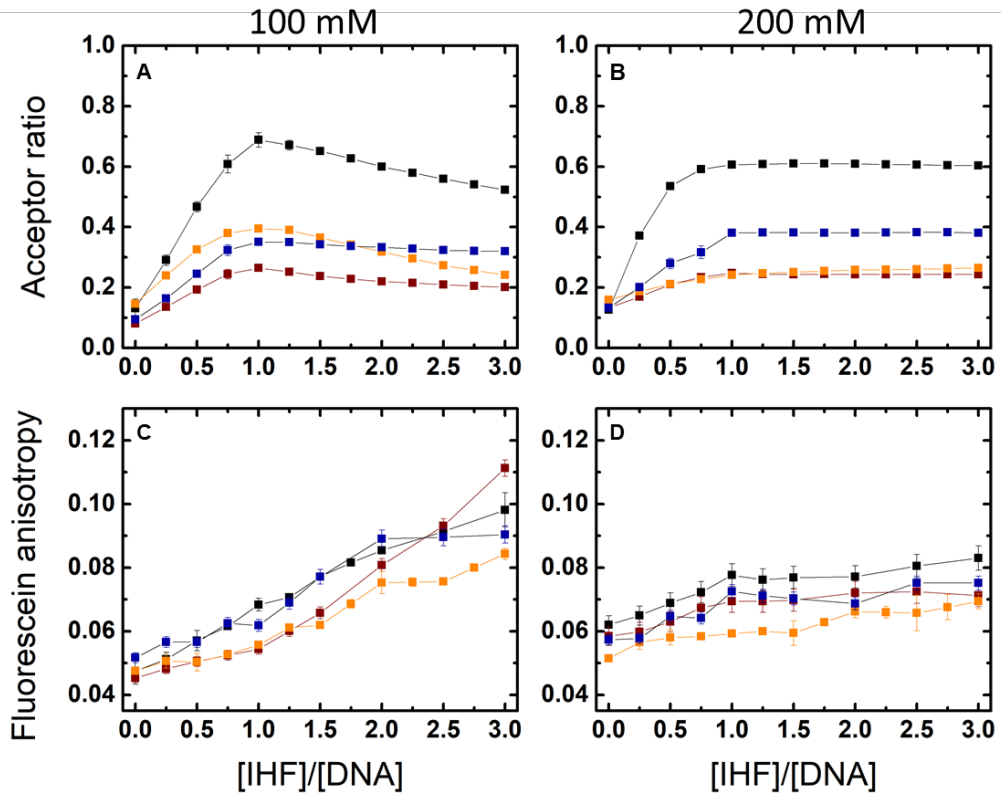
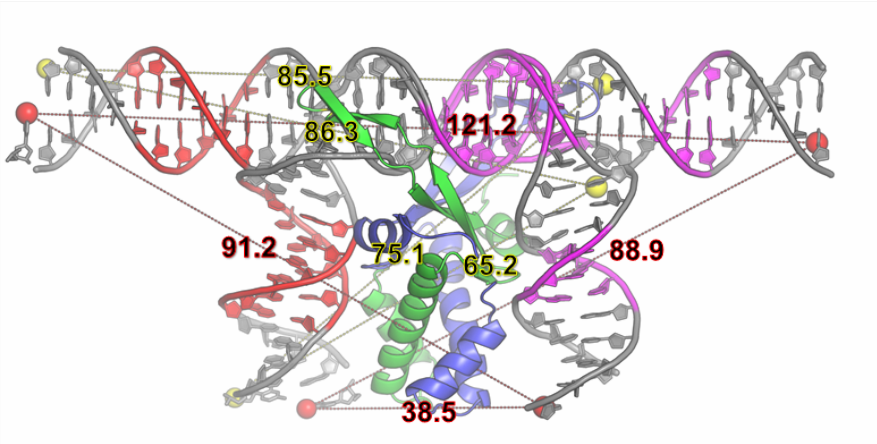


Figure S6. Binding isotherms for different IHF-DNA complexes. **(A, B)** Acceptor ratio values measured on 1 µM DNA_DA samples are plotted versus the ratio of IHF/DNA concentrations, for H' (black), H1 (maroon), H1_CTloop (blue), and H'44A (orange) in (**A**) 100mM KCl and (**B**) 200 mM KCl. **(C, D)** Corresponding anisotropy values are plotted versus IHF/DNA concentrations for measurements in (**C**) 100 mM KCl and (**D**) 200 mM KCl.

A



B

Structure	Estimated Distance (FRET)	
	Design I	Design II
Free DNA	121.2 Å (0.005)	85.5 Å
Fully bent complex	38.5 Å (0.83)	65.2 Å
A-tract side unkinked	91.2 Å (0.03)	86.3 Å
TTG side unkinked	88.9 Å (0.03)	75.1 Å

Figure S7. Computed distances and FRET values from structures. **(A)** Overlay of crystal structures of the fully bent complex with that of free DNA. The colors in the complex are identical to those used in Figure 1. The red (yellow) spheres indicate the atoms where the labels were attached in design I (design II) constructs. The dotted lines indicate the distances between the attachment points (red pair or yellow pair); four lines for each design correspond to distances computed for four different conformations: free DNA, fully bent complex, and two partially bent conformations, one in which the A-tract side (shown by the red nucleotides) is unkinked and the other in which the TTG side (shown by the magenta nucleotides) is unkinked. **(B)** The distances between the attachment points, computed as described above, are tabulated for the four different conformations in both design I and design II constructs. The FRET E values based on these distances are shown in parenthesis for design I constructs, calculated using $R_0 = 50$ Å; for design II constructs R_0 is not well known and hence no corresponding FRET E values are computed. We caution that these distance/FRET estimates are rough approximations and do not represent the actual distance/FRET between the labels in each of these structures, since they do not take into account the lengths and dynamics of the linkers that are used to attach the labels.