Supplementary Information

Probing the rate limiting step for intramolecular transfer of a transcription factor between specific sites on the same DNA molecule by $^{15}N_z$-exchange NMR spectroscopy

Kyoung-Seok Ryu,[1,2] Vitali Tugarinov,[1] and G. Marius Clore[1,*]

[1]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA, and [2]Korean Basic Science Institute, Ochang-Eup, Chunbuk-Do 363-883, South Korea.

EXPERIMENTAL

Sample preparation.

The homeodomain of HoxD9 (residues 285-344 of the full length protein, numbered here as residues 1-60 of the homeodomain) carrying a C6S mutation was prepared as described previously.[S1] The gene for the HoxD9 homeodomain was cloned into the pGEX-4T-1 vector using BamH I/Xho I restriction enzymes, and the plasmid transformed into *E. coli* BL21 DE3 strain. The expressed GST-HoxD9 fusion protein was initially purified using a GSTrap-FF affinity column (GE healthcare), and the GST tag was subsequently removed by thrombin-digestion. The resulting HoxD9 was further purified by ion-exchange column chromatography (Mono-S, GE healthcare). Protein concentration was determined from the UV absorption at 280 nm using a molar extinction coefficient of 11,460 $M^{-1}\cdot cm^{-1}$. Uniformly $^{2}H/^{15}N$- and $^{2}H/^{15}N/^{13}C$-labeled HoxD9 protein was obtained by growing *E. coli* in M9 minimal medium using 99.9% D_2O, $^{15}NH_4Cl$ as the sole nitrogen source, and either D-glucose-1,2,3,4,5,6,6-d_7 or D-glucose-$^{13}C_6$,1,2,3,4,5,6,6-d_7, respectively, as the sole carbon source.

Single-stranded DNA oligonucleotides were purchased from Integrated DNA Technologies. A 1:1 ratio of two complementary DNA strands was determined based on the measured UV absorption hypochromicity upon varying the ratios of the two strands. Double-stranded DNA duplexes were formed by heating (15 min at 90°C) and slow cooling, followed by further purification by gel filtration chromatography (Superdex 75, GE healthcare) and ion-exchange column chromatography (Mono-Q, GE healthcare) to remove any excess single stranded DNA. The concentration of DNA duplexes was determined from the UV absorption at 260 nm using the calculated molar extinction coefficients.[S2]

NMR samples were prepared as follows. HoxD9 was promptly added and mixed into a solution of DNA duplex, followed by dialysis into buffer containing 25 mM HEPES, pH 7.0 and 0.2 mM EDTA. The dialyzed sample was then dispensed in aliquots of appropriate concentration. These aliquots were further dialyzed into NMR buffer containing 20 mM HEPES/5 mM Na^+-HEPES, pH 7.0 and 30 mM NaCl. To achieve the desired concentration of DNA and protein the dialyzed samples were combined

in appropriate ratios. Finally, D_2O buffer (containing 20 mM HEPES/5 mM Na^+-HEPES, pH 7.0, 30 mM NaCl and 85% D_2O/5% H_2O) was added to each sample to yield a final D_2O concentration of 5%.

The molar ratios of HoxD9 to DNA were maintained constant for samples with different DNA concentrations. For the duplexes containing two specific sites (DNA-A^+B^+, DNA-A^+B^- and DNA-B^+A^+; cf. Figure 1B of main paper) the ratio of DNA to protein was 1.2:1 such that 80% of the duplexes have one HoxD9 bound to either one of the two specific sites and 20% of the duplexes have both specific sites occupied. For duplexes containing a single specific site (DNA-A and DNA-B; cf Figure 1B of main paper), the ratio of DNA-A and DNA-B to HoxD9 was kept constant at 1:1:1.

NMR spectroscopy

All NMR experiments were performed at 25°C using a Bruker Avance 600 MHz NMR spectrometer equipped with a z-gradient triple resonance cryoprobe. The $^1H_N/^{15}N$ chemical shifts for $^2H/^{15}N$-labeled HoxD9 complexed with the various DNA duplexes were very similar except for a small number of cross-peaks arising from residues in the N-terminal tail of HoxD9 (which are closest to the sites of mutation immediately 5' to the specific sites; cf. Figure 2 of main paper). Backbone assignments were obtained from through-bond heteronuclear scalar correlations in TROSY-based triple resonance spectra (HNCACB, HN(CO)CACB, HN(CA)CO and HNCO recorded on a sample of 0.3 mM $^2H/^{15}N/^{13}C$-labeled HoxD9 and 0.3 mM DNA-A. Assignment of $^1H_N/^{15}N$ cross-peaks corresponding to HoxD9 bound to the A or B sites of DNA-A^+B^+, DNA-A^+B^- and DNA-B^+A^+ were obtained by reference to the TROSY ^1H-^{15}N correlation spectra of the HoxD9/DNA-A and HoxD9/DNA-B complexes (Figure S1).

TROSY-based $^{15}N_z$-exchange spectra[S3] of $^2H/^{15}N$-labeled HoxD9 complexed with the different DNA duplexes were recorded in an interleaved manner with several *z*-exchange mixing times (10, 20, 40, 70, 130, 200, 300, and 500 ms). Data were processed using NMRpipe[S4] and analyzed with Sparky.[S5] Apparent first-order rate constants for the transfer of HoxD9 between the A and B specific sites were obtained by simultaneously best-fitting the experimental time dependence of the auto and exchange cross-peaks to a phenomenological, pseudo-first order two-site exchange system.[S1]

The time dependence of the magnetization of auto- (M_z^{AA}; M_z^{BB}) and exchange cross- (M_z^{AB}; M_z^{BA}) peaks for a two-site exchange system can be described by a matrix form of the McConnell equations[S6] as follows:

$$\frac{d}{dT_{mix}}\begin{pmatrix} M_z^{AA} \\ M_z^{BB} \\ M_z^{AB} \\ M_z^{BA} \end{pmatrix} = -\begin{bmatrix} (k_{AB}^{app} + R_1^A) & 0 & -k_{BA}^{app} & 0 \\ 0 & (k_{BA}^{app} + R_1^B) & 0 & -k_{AB}^{app} \\ -k_{AB}^{app} & 0 & (k_{BA}^{app} + R_1^B) & 0 \\ 0 & -k_{BA}^{app} & 0 & (k_{AB}^{app} + R_1^A) \end{bmatrix}\begin{pmatrix} M_z^{AA} \\ M_z^{BB} \\ M_z^{AB} \\ M_z^{BA} \end{pmatrix} \quad \text{(S1)}$$

where $k_{AB}^{app} = k_{AB}^{inter}\left[S_B^{free}\right] + k_{AB}^{intra}$, $k_{BA}^{app} = k_{BA}^{inter}\left[S_A^{free}\right] + k_{BA}^{intra}$, $\left[S_k^{free}\right]$ and R_1^k are the concentration of free specific DNA site k and the longitudinal relaxation rate of the protein in complex with specific DNA site k ($k \in A,B$), respectively, and k^{inter} and k^{intra} are the rate constants for intermolecular (second-order) and intramolecular (first order) transfer between the specific sites (cf. Figure 1C, main text), as described in the main text. The mixing time dependence of the $^{15}N_z$-exchange data at five different concentrations of specific DNA sites and for several residues were fit simultaneously in

MATLAB[S7] via exponentiation of the matrix in Eq. (S1), optimizing global values of k_{AB}^{app} and k_{ba}^{app}, together with R_1^A, R_1^B for each individual residue (with $R_1^A = R_1^B$) and the values of $M_z^{AA}(0), M_z^{BB}(0)$ at each DNA specific site concentration for each individual residue. (In the main text, for plotting purposes, we averaged the data and best-fit curves for the different residues). The dependence of the fitted values of k_{AB}^{app} and k_{BA}^{app} on the concentration of free specific DNA sites was analyzed via linear regression and plotted in Figures 3A-D of the main text.

REFERENCES

(S1) Iwahara, J.; Clore, G. M. *J. Am. Chem. Soc.* 2006, *128*, 404-405.
(S2) Tataurov, A. V.; You, Y.; Owczarzy, R. *Biophys. Chem.* 2008, *133*, 66-70.
(S3) Sahu, D.; Clore, G. M.; Iwahara, J. *J. Am. Chem. Soc.* 2007, *129*, 13232-13237.
(S4) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* 1995, *6*, 277-293.
(S5) Goddard, T. D.; Kneller, D. G. *SPARKY 3, University of California, San Francisco*.
(S6) McConnell, H. M. *J. Chem. Phys.* 1958, *28*, 430-431.
(S7) MATLAB, The Language of Technical Computing. www.mathworks.com/products/matlab

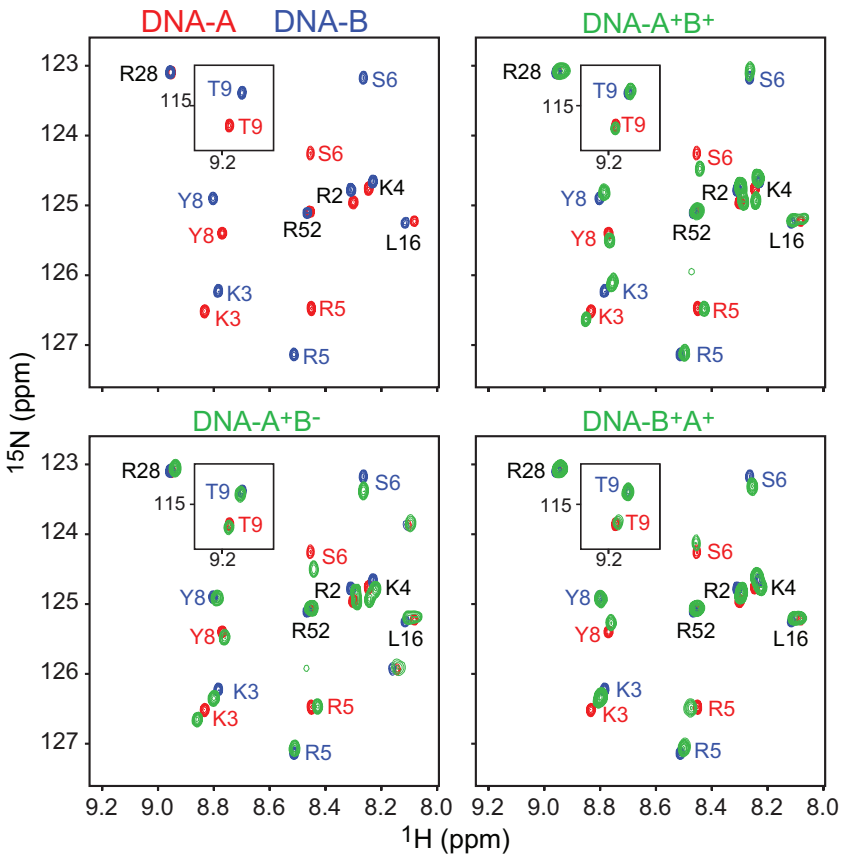

Figure S1. TROSY ^{1}H-^{15}N correlation spectra of ^{2}H/^{15}N-labeled HoxD9 complexed to the five DNA duplexes used in the current study. DNA-A and DNA-B contain a single specific site each and the spectra of the HoxD9/DNA-A (red) and HoxD9-DNA-B (blue) complexes are superimposed in the top left panel. The other panels show the superpositions of the spectra of the complexes (green) with DNA-A^{+}B^{+} (top right, green), DNA-A+B^{-} (bottom left, green), and DNA-B^{+}A^{+} (bottom right, green), each possessing two specific sites (A and B) on a single duplex, onto the spectra of the single site complexes with DNA-A (red) and DNA-B (blue). Cross-peaks with distinct $^{1}H_N/^{15}N$ chemical shifts in the complexes with DNA-A and DNA-B are labeled in red and blue, respectively, in each panel.