Supplementary Information

Detecting Transient Intermediates in Macromolecular Binding by Paramagnetic NMR

Junji Iwahara and G. Marius Clore*

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and

Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520

Lineshape simulations based on the McConnell equations

For two-state exchange, the NMR lineshape is given by the imaginary part of the function $G(v)$ derived from the McConnell equations:[S1,S2]

$$G(v) = \frac{-iC[\lambda_A p_B + \lambda_B p_A + 2 p_A p_B k_{ex}]}{\lambda_A \lambda_B - p_A p_B k_{ex}^2}$$ [1],

where

$$\lambda_A = R_{2,A} + p_B k_{ex} - 2\pi i(v_A - v)$$ [2],

$$\lambda_B = R_{2,B} + p_A k_{ex} - 2\pi i(v_B - v)$$ [3],

v_A and v_B are the chemical shifts of states A and B, respectively, in Hz; p_A and p_B are the populations of the two states; R_{2A} and R_{2B} are the transverse relaxation rates of the two states; and k_{ex} is the overall exchange rate (given by the sum of rate constants for the forward and backward transitions). The transverse relaxation rates are given by:

$$R_{2,A} = R_{2,A}^{diamagnetic} + \Gamma_{2,A}$$ [4],

$$R_{2,B} = R_{2,B}^{diamagnetic} + \Gamma_{2,B}$$ [5],

where $R_{2,A}^{diamagnetic}$ is the transverse relaxation rate arising from non-paramagnetic mechanisms such as nuclear dipole-dipole interactions and CSA; and Γ_2 the transverse relaxation rate arising from paramagnetic relaxation (*i.e.* PRE).

The conditions employed for the simulations shown in Fig. 1 were as follows: $p_A = 0.99$; $p_B = 0.01$; $|v_A - v_B| = 100$ Hz; $R_{2,A}^{diamagnetic} = R_{2,B}^{diamagnetic} = 50$ s^{-1}. For the red lines (with PRE), $\Gamma_{2,A} = 2.0$ s^{-1} and $\Gamma_{2,B} = 5.6 \times 10^3$ s^{-1}; for the black lines (without PRE), $\Gamma_{2,A} = \Gamma_{2,B} = 0$ s^{-1}. The chemical shifts for the diamagnetic and paramagnetic are assumed to be identical since a paramagnetic system with an isotropic electron g-tensor such as Mn^{2+} does not produce a pseudo-contact shift. The value of Γ_2^{app} shown in Fig. 1 was calculated as the difference between the apparent transverse rates with and without PRE, which were obtained by Lorentzian-curve fitting of the simulated lineshapes.

NMR analysis of timescale of intermolecular exchange process for Hox-D9-DNA complex

The timescale of the intermolecular exchange process for the Hox-D9 homeodomain-DNA complex as a function of ionic strength was analyzed using the approach described elsewhere.[S3] In this analysis, we used two 24-bp DNA duplexes, *a* and *b*, that differ at only a single base pair position adjacent to the homeodomain specific target site TAATGG. As a consequence, the ^{1}H/^{15}N-chemical shifts of several backbone amide groups are slightly different in the two complexes, while the affinity of the two oligonucleotides for the homeodomain remains essentially unaltered. When ^{1}H-^{15}N HSQC spectra were measured on 1:1 mixture of the complexes comprising either DNA *a* or DNA *b* (referred to as complexes *a* and *b*, respectively, hereafter), two distinct signals corresponding to complexes *a* and *b* were observed at 20 mM NaCl (Fig. S1B). Upon increase of the salt concentration to 160 mM NaCl, however, the observed signals for the mixture are reduced to single peaks located at the average positions of those measured individually on complexes *a* and *b* at 160 mM NaCl (cf. compare Figs. S1B and S1C). These results indicate that the exchange process in which a protein is transferred from DNA *a* to DNA *b* (and vice versa) is slow on the NMR chemical shift timescale at 20 mM NaCl but fast at 160 mM NaCl. The process in the slow exchange regime was analyzed quantitatively at 20, 30, 40, 50, and 60 mM NaCl using a 2D ^{1}H-^{15}N correlation experiments in which exchange between the ^{15}N z-magnetizations of distinct species occurs during the mixing time following the

$t_1(^{15}N)$-evolution period.[S4] This analysis yields apparent rate constants k_{ab}^{app} and k_{ba}^{app} for the transfer from a to b and from b to a, respectively. The data exhibited a linear relationship between $\log k_{ab}^{app}$ (and $\log k_{ba}^{app}$) and $\log[NaCl]$ (Fig. S1D). The exchange rate, k_{ex}, in Eqs. 2 and 3 corresponds to the sum of these two rate constants. Based on this empirical relationship, the values of k_{ex} for exchange between complexes a and b at 100 and 160 mM NaCl were extrapolated to be ~100 and ~300 s^{-1}, respectively. Values of k_{ex} for the exchange process between DNA cognate sites that includes all four possible transitions, $a \rightarrow a$, $a \rightarrow b$, $b \rightarrow a$ and $b \rightarrow b$, are twice the values reported above, since the measurements were carried out for a 1:1 mixture of complexes a and b. The linear relationship between $\log k_{ab}^{app}$ ($\log k_{ba}^{app}$) and $\log[NaCl]$ seems reasonable considering the following observations. First, the ionic-strength dependence of the equilibrium dissociation constant (K_{diss}) for a protein-DNA complex is theoretically described by linear relationship between $\log K_{diss}$ and $\log[M^+]$, where M^+ represents a cation.[S5,S6] (The ionic-strength dependence is due to competition between cations and the protein for the negatively charged phosphate groups on the DNA resulting in destabilization of electrostatic interactions between the protein and DNA at higher ionic strengths.) Second, an empirical linear relationship has been observed between $\log k^{on}$ (where k^{on} is the association rate constant) and $\log[NaCl]$ for electrostatically assisted binding of protein-protein complexes,[S7] which is likely to also hold true for protein-DNA interactions. It should be noted that k_{ex} obtained from this analysis relates to translocation between cognate (specific) sites. Although translocation between cognate and non-cognate (non-specific) sites is more relevant to the PRE study described in the main text, k_{ex} obtained here represents a lower limit for the exchange process between cognate and non-cognate sites, since the exchange rate is likely to be significantly larger due to the shorter lifetime of a non-specifically bound complex.

Analysis of residual dipolar couplings for the Hox-D9/DNA complex

Using 2D ^1H-^{15}N IPAP-HSQC experiments,[S8] backbone NH ($^1D_{NH}$) residual dipolar couplings were measured for the Hox-D9/DNA complex at 20 and 160 mM NaCl in a dilute liquid crystalline medium of phage Pf1 (12 mg/ml). [S9,S10]. Using the program SSIA,[S11] we compared the observed $^1D_{NH}$ data with those predicted from the 2.4 Å resolution crystal structure of the Antp homeodomain/DNA complex (PDB 9ANT; Ref. S12) by singular value decomposition.

Excellent agreement was obtained for both datasets (Figs. S2A, S2B) with dipolar coupling R-factors[S13] of 14% and 15% for the complexes in 20 mM and 160 mM NaCl, respectively. These values for the dipolar coupling R-factor fall well within the range expected for 1.5-2.5 Å resolution crystal structures.[S14-S16] Since the molecular alignment of the complex is caused by electrostatic interactions with the Pf1 phage, the magnitude of the axial component of the alignment tensor is smaller in 160 mM NaCl.[S17] Nonetheless, the profiles of the two datasets are essentially same (Fig. S2C) and the $^1D_{NH}$ values at the two salt concentrations are highly correlated (Fig. S2D) with a correlation coefficient of 0.99. These results indicate that the structures of the specific (ground state) complex at 20 and 160 mM NaCl are essentially identical.

K_{diss} measurements with fluorescence anisotropy

The equilibrium dissociation constants, K_{diss}, for the complexes between the Hox-D9 homeodomain and DNA 1A (with a cognate site; Fig. 4) or DNA 2B (without a cognate site; Fig. 4) were determined by fluorescence anisotropy measurements[S18] using derivatives with rhodamine conjugated at the 5'-terminius of one of the strands. The rhodamine-labeled DNA fragments were purchased from Midland Certified Reagent Company and the DNA duplexes were purified by polyacrylamide gel eletrophoresis under native conditions. Fluorescence anisotropy (λ_{ex} = 550 nm, λ_{em} = 580 nm) was measured at 25°C on the conjugated rhodamine using a Jobin Yvon FluoroMax-3 spectrometer. For the determination of K_{diss}, the fluorescence anisotropy was recorded upon titrating the Hox-D9 homeodomain (ten concentration points from 0.2 to 60 nM for DNA 1A; and from 0.2-2000 nM for DNA 2B) to a 3-ml solution of rhodamine-labeled DNA (1 nM for DNA 1A; 8 nM for DNA 2B) dissolved in 10 mM Tris·HCl (pH 7.2), 10% Glycerol and 100 mM NaCl (or 160 mM NaCl). The values of K_{diss} were calculated from the titration data as described previously.[S19]
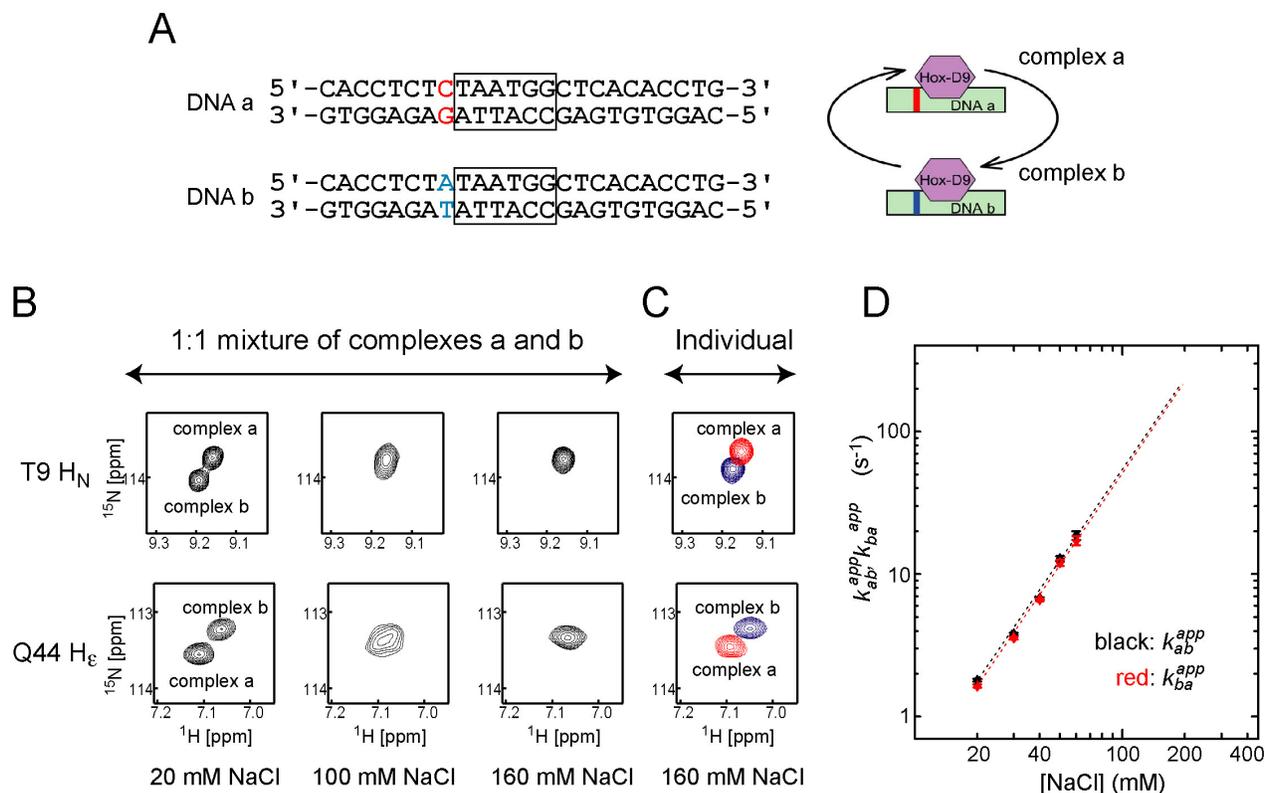
Fig. S1. Analysis of the exchange timescale for the Hox-D9/DNA complex. (A) The two DNA fragments used in this analysis differ at a single base pair (indicated in color). (B) Signals from Thr9 backbone NH and Gln44 side-chain amide groups in the ^{1}H-^{15}N HSQC spectra recorded on 1:1 mixtures of complexes *a* and *b* at 20, 100 and 160 mM NaCl. (C) The same region of the ^{1}H-^{15}N HSQC spectra of complex *a* (red) and complex *b* (blue) measured individually at 160 mM NaCl and superimposed. (D) Exchange rates as a function of salt concentration (plots are with log scales for both axes). Values of the rate constants for the transition from one complex to the other (k_{ab}^{app}, for *a* to *b*; k_{ba}^{app}, for *b* to *a*) were determined from time-course data of intensities of auto- and exchange cross-peaks observed in ^{15}N$_z$-exchange experiments[S4] using the fitting procedure described preciously.[S3] The sample contained 0.4 mM Hox-D9, 0.3 mM DNA *a* and 0.3 mM DNA *b* at pH 6.8 (i.e. ~0.2 mM for complexes *a* and *b*, and ~0.1 mM for free DNA *a* and free DNA *b*).
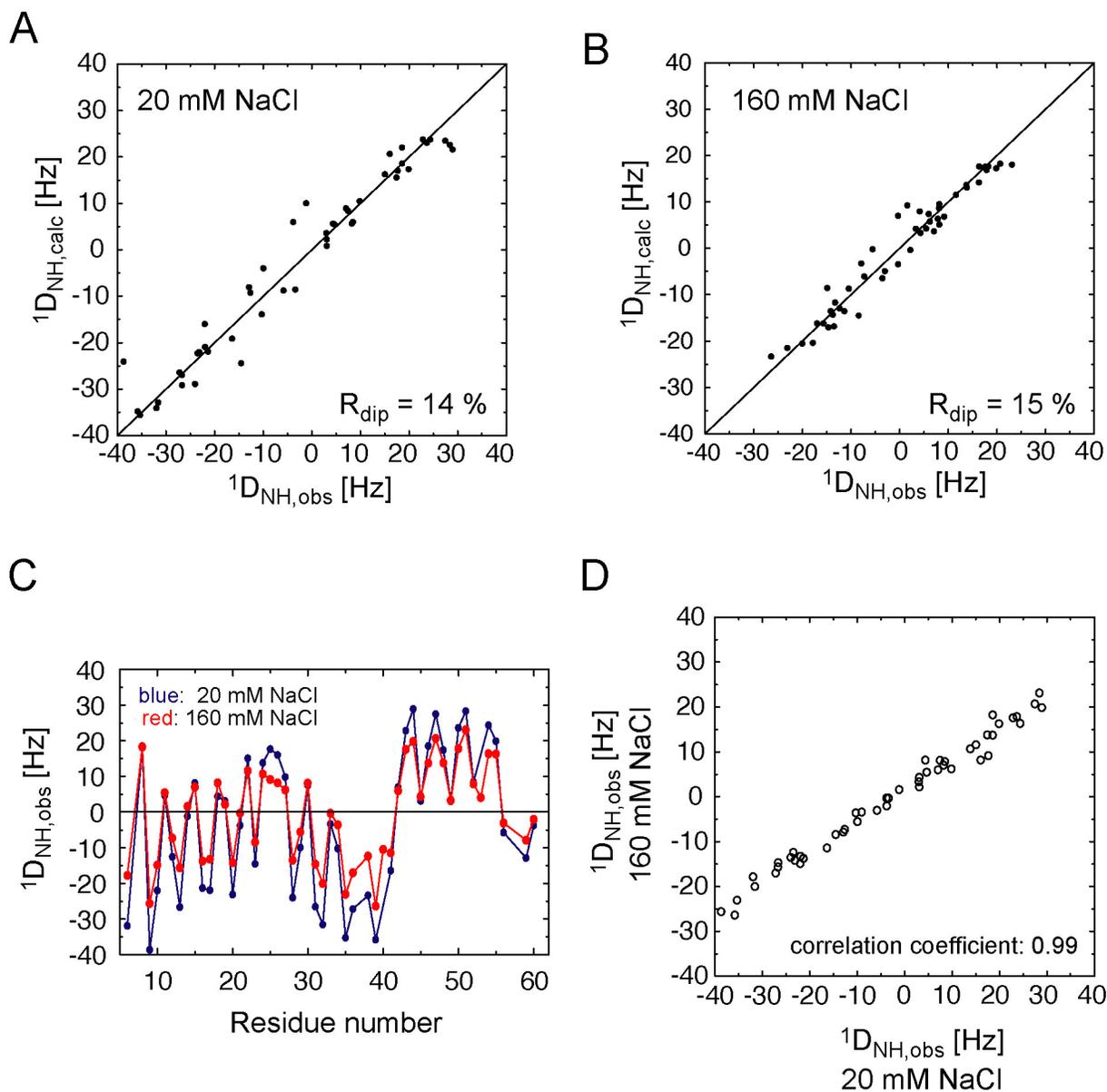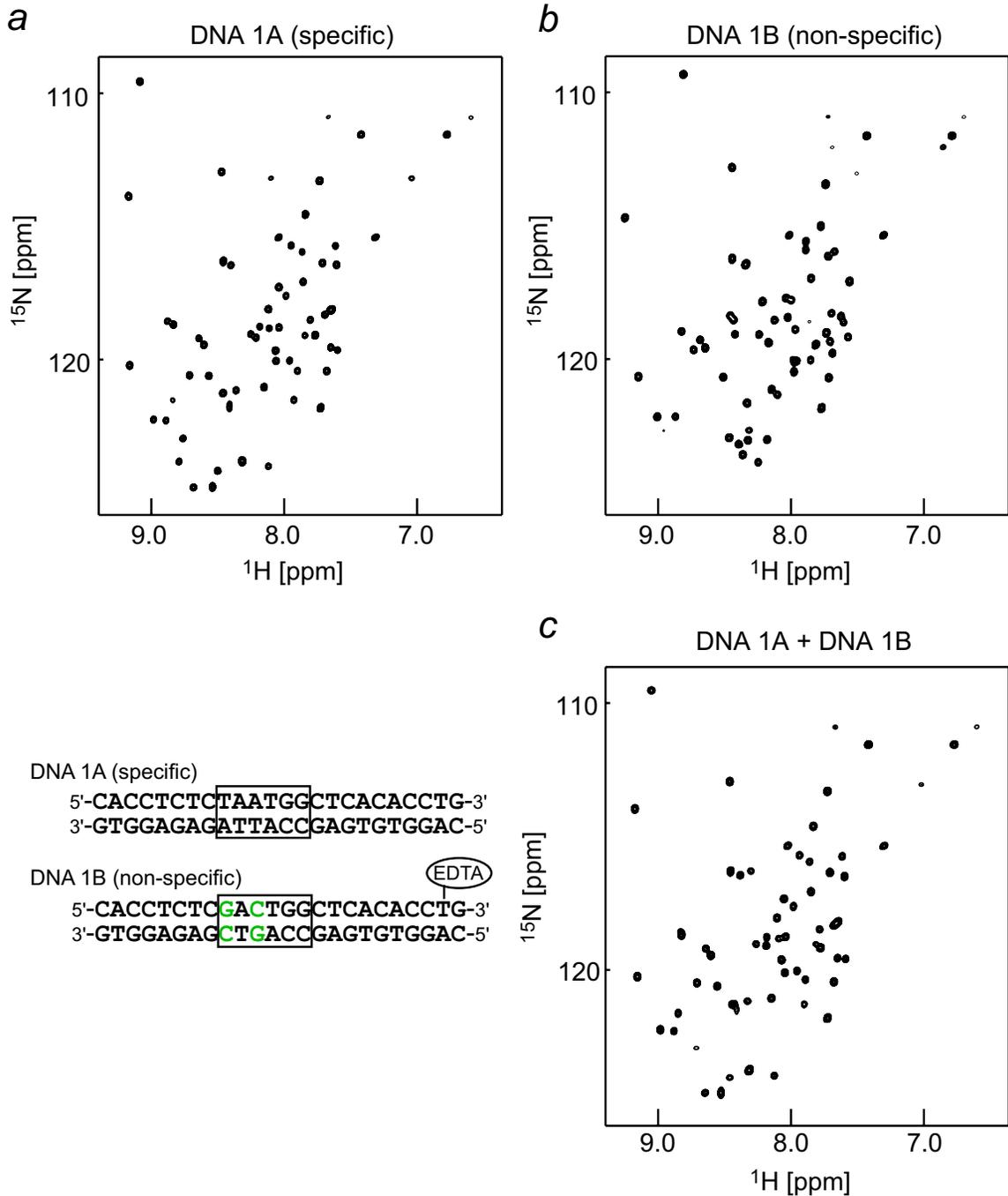
Fig. S2. Residual dipolar coupling $^1D_{NH}$ data measured on the Hox-D9 homeodomain/DNA complex in 20 and 160 mM NaCl. Molecular alignment was induced with 12 mg/ml phage Pf1.[S9,S10] (A, B) Correlations between observed $^1D_{NH}$ values and those predicted from the 2.4 Å crystal structure of the Antp homeodomain/DNA complex (PDB 9ANT; ref. S12) at 20 mM (A) and 160 mM (B) NaCl. In the fitting procedure, the magnitude (D_a^{NH}) and rhombicity (η) of the alignment tensor were calculated by singular value decomposition to be -24.2 Hz and and 0.02, respectively, at 20 mM NaCl, and -15.3 Hz and 0.15, respectively, at 160 mM NaCl. The smaller $|D_a^{NH}|$ value at160 mM NaCl can be attributed to the fact that the molecular alignment by Pf1 phage is caused by electrostatic interactions.[S17] The dipolar coupling R-factor (R_{dip}), which is defined as the ratio of the r.m.s. difference between observed and calculated values and the predicted rms difference if the vectors were randomly oriented (given by $\{2D_a^2[4+3\eta^2]/5\}^{1/2}$; ref. S13), was 14% at 20 mM NaCl and 15% at 160 mM NaCl. (C) Plots of $^1D_{NH}$ for individual residues. (D) Correlation between $^1D_{NH}$ data measured at 20 and 160 mM NaCl.
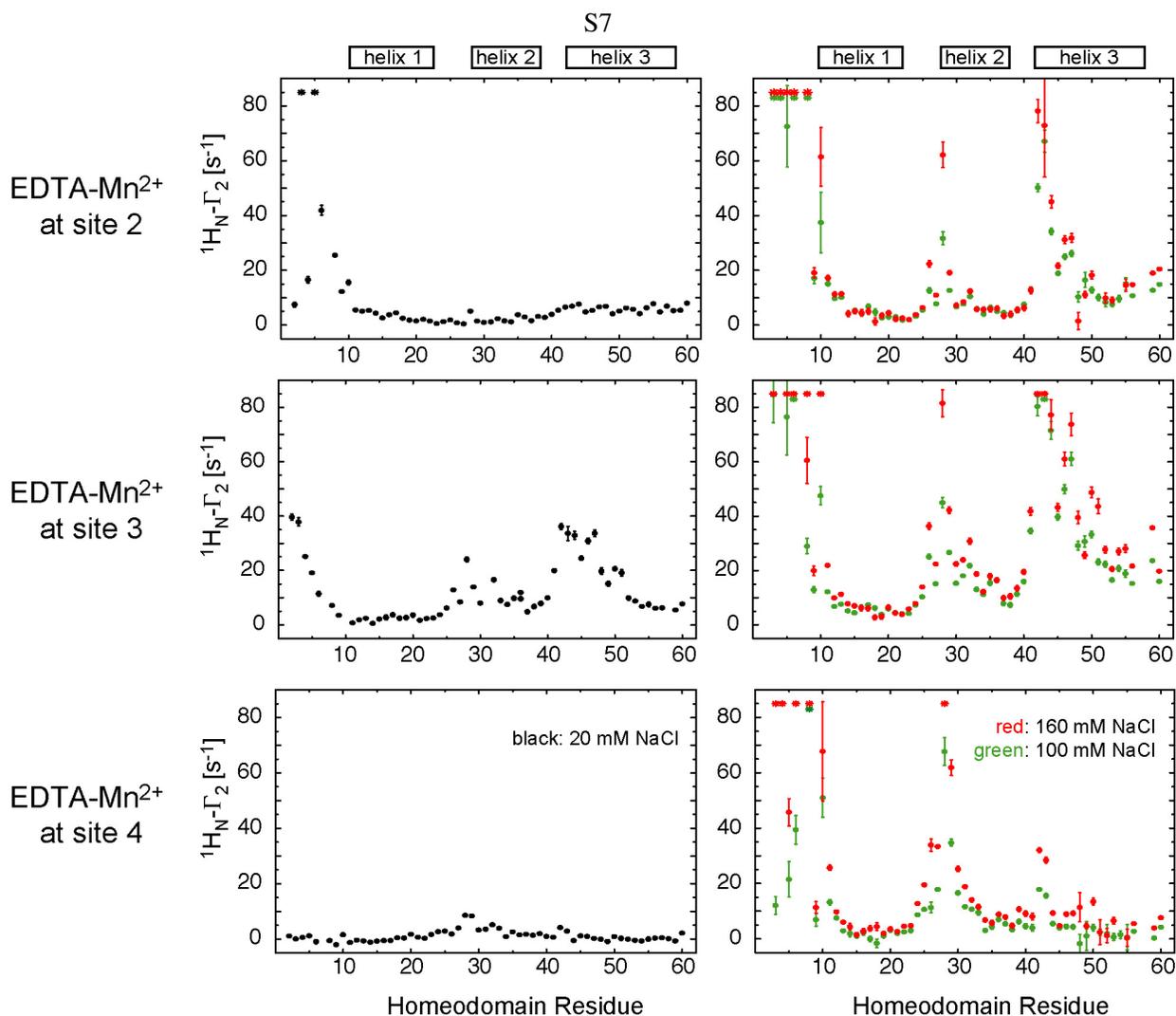
Fig. S3. Plots of intermolecular PRE ^1H_N-Γ_2 data arising from EDTA-Mn^{2+} at sites 2, 3 and 4 as a function of residue number (Black, data at 20 mM NaCl; Green, 100 mM NaCl; Red, 160 mM NaCl). Asterisks represent those residues whose $^1H/^{15}N$ cross-peaks are broadened beyond the limits of detection by the PRE. Data for site 1 is shown in Fig. 2b. These data are mapped on a structural model of the complex in Fig. 3. The concentrations of Hox-D9 and DNA were 0.4 mM and 0.6 mM, respectively. We note that the highly remote possibility of transient binding of Hox-D9 homeodomain to the blunt ends of the DNA duplex as a possible mechanism to account for the PRE data at high salt can be completely excluded on several counts: (a) The magnitude of the PREs observed at 160 mM NaCl for all 4 sites (this figure and Fig. 2c of main text) are comparable, yet sites 2 and 3 are located 5 and 7 base pairs in, respectively, from the ends, while sites 1 and 4 are located only 2 base pairs in from the ends. Given the $\langle r^{-6} \rangle$ dependence of the PRE, transient binding of free protein to the blunt ends of the DNA would predict much lower PREs for sites 2 and 3 than sites 1 and 4. Moreover, the PREs observed for residues at the N-terminus of helix 3 are in fact much larger for sites 2 and 3 than sites 1 and 4. (b) Translocation between specific and non-specific binding sites occurs via *direct* transfer of the protein between DNA sites either through sliding along the DNA (intramolecular transfer) or by translocation between DNA molecules via intermolecular transfer without ever involving the intermediary of free protein;[S3] hence transient binding of free protein to the blunt ends of the DNA is simply never an issue. (c) Previous

studies using the PRE to study a non-specific protein-DNA complex involving HMG-1A in which multiple sites are occupied on the DNA provide absolutely no evidence for binding to the blunt ends of the DNA.[S19] (d) Any possible transient binding to the blunt ends of the DNA will be orders of magnitude weaker (and therefore orders of magnitude lower occupancy) than regular non-specific binding ($K_D \sim 270$ nM at 100 mM NaCl) and would certainly not be expected to occur via the same mode (and therefore involve the same interaction surface) as that employed for specific and non-specific binding since the structure of the blunt ends of the DNA bear no resemblance to the major or minor grooves that constitute the DNA binding site for Hox-D9; (e) Transient binding to the blunt ends of the DNA cannot account for the data shown in Fig. 4 of the main text which provides a direct demonstration of the existence of sliding along the DNA. Another highly remote theoretical consideration is that transient collisions between the EDTA-Mn^{2+} moiety of one molecule of DNA and the homeodomain bound to another DNA molecule could potentially give rise to a solvent PRE effect whereby solvent exposed regions of the protein would be selectively broadened. This remote possibility can also be completely excluded since: (a) the solvent PRE would not be affected by salt concentration and translocation exchange rates; (b) the observed PRE profiles at both low salt and high salt are inconsistent with a solvent PRE effect (e.g. the highly solvent exposed regions in the complex, such as the C-terminal ends of helices 1 and 2, do not exhibit any significant line broadening; cf. Fig. 3 of main text); and (c) the concentration of dT-EDTA-Mn^{2+} DNA employed is simply too low (sub-millimolar) to give rise to any measurable solvent PRE.[S20] As noted in the main text, the PRE data at high salt are fully consistent with a non-specific mode of DNA binding adopted during the target search process that is very similar to that in the specific complex, and therefore the population of any potential species involving alternate interaction surfaces on the protein, should these exist, are below the limits of detection. In this regard, it is worth mentioning that intermolecular translocation involves the formation of a transient ternary encounter species. One might therefore postulate that this would require the utilization of protein contact surface that is distinct from that in the specific complex given that the same part of the protein surface cannot interact simultaneously with two DNA molecules. However, as discussed in ref. S3, there is no need to invoke the utilization of alternate protein contact surfaces if one envisions a simple physical model in which the homeodomain is considered as a two sub-domain protein comprising the helix-turn-helix core in the major groove and the N-terminal tail in the minor groove, each of which can come on and off the DNA independent of the other. The microscopic equilibrium dissociation constants would be expected to be quite large (particularly as truncation of the N-terminal domain results in weak DNA binding) with rapid dissociation and reassociation of the individual sub-domains, while still maintaining a global binary protein-DNA complex (and therefore a very low overall macroscopic equilibrium dissociation constant). At high concentrations of free DNA, the dissociated sub-domain could readily attach itself to the relevant groove on a free DNA molecule, leading to translocation of the homeodomain from one oligonucleotide to the other.[S3] This type of mechanism is likely to be general to virtually all DNA binding proteins since these invariably consist of two or more domains (e.g. core plus positively charged N- or C-terminal tails, repeating units, or multiple subunits).

a DNA 1A (specific)

b DNA 1B (non-specific)

c DNA 1A + DNA 1B

DNA 1A (specific)
5'-**CACCTCTC**|**TAATGG**|**CTCACACCTG**-3'
3'-**GTGGAGAG**|**ATTACC**|**GAGTGTGGAC**-5'

DNA 1B (non-specific)
5'-**CACCTCTC**|**GAC TGG**|**CTCACACCTG**-3'
3'-**GTGGAGAG**|**CTG ACC**|**GAGTGTGGAC**-5'

EDTA

Fig. S4. (a, b) $^1H/^{15}N$ HSQC spectra of $^2H/^{15}N$-labeled Hox-D9 homeodomain bound to specific (panel a) or non-specific (panel b) DNA fragments at 160 mM NaCl (in the fast exchange regime). DNA fragments 1A and 1B are the same as those shown in Fig. 4. (c) $^1H/^{15}N$ HSQC spectrum measured on Sample 1 (as defined in Fig. 4 comprising Protein:DNA 1A:DNA 1B in a ratio of 1:1.5:1.5 with a protein concentration of 0.4 mM) at 160 mM NaCl. For these spectra, dT-EDTA of DNA 1B is in the Ca^{2+}-chelated state. The spectrum of the mixture in (c) is very similar to that of the specific complex in (a) but very different from that of the non-specific complex shown in (b).

References for Supplementary Information

S1. McConnell, H. M. Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430-431 (1958).

S2. Reuben, J. & Fiat, D. Nuclear Magnetic Resonance Studies of Solutions of the Rare-Earth Ions and Their Complexes. IV. Concentration and Temperature Dependence of the Oxygen-17 Transverse Relaxation in Aqueous Solutions. *J. Chem. Phys.* **51**, 4918-4927 (1969).

S3. Iwahara, J. & Clore, G. M. Direct Observation of enhanced translocation of a homeodomain between DNA cognate sites by NMR exchange spectroscopy. *J. Am. Chem. Soc.* **128**, 404-405 (2006).

S4. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727-734 (1994).

S5. Record, M. T. Jr., Lohman, T. M. & de Haseth, P. Ion effects on ligand-nucleic acid interactions. *J. Mol. Biol.* **107**, 145-158 (1976).

S6. Ha, J.-H., Spolar, R. S. & Record, M. T. Jr. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J. Mol. Biol.* **208**, 801 (1989).

S7. Schreiber, G. & Fersht, A. R. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* **3**, 427-431 (1996).

S8. Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373-378 (1998).

S9. Clore, G.M., Starich, M. R. & Gronenborn, A.M. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am. Chem. Soc.* **120**, 10571-10572 (1998).

S10. Hansen, M. R., Mueller, L. & Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**, 1065-1074 (1998).

S11. Zweckstetter, M. & Bax, A. Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR. *J. Am. Chem. Soc.*, **122**, 3791-3792 (2000).

S12. Fraenkel, E. & Pabo, C. O. Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. *Nat. Struct. Biol.* **5**, 692-697(1998).

S13. Clore, G. M. & Garrett, D. S. R-factor, Free R, and Complete Cross-Validation for Dipolar Coupling Refinement of NMR Structures. *J. Am. Chem. Soc.* **121**, 9008-9012 (1999).

S14. Bax, A., Kontaxis, G. & Tjandra, N. Dipolar couplings in macromolecular structure determination. *Methods Enzymol.* **339**, 127-174 (2001).

S15. Williams, D. C., Cai, M. & Clore, G. M. Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1·Sox2·Hoxb1-DNA ternary transcription factor complex. *J. Biol. Chem.* **279**, 1449-1457 (2004).

S16. Williams, D. C., Lee, J.-Y., Cai, M., Bewley, C. A. & Clore, G. M. Crystal structures of the HIV-1 inhibitory cyanobacterial protein MVL free and bound to $Man_3GlcNAc_2$: structural basis for specificity and high-affinity binding to the core pentasaccharide from N-linked oligomannoside. *J. Biol. Chem.* **280**, 29269-29276 (2005).

S17. Zweckstetter, M. & Bax, A. Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage *J. Biomol. NMR* **20**, 365-377 (2001).

S18. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd Ed., Kluwer Academic/Plenum Publishers, New York (1999).

S19. Iwahara, J., Schwieters, C. D. & Clore, G. M. Characterization of nonspecific protein-DNA interactions by ^1H paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **126**, 12800-12008 (2004).

S20. Donaldson, L. W. *et al*. Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *J. Am. Chem. Soc.* **123**, 9843-9847 (2001).