Supporting Information

Direct Observation of Enhanced Translocation of a Homeodomain between DNA Cognate Sites by NMR Exchange Spectroscopy

Junji Iwahara and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

1. Determination of apparent translocation rates k_{ab}^{app} and k_{ba}^{app}

The phenomenological mixing time-dependence of the auto- and exchange cross peaks for a simple two-site exchange system

$$\mathrm{A} \underset{k_{ba}^{app}}{\overset{k_{ab}^{app}}{\rightleftarrows}} \mathrm{B}$$

can be described by the McConnell equations (S1) comprising Eq. 1-4 (Note that in this experiment $M_z^{aa}(\infty) = M_z^{bb}(\infty) = 0$ instead of the Boltzman magnetization):

$$\frac{dM_z^{aa}}{dt} = k_{ba}^{app} M_z^{ab} - M_z^{aa}(k_{ab}^{app} + R_1^{\text{complex } a}) \qquad [1]$$

$$\frac{dM_z^{bb}}{dt} = k_{ab}^{app} M_z^{ba} - M_z^{bb}(k_{ba}^{app} + R_1^{\text{complex } b}) \qquad [2]$$

$$\frac{dM_z^{ab}}{dt} = k_{ab}^{app} M_z^{aa} - M_z^{ab}(k_{ba}^{app} + R_1^{\text{complex } b}) \qquad [3]$$

$$\frac{dM_z^{ba}}{dt} = k_{ba}^{app} M_z^{bb} - M_z^{ba}(k_{ab}^{app} + R_1^{\text{complex } a}) \qquad [4]$$

M_z^{aa}: z-magnetization of complex *a* that arises auto peak (F_1, *a*; F_2, *a*)
M_z^{bb}: z-magnetization of complex *b* that arises auto peak (F_1, *b*; F_2, *b*)
M_z^{ab}: z-magnetization that arises exchange cross peak (F_1, *a*; F_2, *b*)
M_z^{ba}: z-magnetization that arises exchange cross peak (F_1, *b*; F_2, *a*)
$R_1^{\text{complex } a}$: ^{15}N longitudinal relaxation rates for complex *a*
$R_1^{\text{complex } b}$: ^{15}N longitudinal relaxation rate for complex *b*

Since $M_z^{ab}(0) = M_z^{ab}(0) = 0$ (see Fig. 1B) and $R_1^{\text{complex } a}$ and $R_1^{\text{complex } b}$ are virtually identical, Eq. (1-4) were fitted to 32 experimental data points (8 time-points for four peaks) ny optimizing the values of five parameters: k_{ab}^{app}, k_{ba}^{app}, $R_1^{\text{complex } a}$ $(= R_1^{\text{complex } b})$, $M_z^{aa}(0)$ and $M_z^{bb}(0)$.

Integration of the differential equations and all best-fitting and optimization of parameters was carried out using the FACIMSILE package, a computer program for general initial value problems (S2-S4).

2. Scheme 1

Scheme 1 consists of a simple dissociation/reassociation mechanism:

$$PD_a \underset{k_a^{on}}{\overset{k_a^{off}}{\rightleftarrows}} P+D_a \quad , \quad P+D_b \underset{k_b^{off}}{\overset{k_b^{on}}{\rightleftarrows}} PD_b$$

If Scheme 1 is the only mechanism of translocation, the time-dependence of the auto- and cross-peaks is described by the following differential equations (Eqs. 5-10):

$$\frac{dM_z^{aa}}{dt} = k_a^{on}[D_a]M_z^{Pa} - M_z^{aa}(k_a^{off} + R_1^{\text{complex } a}) \quad [5]$$

$$\frac{dM_z^{bb}}{dt} = k_b^{on}[D_b]M_z^{Pb} - M_z^{bb}(k_b^{off} + R_1^{\text{complex } b}) \quad [6]$$

$$\frac{dM_z^{ab}}{dt} = k_b^{on}[D_b]M_z^{Pa} - M_z^{ab}(k_b^{off} + R_1^{\text{complex } b}) \quad [7]$$

$$\frac{dM_z^{ba}}{dt} = k_a^{on}[D_a]M_z^{Pb} - M_z^{ba}(k_a^{off} + R_1^{\text{complex } a}) \quad [8]$$

$$\frac{dM_z^{P,a}}{dt} = k_a^{off}M_z^{aa} + k_b^{off}M_z^{ab} - M_z^{P,a}(k_a^{on}[D_a] + k_b^{on}[D_b] + R_1^{\text{free}}) \quad [9]$$

$$\frac{dM_z^{Pb}}{dt} = k_b^{off}M_z^{bb} + k_a^{off}M_z^{ba} - M_z^{Pb}(k_a^{on}[D_a] + k_b^{on}[D_b] + R_1^{\text{free}}) \quad [10]$$

M_z^{Pa}: z-magnetization of free protein transferred from complex *a*
M_z^{Pb}: z-magnetization of free protein transferred from complex *b*
$[D_a]$: concentration of free DNA *a*
$[D_b]$: concentration of free DNA *b*
R_1^{free}: ^{15}N longitudinal relaxation rate for free protein

Fig. 2C was obtained by integration of Eq.s 5-10 within initial conditions $M_z^{ab}(0) = M_z^{ba}(0) = M_z^{Pa}(0) = M_z^{Pb}(0) = 0$, R_1^{complex} = 1.45 s^{-1} (experimental), and R_1^{free} = 2.0 s^{-1} (estimated). Under conditions where $k^{off} \ll k^{on}[D]$ (which is the case for the present study), the time-courses of the auto- and exchange peaks simulated using Eqs. 5-10 at a single DNA concentration are indistinguishable from those simulated using Eq. 1-4 with $k_{ab}^{app} = k_a^{off}/2$, $k_{ab}^{app} = k_a^{off}/2$. However, Scheme 1 cannot account for the linear concentration dependence of k_{ab}^{app} and k_{ba}^{app}, since the first order rate constants k_a^{off} and k_b^{off} in Scheme 1 are independent of DNA concentration. Note that when $k^{off} \geq k^{on}[D]$, the time-dependence of the auto- and cross-peaks described by Eqs. 5-9 is different from that resulting from Eqs. 1-4.

3. Scheme 2

Scheme 2 consists of reversible direct transfer of the protein from the specific target site on DNA *a* to the specific target site on DNA *b* via collision of free DNA with the protein-DNA complex:

$$\mathrm{PD_a + D_b} \underset{k_{ba}^{it}}{\overset{k_{ab}^{it}}{\rightleftarrows}} \mathrm{PD_b + D_a}$$

The time-dependence of the auto- and cross-peaks is described by the following differential equations:

$$\frac{dM_z^{aa}}{dt} = k_{ba}^{it}[D_a]M_z^{ab} - M_z^{aa}(k_{ab}^{it}[D_b] + R_1^{\text{complex a}}) \qquad [11]$$

$$\frac{dM_z^{bb}}{dt} = k_{ab}^{it}[D_b]M_z^{ba} - M_z^{bb}(k_{ba}^{it}[D_a] + R_1^{\text{complex b}}) \qquad [12]$$

$$\frac{dM_z^{ab}}{dt} = k_{ab}^{it}[D_b]M_z^{aa} - M_z^{ab}(k_{ba}^{it}[D_a] + R_1^{\text{complex b}}) \qquad [13]$$

$$\frac{dM_z^{ba}}{dt} = k_{ba}^{it}[D_a]M_z^{bb} - M_z^{ba}(k_{ab}^{it}[D_b] + R_1^{\text{complex b}}) \qquad [14]$$

The time-courses described by Eqs. 11-14 are indentical to those given by Eq 1-4 when $k_{ab}^{app} = k_{ab}^{it}[D_b]$ and $k_{ba}^{app} = k_{ba}^{it}[D_a]$. All the chemical exchange data obtained at five different DNA concentrations (total of 160 data points) was fitted simultaneously, optimizing the values of k_{ab}^{it}, k_{ba}^{it} and $R_1^{\text{complex } a}$ $(= R_1^{\text{complex } b})$, as well as the values of $M_z^{aa}(0)$ and $M_z^{bb}(0)$ at each DNA concentration.

4. Combined Schemes 1 and 2

If both Schemes 1 and 2 contribute the translocation process, the Eqs. 5-8 are replaced by:

$$\frac{dM_z^{aa}}{dt} = k_a^{on}[D_a]M_z^{P,a} + k_{ba}^{it}[D_a]M_z^{ab} - M_z^{aa}(k_a^{off} + k_{ab}^{it}[D_b] + R_1^{\text{complex a}}) \qquad [15]$$

$$\frac{dM_z^{bb}}{dt} = k_b^{on}[D_b]M_z^{P,b} + k_{ab}^{it}[D_b]M_z^{ba} - M_z^{bb}(k_b^{off} + k_{ba}^{it}[D_a] + R_1^{\text{complex b}}) \qquad [16]$$

$$\frac{dM_z^{ab}}{dt} = k_b^{on}[D_b]M_z^{P,a} + k_{ab}^{it}[D_b]M_z^{aa} - M_z^{ab}(k_b^{off} + k_{ba}^{it}[D_a] + R_1^{\text{complex b}}) \qquad [17]$$

$$\frac{dM_z^{ba}}{dt} = k_a^{on}[D_a]M_z^{P,b} + k_{ba}^{it}[D_a]M_z^{bb} - M_z^{ba}(k_a^{off} + k_{ab}^{it}[D_b] + R_1^{\text{complex b}}) \qquad [18].$$

and Eqs. 9-10 remain unchanged. Fitting all the exchange data to this combined scheme indicates that Scheme 1 does not make a significant contribution to the translocation process (see main text).

5. Practical Consideration

In this paper we make use of simultaneous fitting of four cross-peaks (two auto-peaks and two exchange cross-peaks) to investigate the kinetics of protein-DNA interactions. From a practical standpoint the intensities of the auto-peaks at T = 0, and the maximum intensities of the cross-peaks should be significantly higher than the noise level to permit reasonably accurate peak intensity determination. The ratio of the auto-peak intensities at T=0 (or the corresponding peaks in normal HSQC spectrum) is given by:

$$\frac{M_z^{aa}(0)}{M_z^{bb}(0)} = \frac{[PD_a]}{[PD_b]} = \frac{K_{diss}^b}{K_{diss}^a} \cdot \frac{[D_a]}{[D_b]} \qquad [19]$$

where K_{diss}^a and K_{diss}^b are the equilibrium dissociation constants for complexes *a* and *b*, respectively; and $[D_a]$ and $[D_b]$ are the free concentrations of the corresponding two DNA oligonucleotides. It is also desirable that the intensities of the two auto-peaks at T = 0 are approximately comparable. If these intensities are very different under conditions comprising equimolar amounts of the two DNA species, one can conclude that the equilibrium dissociation constants for the two species are very different. Even for such a case, however, it is still possible to make the intensity ratio of the auto-peaks at T = 0 close to 1 by appropriate adjustment of the molar ratio between protein, DNA *a* and DNA *b*. For example, consider the case where $K_{diss}^a << K_{diss}^b$(i.e. the protein binds to DNA *a* with much higher affinity) and $[P]_{total} >> K_{diss}^b$ (where *P* is protein). If the three components are mixed at a ratio $[P]_{total} : [D_a]_{total} : [D_b]_{total}$ = 1 : 0.5 : 2, the concentrations of the two complexes will be approximately equal to $0.5[P]_{total}$ and the ratio in Eq. 19 will be close to 1. (Note that $[D_a]$will be very small in this case). Thus, providing the translocation process between complexes *a* and *b* is in the slow exchange regime, it is a simple matter to adjust the experimental conditions such that the kinetics of protein-DNA interactions can be analyzed using the approach described in this paper.

6. Sample preparation

DNA encoding the human Hox-D9 homeodomain (hox-D9 residues 275-334) was obtained by PCR using chemically synthesized DNA fragments and inserted into the pGEX-4T-1vector (Amersham). For sample stability, Cys280 was substituted by Ser and it was confirmed that this mutation does not affect the affinity sequence specific DNA binding (data not shown). $^{13}C/^{15}N$-labeled or $^{2}H/^{15}N$-labeled proteins with a GST-tag were expressed in *E. coli* strain BL21 grown in 1:1 mixture of minimal M9 and Silantes OD2 media ($^{13}C/^{15}N$ or $^{2}H/^{15}N$). The GST-fusion protein was purified using a GST-Fast Flow column and cleaved with thrombin. The resultant homeodomain with two extra residues (Gly-Ser) at the N-terminus was purified by Superdex-75 gel-filtration and Mono-S cation exchange chromatography. All purification procedures were carried out at an ionic strength greater than 300 mM, since the solubility of the free Hox-D9 homeodomain (and also its fusion product with GST) is much reduced at lower ionic strength. Numbering for the Hox-D9 residues within the homeodomain in the main text is in accordance with the conventional numbering scheme employed for homeodomains (S5).

Single-stranded DNA oligonucleotides used to prepare the duplex DNA oligonucleotides *a* and *b* (Fig. 1A) were synthesized using solid-phase phosphoreamidite chemistry (Midland Certified Reagent Inc.). The DNA strands were

purified by Mono-Q anion exchange chromatography. Equimolar complementary strands were mixed and annealed to obtain duplexes. The DNA duplexes for NMR study were further purified by Mono-Q chromatography to remove any excess single-strand DNA. The rhodamine-conjugated DNA used for the affinity measurements by fluorescence (see below) was purified by native polyacrylamide gel electrophoresis.

7. K_{diss} measurement

The equilibrium dissociation constant K_{diss} for the complex between DNA *a* and the Hox-D9 homeodomain was determined by fluorescence anisotropy measurements using Jobin Ybon FluoroMax-3 spectrometer. Fluorescence anisotropy (λ_{ex} = 550 nm, λ_{em} = 580 nm), measured at 25 °C on rhodamine conjugated at the 5'-terminus of one of the strands of DNA *a*, was measured upon titrating the Hox-D9 homeodomain (0.2–60 nM) to a solution of 1 nM rhodamine-DNA *a*, 10 mM Tris·HCl (pH7.2), 10% Glycerol and 100 mM NaCl. The value of K_{diss} was calculated from the data as described previously (S6).

8. NMR experiments

The Hox-D9 homeodomain-DNA complex for NMR spectroscopy was made by initially mixing protein with a slight excess DNA in a high ionic-strength buffer (10 mM Tris·HCl pH 6.8 and 500 mM NaCl). The salt concentration was then gradually reduced to 40 mM NaCl using Amicon Ultra-4 spin concentrators. The final protein and DNA concentration were determined using the BCA assay (Pierce) and UV absorbance at 260 nm, respectively.

NMR experiments were carried out at 35 °C using Bruker DMX-500 and DRX-600 spectrometers equipped with cryogenic triple resonance *z*-gradient probes. Protein backbone ^{1}H, ^{13}C, ^{15}N resonances were assigned using 3D-HNCO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HBHA(CO)NH, C(CO)NH, ^{1}H-^{15}N NOESY-HSQC and ^{1}H-^{15}N HMQC-NOESY-HSQC spectra recorded on a sample of 0.5 mM $^{13}C/^{15}N$-labeled Hox-D9 homeodomain-DNA complex (complex *a* in Fig. 1). The ^{1}H-^{15}N correlation experiments for measurement of the exchange rates (S7) were carried out at ^{1}H-frequency of 500 MHz using an approximately 1:1 mixture of $^{2}H/^{15}N$-labeled (HOX-D9) complexes *a* and *b*. The initial sample conditions were 0.68 mM protein, 0.44 mM DNA *a* and 0.44 mM DNA *b* in a buffer of 10 mM Tris·HCl (pH6.8), 40 mM NaCl and 7% D_2O. (i.e. This corresponds to ~0.34 mM complex *a*, ~0.34 mM complex *b*, ~0.1 mM free DNA *a* and ~0.1 mM free DNA *b*). Small volumes of a 1.9 mM stock DNA solution containing free DNA *a* and *b* (dissolved in the same buffer) were added to increase the free DNA concentration as shown in Fig. 2A. Eight different mixing times were used: 53.9, 77.9, 108.9, 137.9, 167.9, 287.9, 467.9, and 647.9 ms.

References for Supporting Information

(S1) McConnell, H. M. *J. Chem. Phys.* **1958**, *28*, 430.
(S2) Chance E. M.; Curtis, A. R.; Jones, I. P.; Kirby, C. R. U.K. *At. Energy. Auth., Harwell Lab.* **1979**, *AERE-R* 8775, HM Stationary Office, London, U.K.
(S3) Curtis, A. R. U.K. *At. Energy. Auth., Harwell Lab.* **1979**, *AERE-R* 9352, HM Stationary Office, London, U.K.
(S4) Clore, G. M. In *Computing in Biological Science*; Geisow, M. J., Barrett, M., Eds.; Elsevier North-Holland: Amsterdam, 1983; pp 313-348.
(S5) Branden, C.; Tooze, J. In *Introduction to protein structure* 2nd Ed.; Garland Publishing, New York, 1998.
(S6) Iwahara, J; Schwieters, C.D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 12800-12808.
(S7) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *J. Biomol. NMR* **1994**, *4*, 727-734.