Supplementary Information

EDTA-Derivatized Deoxythymidine as a Tool for Rapid Determination of Protein Binding Polarity to DNA by Intermolecular Paramagnetic Relaxation Enhancement

Junji Iwahara, D. Eric Anderson, Elizabeth C. Murphy and G. Marius Clore*

Laboratories of Chemical Physics and Bioorganic Chemistry, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0510

1. MASS SPECTROMETRY

Chromatography

Samples of duplex DNA (oligo1 shown in Fig. 1b of main paper) containing one non-modified strand (5'd-GGACGTGTTTGTGGAC) and one strand in which the dT-EDTA (T*) was incorporated (5'd-CCTGCACAAACACCT*G) were incubated for at least 7 days at room temperature prior to analysis using HPLC/electrospray mass spectrometry. Chromatography was performed using an Agilent HP1100 stack including a solvent degasser, binary pump, auto-sampler, variable temperature column compartment, and a diode array detector. An Agilent Zorbax 300SB-C8 4.6 x 150 mm RP column was maintained at 60°C. A hexafluoro-isopropanol/triethylamine/methanol solvent system was used (Apffel A; Chakel, J.A.; Fischer, S.; Lichtenwalter, K.; Hancock, W.S. *Journal of Chromatograhy A* **1997**, 3-21). Injection of approximately 20 picomoles of duplex conjugate onto the column was followed by a 60 minute wash with solvent A (400mM 1,1,1,3,3,3-hexafluoro-isopropanol/adjusted to pH 6.8 with triethylamine). Different gradients were used to elute the oligonucleotides. In the series reported here a low-resolution elution was used which involved a 15 minute gradient from 0 to 80% solvent B (solvent A plus 50% methanol). The set of experiments performed on the duplex DNA used in subsequent NMR studies were performed with and without dithiotreitol (DTT) in the absence of metal ions and in the presence of chelated Ca^{2+}, Mn^{2+}, and Fe^{3+}.

Mass Spectrometry

After passing through the UV detector the sample was introduced into an HP MSD1946A electrospray mass spectrometer operating in negative mode using a parameter set optimized for the experiment using a standard sample. Importantly the capillary potential was lowered to 2.5 kV from default values to reduce the observation of chemical noise in the mass chromatographs.

Results

Except for one case (Fe^{2+} + DTT), these experiments yielded similar overall results. The results for the Ca^{2+}/DTT experiment (Fig. S1.1) are representative of this set of similar results. In panels A and B below are the UV traces at 280 nm and 260 nm respectively. Immediately below these is the trace of the total ion current from the mass spectrometer. Note that partial resolution of the two peaks is achieved and that a delay of about 12 seconds and band broadening appear to take place between the UV detector and the mass spectrometer. The two dominant species present in these samples eluted independently. Panel D shows the elution of a species observed at 619.8 Th which originates from the non-derivatized oligonucleotide while Panel E shows the elution of a species observed at 649.5 Th which originates from the derivatized species in which two protons have been replaced by a divalent calcium ion.
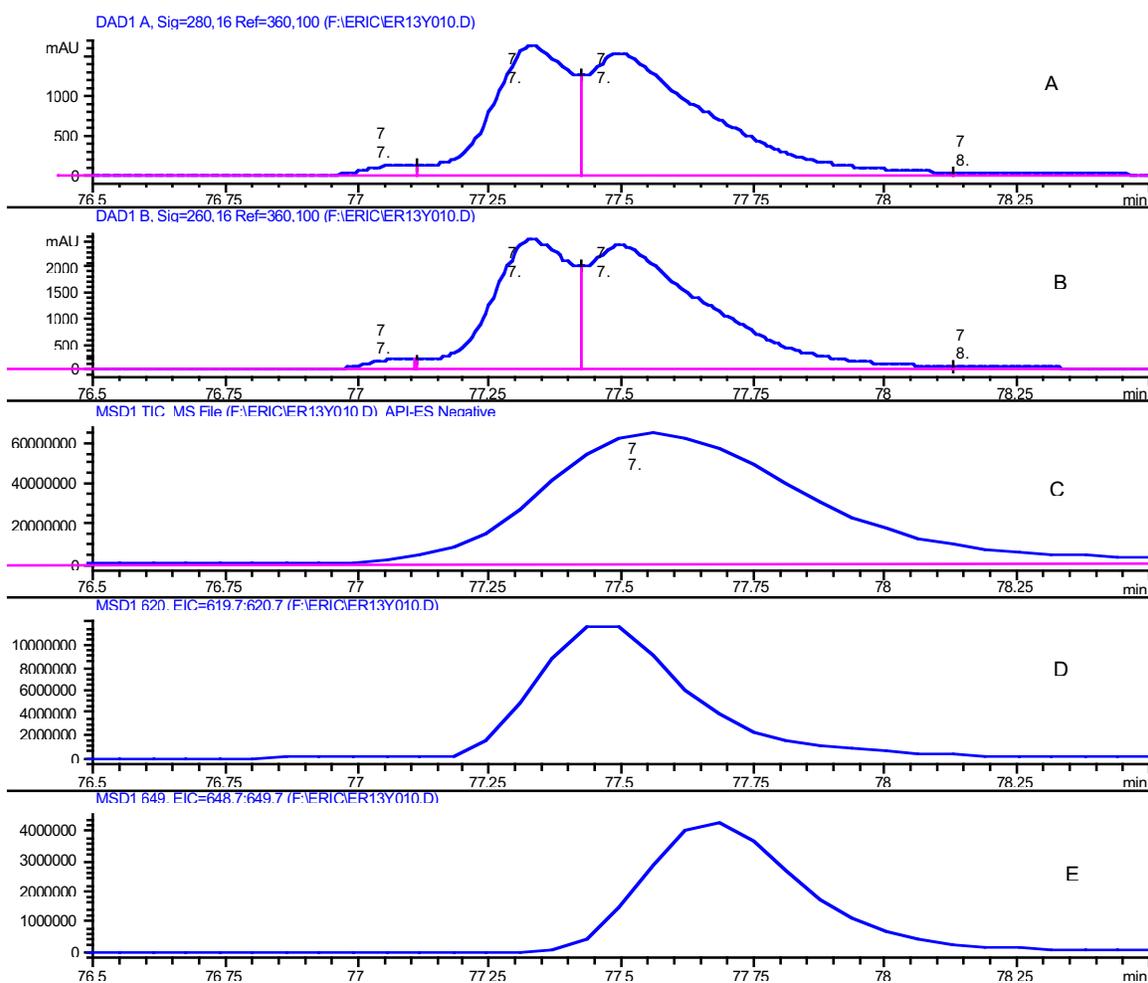
Ca^{2+} + DTT

Figure S1.1

The calculated mass for the underivatized oligonucleotide was 4965.85 with measured mass corresponding closely at 4966.6. The calculated mass of the derivatized species is 5167.6 with a measured mass of 5166.1.

As expected there was a significant difference observed for the sample containing Fe^{2+} and DTT. A series organized in a similar way to that shown above for the Ca^{2+} sample (panels A-E) is displayed below (Fig. S1.2).
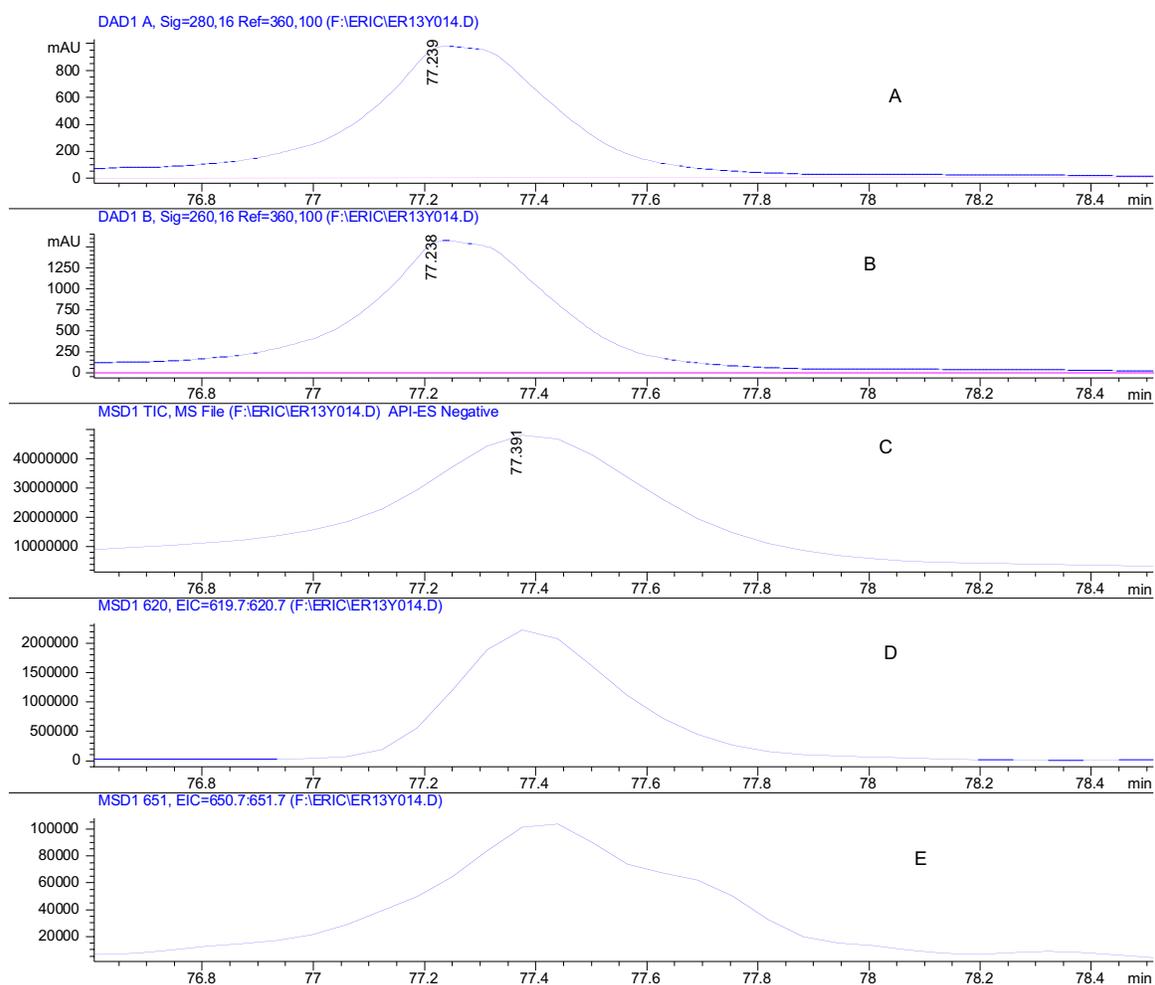
Fe $^{2+}$ + DTT

Figure S1.2

Several important differences between the Fe^{2+}/DTT and Ca^{2+}/DTT data sets can be observed in these panels (Fig. S1.2). First there is only one peak in the absorbance trace. In addition, there is a non-zero absorbance and total ion current (TIC) indicated very early in the chromatogram. There remains an apparent peak from the smaller oligonucleotide though the peak height may be less than observed before. The co-added mass spectra from 77.2 to 77.7 minutes for this species is shown below (Fig. S1.3, top) and reveals that there remains a discernable series of peaks from the smaller oligonucleotide but that there appears to be a large number of other species present. Contrast this with the co-added spectrum from the half-minute under the same species peak from the Ca^{2+}/DTT experiment shown just below this spectrum (Fig. S1.3, bottom). We interpret this difference to originate from partial degradation of the smaller oligonucleotide and presumably nearly complete degradation of the larger oligonucleotide. Indeed the signal shown in panel E above probably does not originate from the Fe^{2+} complexed oligonucleotide but is simply isobaric with some earlier eluting degradation product of one or both of the oligonucleotides.

A detail from the Fe^{2+}/DTT experiment
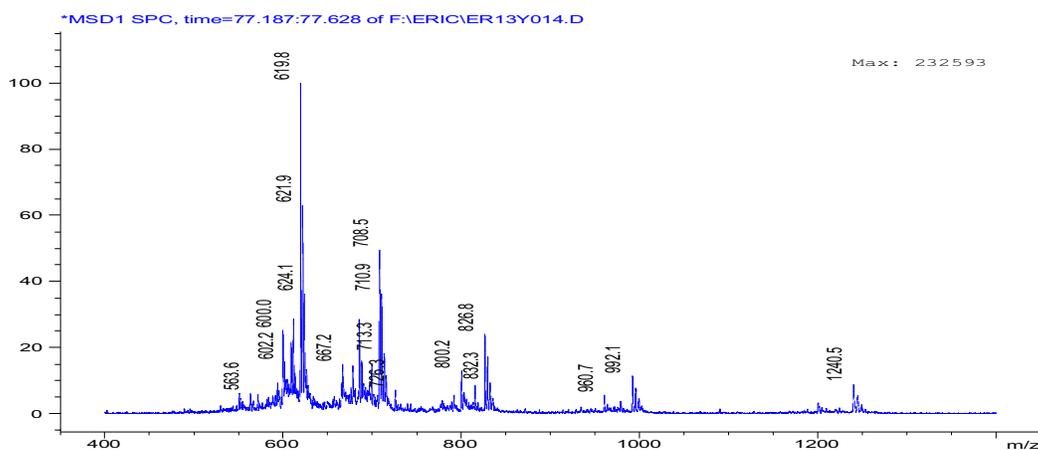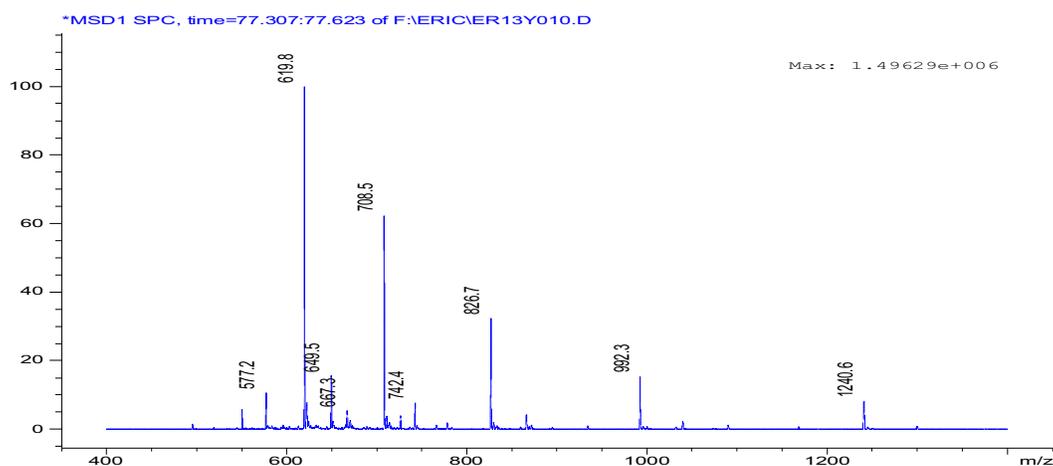

A detail from the Ca^{2+}/DTT experiment

Figure S1.3

Below (Fig. S1.4) are each of the three channel chromatographs from the experiments for "no metal" and Mn^{2+} together with DTT in the reaction shown on a similar scale to that of the panels shown above. Importantly there is similar intensity observed in the optical trace of the non-metal control supporting the interpretation that little or no degradation is occurring in other non-iron containing samples.

No Metal + DTT
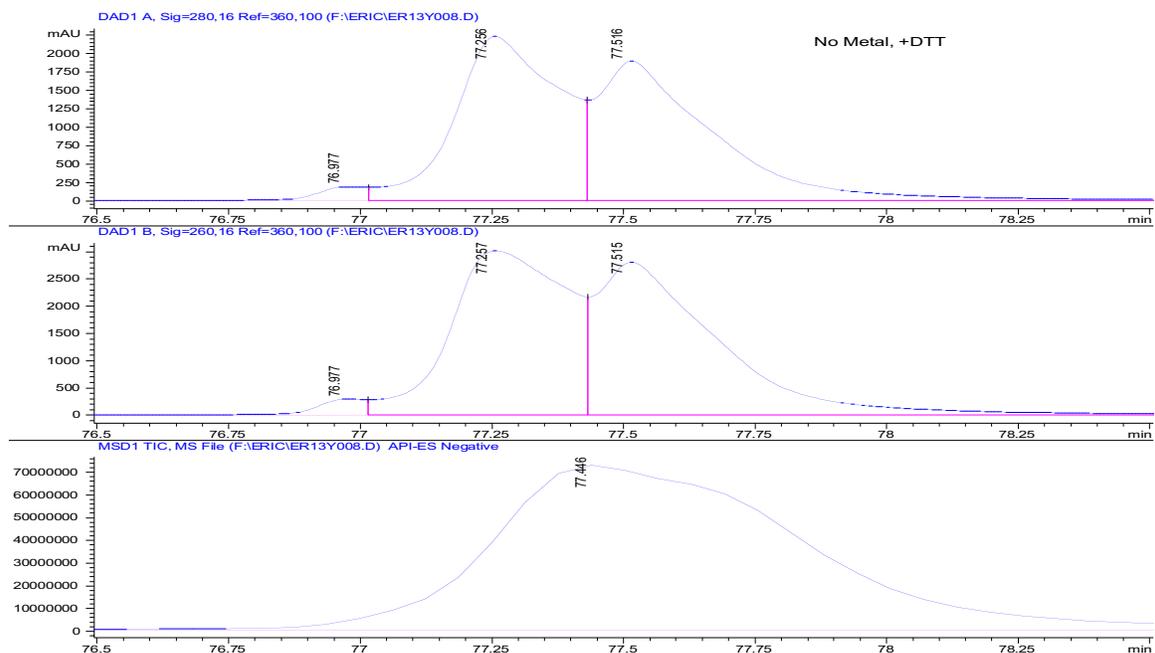
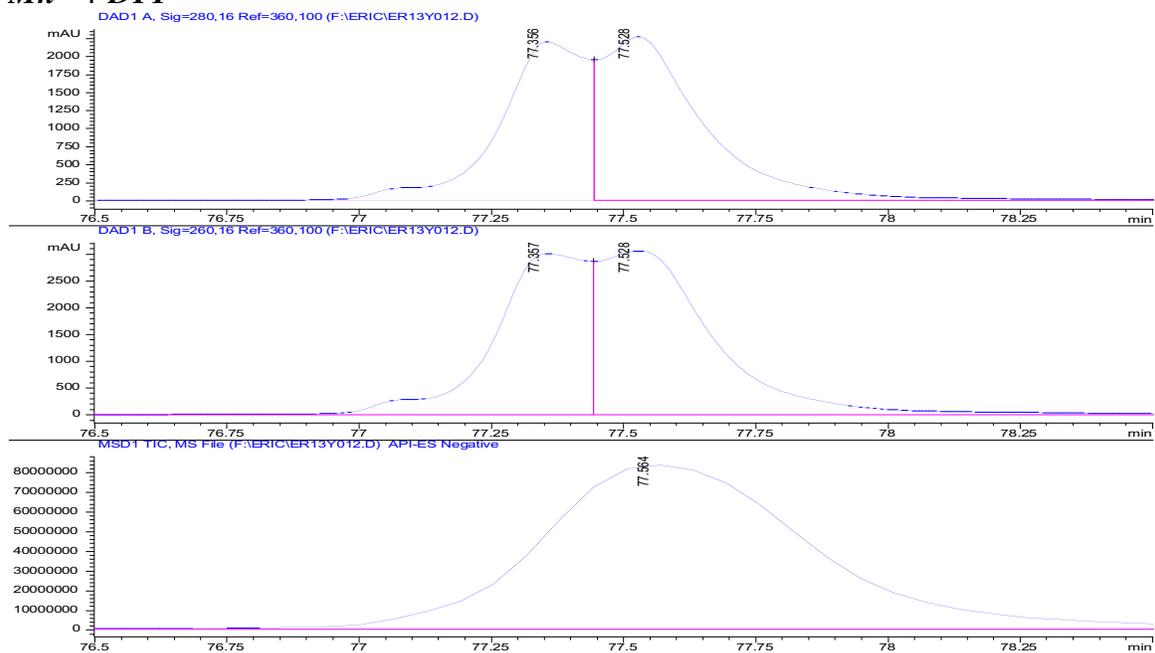## *Mn^{2+} + DTT*

Figure. S1.4

A summary of the mass spectrometry results is provided in Table S.1

Table S.1 Electrospray mass spectrometry analysis[a] of the dT-EDTA derivatized 16 base pair DNA oligonucleotide (oligo1) chelated to various metal ions in the presence and absence of the reducing agent DTT.

Strand 1: 5'd-CCTGCACAAACACCT*G   Strand 2: 5'd-GGACGTGTTTGTGGAC

T* is dT-EDTA

Metal	DTT	Low Mass (strand 2)[b]	High Mass (strand 1)	Predicted Mass strand 1[c]
None	No	4966.64	5166.06	5167.6
None	Yes	4966.67	5166.11	5167.6
Ca^{2+}	No	4966.65	5204.17	5205.7
Ca^{2+}	Yes	4966.60	5204.22	5205.7
Mn^{2+}	No	4966.63	5219.02	5220.5
Mn^{2+}	Yes	4966.63	5218.83	5220.5
Fe^{3+} [d]	No	4966.63	5219.01	5220.5
Fe^{2+}	Yes	4966.56	[e]	5220.5

[a]Electrospray mass spectra were measured in negative mode after HPLC separation. [b]The predicted mass of strand 2 is 4965.9. [c]The predicted mass of strand 1 is the mass of the derivatized oligonucleotide plus the mass of the metal ion minus the mass of the displaced protons (2 for Ca^{2+} and Mn^{2+}, and 3 for Fe^{3+}). [d]Added as Fe^{2+} but undergoes spontaneous air oxidation to Fe^{3+}. [e]Mass spectra and chromatography suggest some remaining 4966.56 material, and numerous cleaved species.

2. NMR SPECTROSCOPY

2. 1 Pulse sequences employed for the measurement of amide proton T_2 and T_1 values
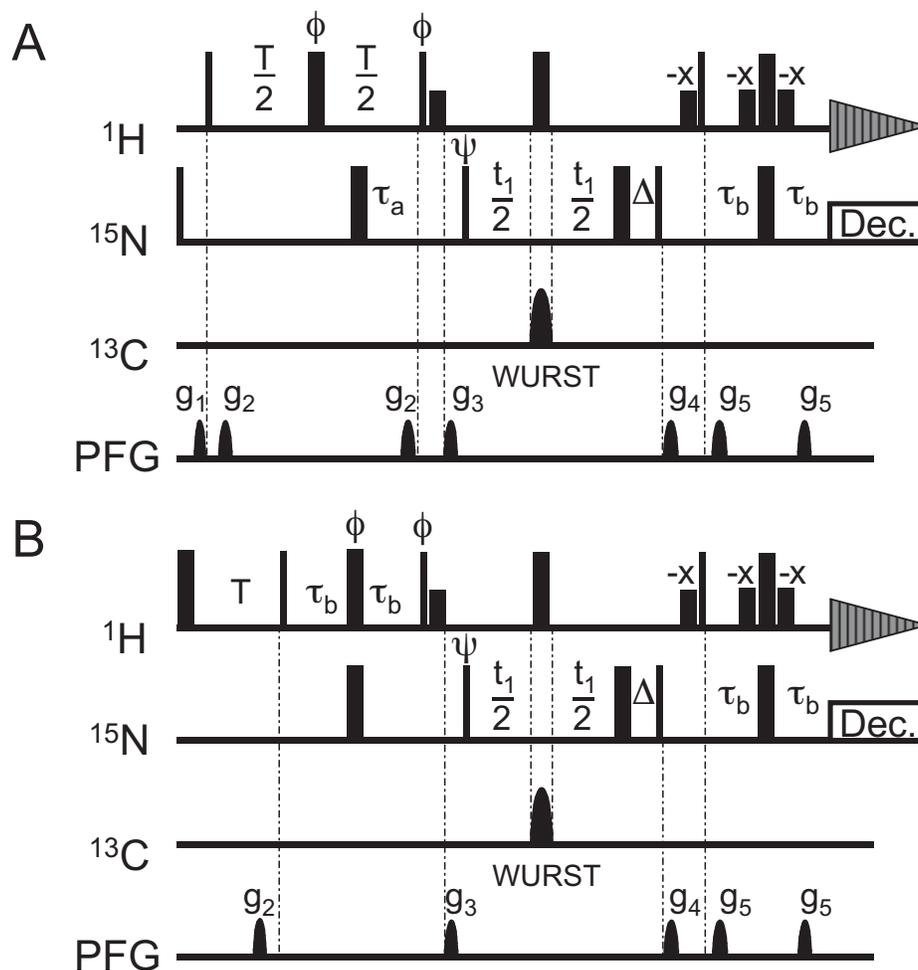

Figure S2.1 (A) Pulse sequence for measurement of amide proton T_2. Phases are along x unless indicated otherwise. Small bold ^{1}H pulses represent water-selective rectangular 90° pulses (1.4 ms). The ^{13}C-WURST pulse (500 μs) is for ^{13}Cα- and ^{13}CO- decoupling. Delays τ_a, τ_b, and Δ are 2.75 ms, 2.25 ms, and 500 μs (length of the ^{13}C WURST pulse), respectively. The phase cycles are as follows: $\psi = $ (x, -x), $\phi = $ (y, y, -y, -y), receiver = (x, -x, -x, x). Phase ψ was incremented in the States-TPPI manner for quadrature detection in the F1-dimension. The gradient lengths and strengths are as follows: g_1 (0.6 ms, 10 G/cm), g_2 (0.6 ms, 17 G/cm), g_3 (1.8 ms, 17 G/cm), g_4 (1.4 ms, -27 G/cm), g_5 (0.5 ms, 13 G/cm). The peak intensity modulation can be fitted to $I = I_0\exp(-T/T_2)\cos{(\pi\,^3J_{HN\text{-}H\alpha}\,T)}$. The effect of $^3J_{HN\text{-}H\alpha}$ on Γ_2 is negligible, if the protein is α-helical (as is the case for SRY) and $T_{max} \ll 1/2\,^3J_{HN\text{-}H\alpha}$. Calculations indicate that even for $^3J_{HN\text{-}H\alpha} = 10$ Hz, the value of ^1H_N-Γ_2 is only underestimated by 3-4%. It is also possible to complete eliminate $^3J_{HN\text{-}H\alpha}$ modulation

by employing a H_N-selective 180° pulse (*e.g.* REBURP) at the center of T. (B) Pulse sequence for measurement of amide proton T_1. Experimental conditions are as described for (A). The T-delay in sequence (B) should be sufficiently long (e.g. T ≥ 100 ms) to let water magnetization return to the +z axis through radiation damping (Talluri,S.; Wagner, G. *J.Magn.Reson. Series B*, **1996**, *112*, 200-205). The peak intensity modulation was fitted to $I = I_0\{ 1- ρ \exp(-T/T_1)\}$, where ρ is related to a recovery level during a repetition delay. Since ^1H_N-T_1 values obtained in this manner are affected by cross-relaxation and water-H_N exchange, ^1H_N-$Γ_1$ values were used only for determination of the correlation times for Mn^{2+}-^1H_N interaction vectors as described in Section 3 (Eq. 3.1).

2.2 Quantitative ^1H_N- $Γ_2$ measurements.

Quantitative measurements were carried out on the SRY-oligo1 and SRY-oligo2 complexes (Fig. S2.2).

Oligo1
`5'd-CCTGCACAAACACC`**`T`**`G`
`3'd-GGACGTGTTTGTGGAC`

Oligo2
`5'd-CACCTGCACAAACACC`
`3'd-G`**`T`**`GGACGTGTTTGTGG`

dT-EDTA is depicted in red for oligo1 and green for oligo2.
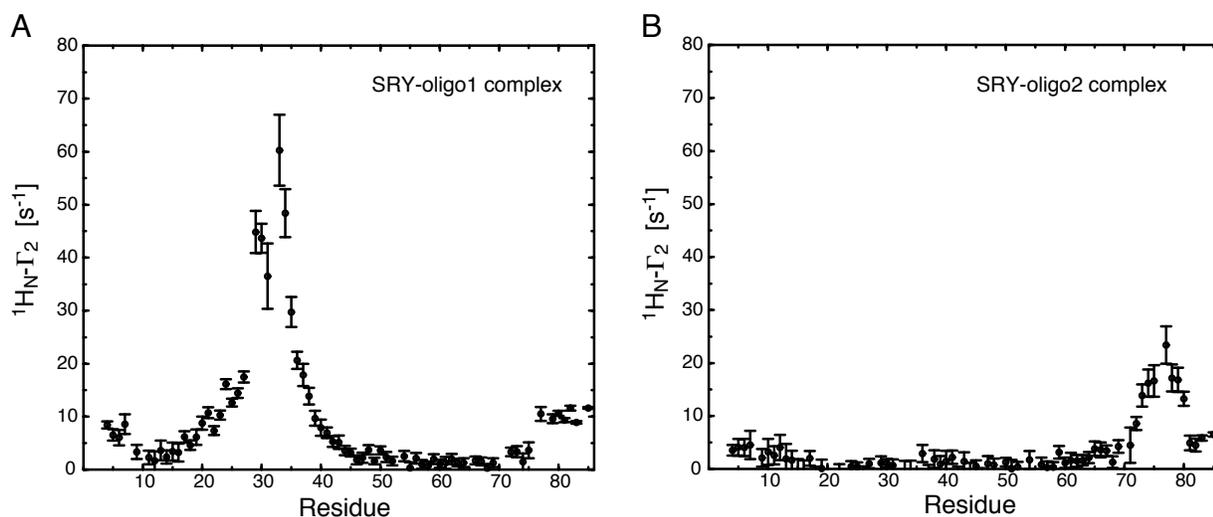

Figure S2.2 Protein backbone ^1H_N- $Γ_2$ measured on 0.3 mM SRY/DNA-EDTA complexes (**A**, oligo1; **B**, oligo2) in 20 mM Tris•HCl (pH 6.8), 20 mM NaCl and 93% H_2O/7% D_2O. Mn^{2+}- and Ca^{2+}-chelated complexes were prepared and values of H_N-$Γ_2$ are determined as $1/T_2^{paramagnetic}$ − $1/T_2^{diamagnetic}$. To remove divalent ions at sites other than the attached EDTA, complexes at relatively low concentration (~30 μM) were extensively washed with a high ionic strength buffer (20 mM Tris•HCl, pH 6.8, 0.5 M NaCl). The relaxation rates of Mn^{2+}- and Ca^{2+}-bound states were measured at 35°C with a cryogenic probe on a Bruker

DMX-500 spectrometer. Eight 2D ^{1}H-^{15}N spectra with ^{1}H T_2-relaxation delays of 0.3, 2, 4, 8, 12, 16, 20, and 24 ms were analyzed to determine relaxation rates. Errors were estimated based on noise levels of the spectra.

3. STRUCTURE REFINEMENT

Structures were refined by simulated annealing and conjugate gradient minimization using the molecular structure determination package Xplor-NIH (Schwieters, C. D.; Kuszewski, J.; Tjandra, N.; Clore, G.M. *J. Magn. Reson.* **2003**, *160*, 65-73). Two base pairs (b.p.) including the dT-EDTA-Mn^{2+} moiety we added in a B-DNA geometry to each end of the structure of the original 14 b.p. DNA in the SRY/DNA complex (Murphy, E.C.; Zhurkin, V. B.; Louis, J. M.; Cornilescu, G.; Clore, G. M. *J. Mol. Biol.* **2001**, *312*, 481-499). The resulting oligonucleotide of 18 b.p. enables one to combine the Γ_2 measurements on the SRY/oligo1 and SRY/oligo2 complexes. The coordinates of the last two b.p. at either end including the dT-EDTA-Mn^{2+} moieties were refined by simulated annealing in torsion angle space (Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *152*, 288-302) against a target function comprising the ^{1}H$_N$-Γ_2 restraints (Donaldson, L. W.; Skrynykov, D. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843-9847), B-DNA torsion angle restraints for the sugar-phosphate backbone, and non-bonding terms comprising a van der Waals repulsion potential and conformational database torsion angle (Clore, G. M.; Kuszweski, J. *J. Am. Chem. Soc.* **2003**, *125*, 1518-1525) and DNA base-base positional (Kuszewski, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2001**, *123*, 3903-3918) potentials of mean force; the protein and the central 14 b.p. corresponding to the original structure were held fixed. Subsequently, all coordinates were subjected to conjugate gradient minimization in cartesian coordinate space against all the experimental restraints [i.e. interproton distances, torsion angles, ^{13}Cα/β shifts, dipolar couplings used in the original structure determination (Murphy, E.C.; Zhurkin, V. B.; Louis, J. M.; Cornilescu, G.; Clore, G. M. *J. Mol. Biol.* **2001**, *312*, 481-499), together with the ^{1}H$_N$-Γ_2 restraints]; bond, angle and improper torsion restraints to minimize deviations from idealized covalent geometry; and the three non-bonding interaction terms listed above.

The global value of τ_c, defined as $(1/\tau_r + 1/\tau_e)^{-1}$ (where τ_r is the rotational correlation time, and τ_e the electronic relaxation correlation time) is optimized during refinement within a pre-specified range restraints (Donaldson, L. W.; Skrynykov, D. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843-9847). The range of τ_c can be readily estimated from the experimentally measured ^{1}H$_N$-Γ_1 and ^{1}H$_N$-Γ_2 values and the equation

$$\tau_c = (1.5\ \Gamma_2/\Gamma_1 - 1.75)^{1/2}/\omega_H \qquad\qquad (3.1)$$

The average value of τ_c for the protein core (using residues with ^{1}H$_N$-$\Gamma_2 \geq 5$ s^{-1} and ^{1}H$_N$-$\Gamma_1 \geq 0.2$ s^{-1}) is 2.9±0.4 ns. For the N- and C-terminal regions (residues 1-6 and 80-85) of SRY, however, the residue specific τ_c values were much lower ranging from 1-1.5 ns, indicative

of internal motion. Consequently, only the ^1H_N-Γ_2 data for residues 7-79 were employed in the refinement calculations.

20 structures were calculated, and a superposition of the structures is shown in Fig. S3.1. All the structures satisfy the experimental restraints within their experimental error (Table S3.1). The atomic rms shift between the regularized mean coordinates of the present structure and the previous structure determined without ^1H_N-Γ_2 data is 0.25 Å for all heavy atoms of the protein (residues 7-79) and the central 14 b.p. of the DNA, and 0.16 Å for the backbone atoms of the proten (residues 7-79) and the central 14 b.p. of the DNA (Fig. S3.2). These atomic r.m.s. shifts are within the errors of the coordinates.
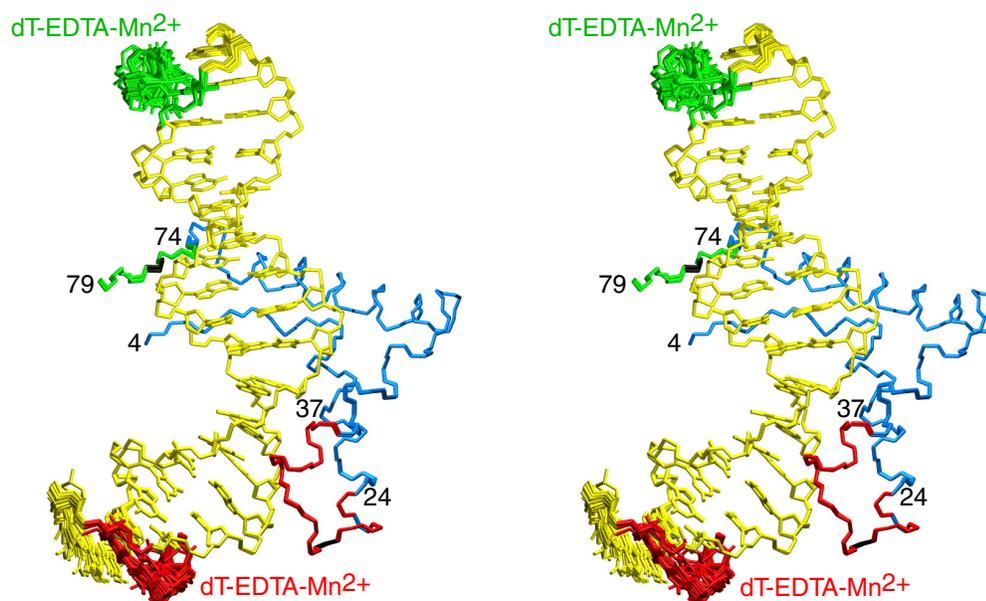

Figure S3.1 Superposition of refined structures (best-fitted to residues 7-79 of SRY and the central 14 b.p. of the DNA). The backbone of residues with ^1H_N-$\Gamma_2 \geq 15$ s^{-1} for the SRY-oligo1 complex are displayed in red and for the SRY-oligo2 complex in green, with intervening prolines in black. The dT-EDTA moieties from oligo1 and oligo2 are shown in red and green, respectively.
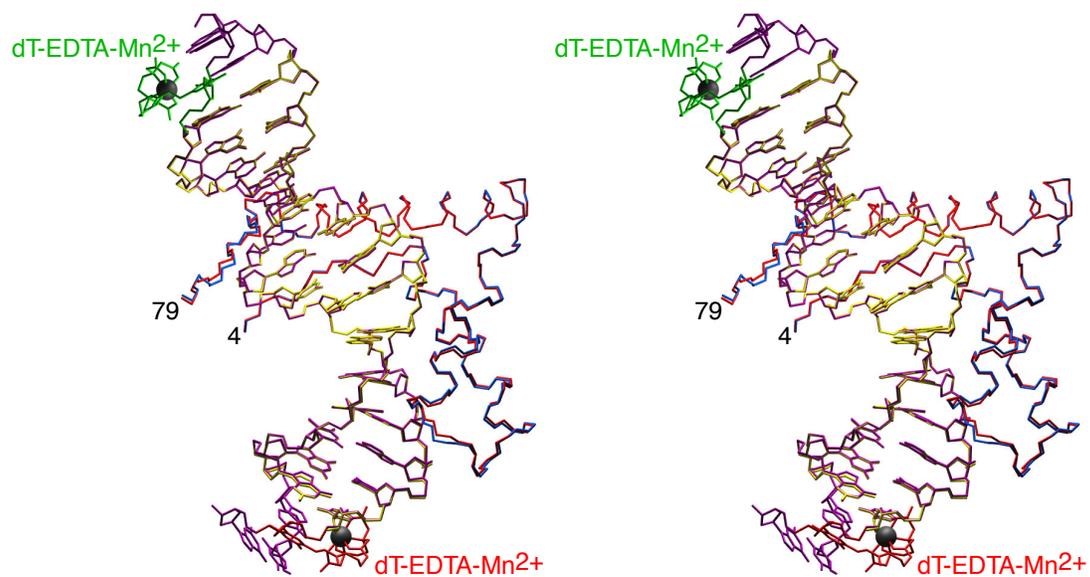
Figure S3.2 Superposition of the restrained regularized mean structures obtained with (SRY in red, DNA in purple, dT-EDTA from oligo1 and oligo2 in red and green, respectively, and Mn^{2+} ions in blue) and without (SRY in blue and DNA in yellow; PDB accession code 1J46) refinement against the ^1H_N-Γ_2 data.

Table S3.1 Structural statistics of the restrained regularized mean structures[a]

	Refined without ^1H_N-Γ_2 data[b]	Refined with ^1H_N-Γ_2 data
Agreement with ^1H_N-Γ_2 data (130)		
r.m.s. deviation (s^{-1})	-	2.0
Q-factor (%)	-	17.0
R.m.s. deviations from experimental restraints		
Distance restraints (Å) (1755)	0.04	0.04
Torsion angle restraints (°) (433)	0.34	0.68
$^3J_{HN\alpha}$ coupling restraints (Hz) (70)	0.84	0.98
$^{13}C\alpha$/$^{13}C\beta$ shift restraints (ppm) (165)	0.99	1.05
DNA D_{HH} dipolar couplings (Hz) (55)	0.56	0.70
Fixed distance heteronuclear dipolar coupling R-factors (%)		
protein $^1D_{NH}$ (71)	5.5	6.1
protein $^1D_{CH}$ (67)	6.3	6.1
protein $^1D_{NC'}$ (68)	18.9	22.9
protein $^2D_{HNC'}$ (68)	18.8	21.9
DNA $^1D_{NH}$ (9)	10.2	11.2
DNA $^1D_{CH}$ (37)	11.2	13.7
Deviations from idealized covalent geometry		
bonds (Å)	0.004	0.006
angles (°)	0.81	0.99
improper torsions (°)	0.79	0.73

[a]The numbers in parentheses indicate the number of restraints in each category
[b]The structure (PDB accession code 1J46) and experimental restraints, excluding the ^1H_N-Γ_2 data are from Murphy, E.C.; Zhurkin, V. B.; Louis, J. M.; Cornilescu, G.; Clore, G. M. *J. Mol. Biol.* **2001**, *312*, 481-499.
[c]The Q-factor for the agreement between observed and calculated ^1H_N-Γ_2 values is given by $100[\Sigma(\Gamma_2^{obs}-\Gamma_2^{calc})^2/\Sigma(\Gamma_2^{obs})^2]^{1/2}$

A comparison between observed and calculated values of ^1H_N-Γ_2 is shown in Fig. S3.3.

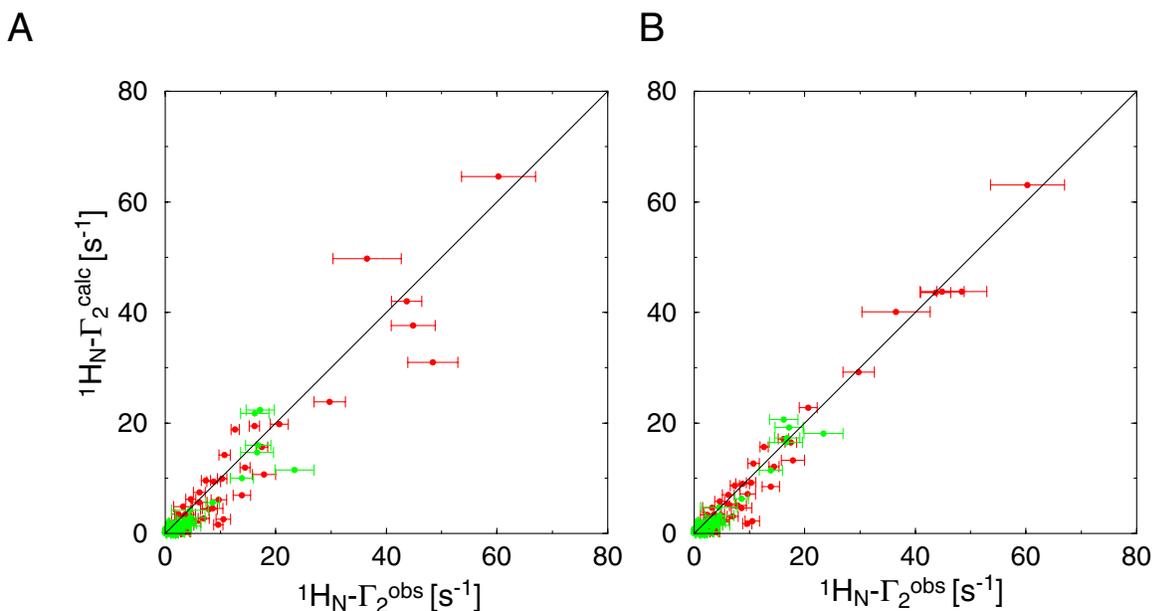**Figure S3.3** Comparison between observed and calculated values of ^1H_N-Γ_2 (A) after simulated annealing refinement of the coordinates of the dT-EDTA-Mn^{2+} moieties with the protein and central 14 b.p. of DNA held fixed, and (B) after conjugate gradient minimization of all coordinates against all experimental restraints. The Q-factor, rms difference and correlation coefficient are 28%, 3.2 s^{-1} and 0.95, respectively, in (A) and 17%, 2.0 s^{-1} and 0.98, respectively, in (B). The improvement in the agreement between observed and calculated ^1H_N-Γ_2 values is achieved with minimal atomic rms shifts which are within coordinate errors (see above). The data from the SRY-oligo1 and SRY-oligo2 complexes are shown in red and green, respectively.

4. Theoretical dependence of ^1H_N-Γ_2 values on metal ion-proton distance

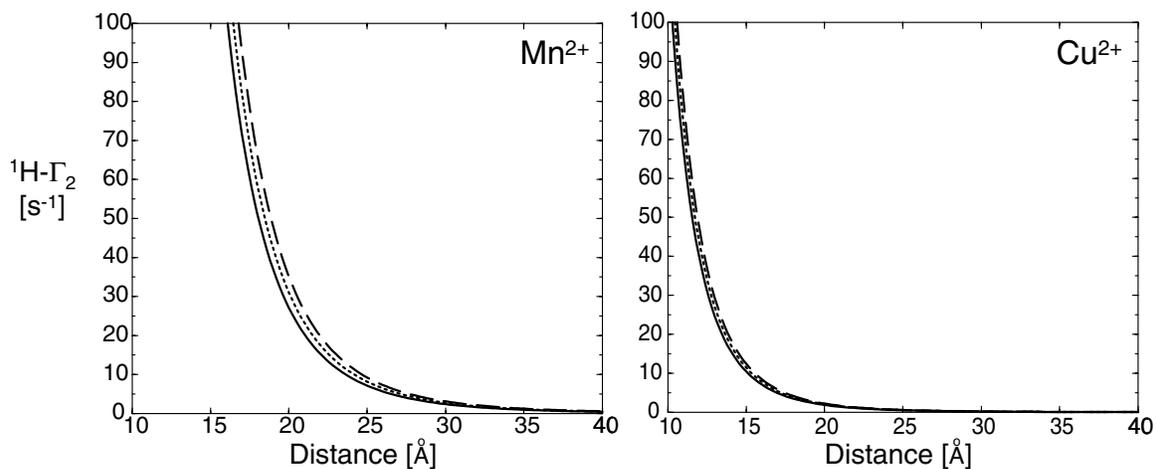**Figure S4.1** Theoretical relationships between distance and ^1H-Γ_2 for Mn^{2+} (electron spin quantum number S = 5/2) and Cu^{2+} (S = 1/2). Solid, dotted, and dashed lines are for molecules with rotational correlation times of 8, 12 and 20 ns, respectively. Electronic relaxation correlation times for Mn^{2+} and Cu^{2+} are assumed to be 5 and 3 ns, respectively, which are based on the current work and Donaldson et al. (Donaldson, L. W.; Skrynykov, D. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843-9847). The *g* factors for Mn^{2+} and Cu^{2+} are 2.0 and 2.1, respectively. Only dipolar interactions between electrons and ^1H nuclei are considered since the contribution of Curie spin relaxation is negligible for these cases. The curves are calculated for a spectrometer frequency of 500 MHz but the field-dependency is very small for both metals.