Supporting Information

Cu^{2+}-based distance measurements by pulsed EPR provide distance constraints for DNA backbone conformations in solution

Shreya Ghosh, Matthew J. Lawless, Hanna J. Brubaker, Kevin Singewald, Michael R. Kurpiewski, Linda Jen-Jacobson and Sunil Saxena*[ǂ]

[*] – Department of Biological Sciences, University of Pittsburgh, 320 Clapp Hall, Pittsburgh, PA 15260 (USA)
[ǂ] - Department of Chemistry, University of Pittsburgh, 219 Parkman Ave., Pittsburgh, PA 15260 (USA)

Mass Spectroscopy

The mass spectrum of each strand of the DNA duplexes (strands A and B) was obtained after synthesis by ATDBio. All the spectra and the corresponding mass of the strands in amu are shown in Figure S1.

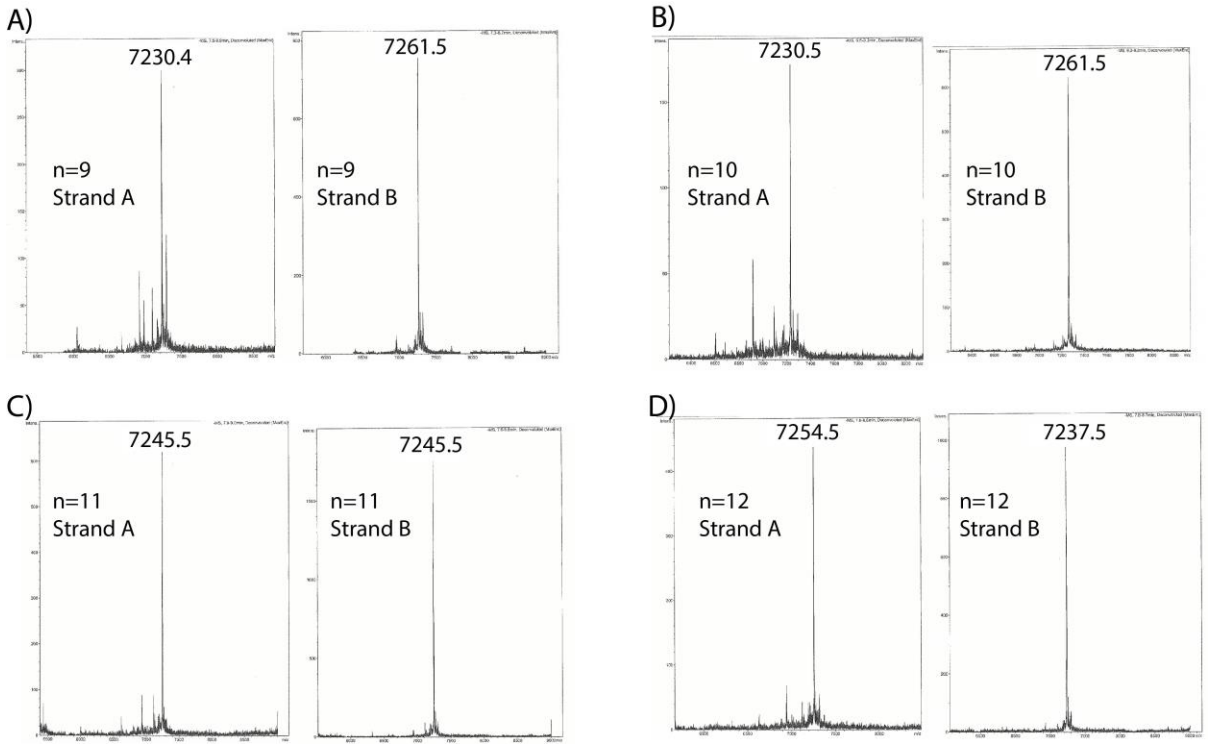

Figure S1: The mass spectrum of each strand was completed upon synthesis by ATDBio. The mass of each individual strand (in amu) of each duplex is shown above.

Calculation of DNA concentration:

Concentrations of purified single-stranded oligodeoxynucleotides were determined spectrophotometrically from molar extinction coefficients calculated by a nearest neighbor method[1–3]. Concentrations of duplexes formed from stoichiometric amounts of the complementary single strands were also determined spectrophotometrically using molar extinction coefficients calculated by the nearest neighbor method[1–3]. The contribution of the DPA to the 24-base pair duplex was accounted for in three ways: a) Assuming an unpaired G in the position of the DPA b) Assuming an unpaired A in the position of the DPA c) Assuming a G-C base-pair for the DPA and opposing d-Spacer. Molar absorptivity coefficients for duplexes calculated according to a), b) and c) differed only by ~2%. The overall accuracy of these DNA concentrations is within 5%.

CW-EPR of control DNA

Figure S2 shows the CW-EPR spectrum of the control DNA (red) and the corresponding simulation (black dashed). $CuCl_2$ was added in a 1:1 ratio with the DNA, with the DNA concentration being 100 μM.

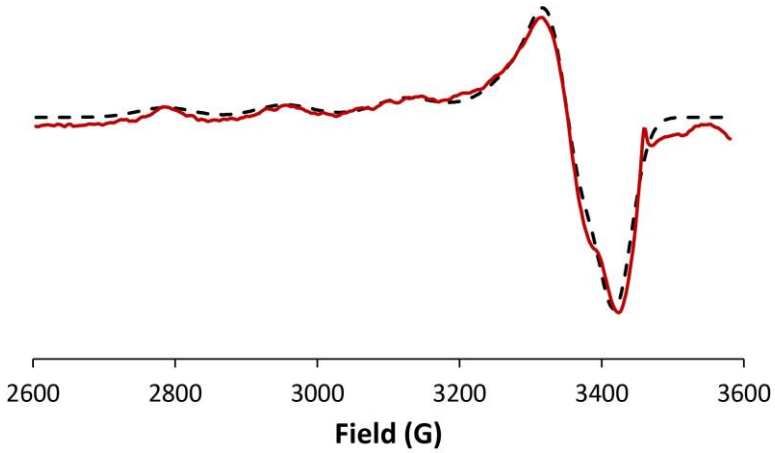

Figure S2: CW simulation of the control DNA shows a single component fit with $g_{||}$ and $A_{||}$ values of 2.280 and 163 G.

DEER at Q-band frequency

We used DNA duplex with a base pair separation of 11 between the Cu^{2+}-DPA labels to probe orientational selectivity effects at Q-band frequency. For the DEER, we used observer pulses, ($\pi/2$) and (π), of 14 ns and 28 ns, respectively and a pump pulse, (π) of 32 ns. The observer frequency was set 100 MHz higher in frequency than the pump. The increment of time after the second pulse was 24 ns. DEER spectra were acquired from 11220 G to 11820 G at 8 different magnetic field values. All DEER data were collected at 20 K. All DEER data were analysed using DEERAnalysis2018[4].

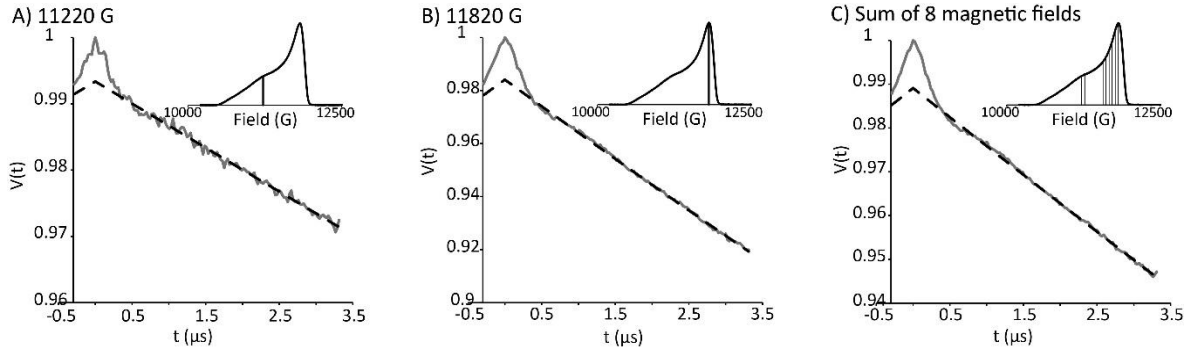

Figure S3: DEER signals on Cu^{2+}-DPA-DNA (n=11) performed at Q-band frequency. DEER performed at A) $g_{\|}$ region (11220 G) B) $g_{\perp}$ region (11820 G) C) 8 different magnetic fields.

DEER at X-band frequency

For X-band DEER experiments, the observer pulse lengths, (π/2) and (π), were 16 ns and 32 ns, respectively while the pump pulse length, (π), was 16 ns. The delay after the second pulse was incremented using step sizes that varied from 8 ns up to 28 ns, depending on the DPA base pair separation. Figure S4 shows raw time-domain DEER data for experiments performed at both $g_{\parallel}$ (grey) and $g_{\perp}$ (blue) regions for different DNA duplexes. Figure S5 shows the background subtracted time domain DEER data. Figure S6 shows the validation for distance distributions obtained in the $g_{\perp}$ region.

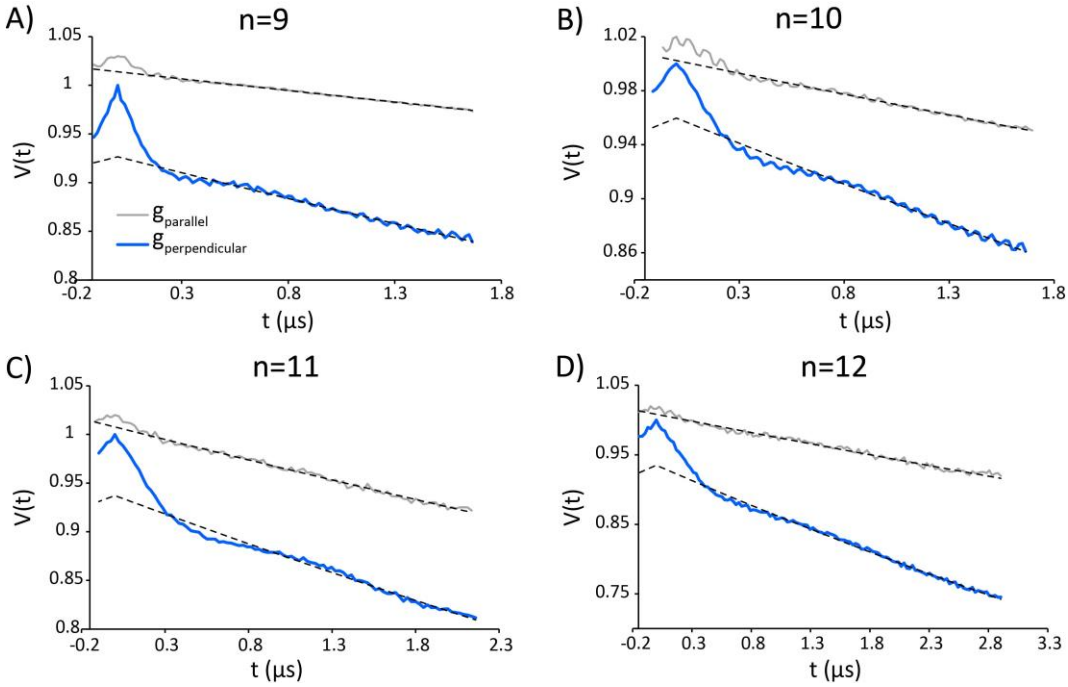

Figure S4: DEER signals for DNA duplexes with n ranging from 9 to 12, performed at fields corresponding to $g_{\parallel}$ (grey) and $g_{\perp}$ (blue) at X-band frequency.

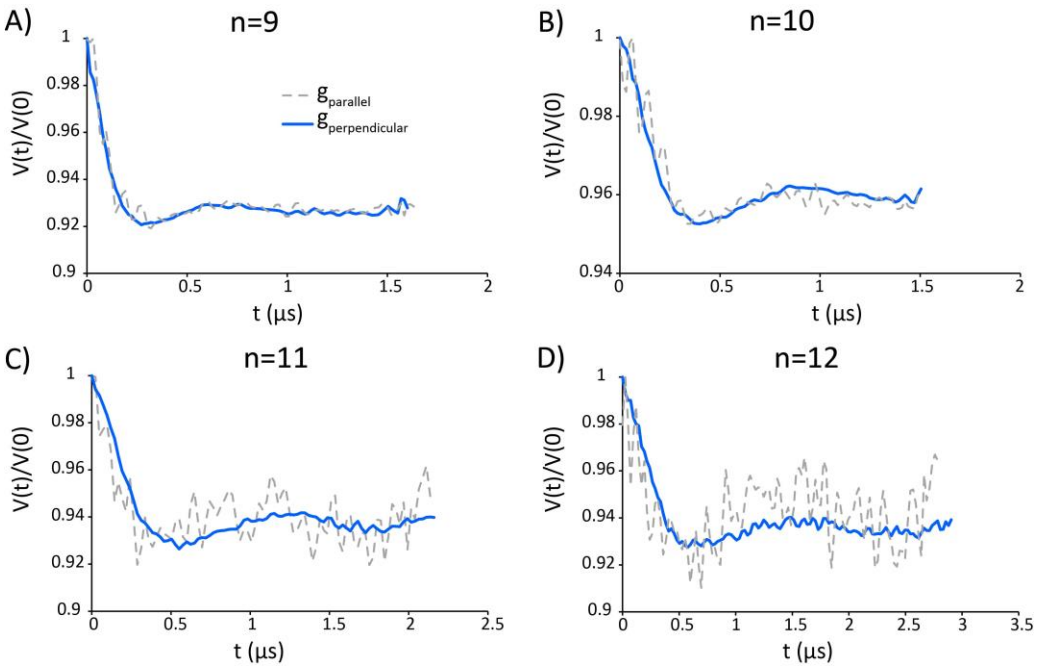

Figure S5: The background subtracted time domain DEER data of $g_{\parallel}$ (grey) and $g_{\perp}$ (blue) regions. Minimal differences in dipolar frequency are observed, thereby showing no orientational selectivity effects at X-band frequency.

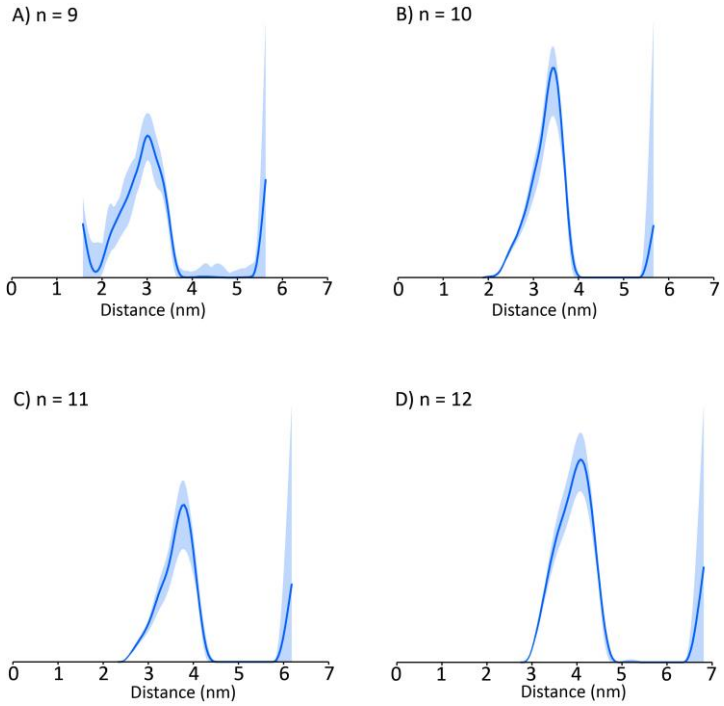

Figure S6: Validation of distance distributions obtained via Tikhonov Regularization for duplexes n=9 to n=12. All the corresponding DEER were performed at X-band frequency at the field of highest intensity.

MD simulations

Molecular dynamics simulations were performed on an unmodified DNA for a total of 100 ns using the AMBER parmbsc1 (bsc1)[5] force field. The B-DNA helix used in the simulations were built using the Nucleic Acid Builder (NAB) module in the AMBER suite[6]. The DNA duplex was then solvated in an explicit 12 Å water box using the TIP3P water model[7] and neutralized with Na^+ and Cl^- ions with a final salt concentration of 75 mM. All simulations were performed using the pmemd program in the AMBER16 software package. The last 95 ns of the production runs were used for analysis, to ensure that the system has been fully equilibrated.

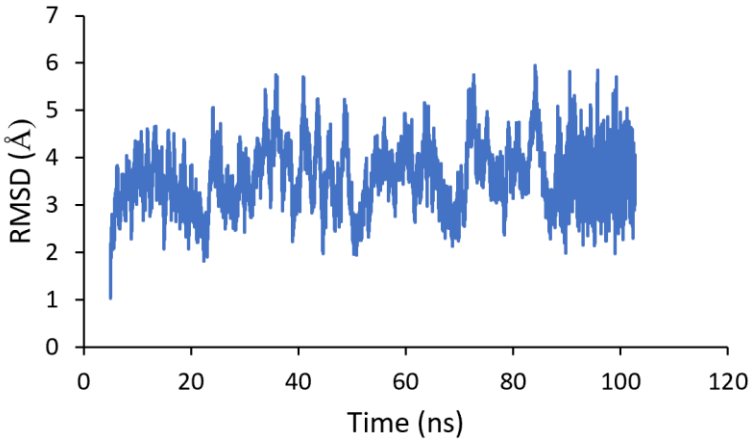

Figure S7: Root mean square deviation (RMSD) of the DNA duplex over the 100 ns of the MD run. The small fluctuation in RMSD shows a stable duplex during the simulations.

Twist-stretch model of DNA to predict Cu^{2+}-DPA flexibility

We constructed a geometric model of the DNA, considering the twist-stretch or breathing motion of the DNA. This DNA motion is associated with a varying radius while the pitch of the helix remains constant. Previous work reporting this model were able to predict the flexibility of the label by comparing the standard deviations of experimental distance distributions with the model[8]. In our work, we followed a similar approach whereby we considered the length of our label, Cu^{2+}-DPA and roughly estimated the flexibility of the label.

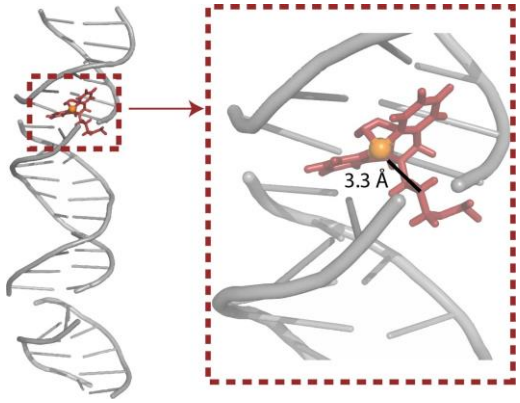

Figure S8: DPA structure, built on available crystal structure parameters, is incorporated in a DNA duplex and the approximate length of the label is calculated to be 3.3 Å.

First, to estimate the length of the label, we built a DPA-DNA duplex, where we incorporated the DPA at two specific sites in the DNA using PyMOL software[9]. The DPA structure itself was built on the bond length and bond angle information available from crystal structure[10]. Figure S8 shows the estimated length of the DPA label, that is, the length of the Cu^{2+} from the DNA backbone is roughly 3.3 Å.

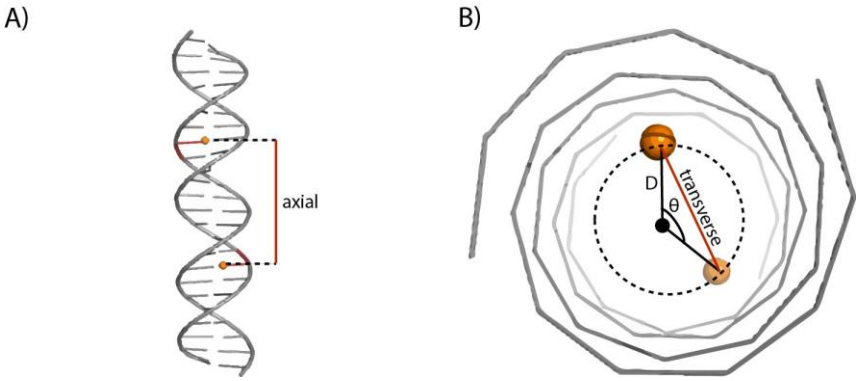

Figure S9: Geometric model of the B-DNA helix. The model comprises of two components. A) axial and B) transverse. To calculate the angle θ, the DNA helix is viewed from the top. The Pythagorean sum of the two components give the final length between two labels placed inside a DNA.

Second, we constructed the physical model of the DNA twist-stretch motion. This DNA motion is built on two main parameters: an axial and a transverse distance between the two probes[8,11], as shown in Figure S9. The axial distance is calculated as the following:

$$axial = axial_0 + rn \quad (1)$$

where $axial_0$ is the separation between the probes when separated by zero base pairs, r is the rise per base pair and n is the number of base pairs between the probes.

The transverse distance is calculated as:

$$transverse = 2Dsin(\frac{\theta}{2}) \quad (2)$$

where D is the radial displacement of the probes of off the helical axis and θ is the angle between the probes when viewed in projection along the axis of the helix. The angle θ can be further calculated as:

$$\theta = \theta_0 + 2\pi n/N \quad (3)$$

where θ_0 is the angle between the probes at zero base pair separation and N is the number of base pairs per turn (since N base pairs make one full turn or 2π).

Finally, the length between the two probes can be given as the Pythagorean sum of the axial and the transverse components as:

$$Length\ (n) = \sqrt{axial^2 + transverse^2} \quad (4)$$

Third, using the above equations we obtained a distance distribution for different base pair separation for the twist-stretch motion of the DNA. Figure S10 shows the comparison of the experimental distance distributions with that from the model. One important aspect while obtaining these distributions is the flexibility of the DNA as well as the label. Previous work on DNA breathing motion estimated the DNA mobility to be ~0.65 Å. Here, along with the DNA mobility, we also added an additional label flexibility.

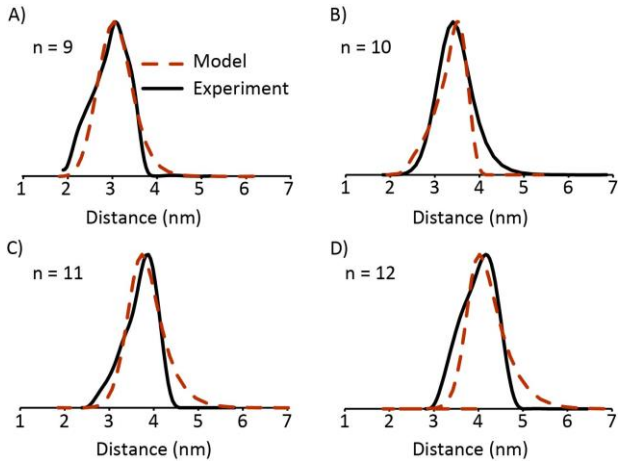

Figure S10: Distance distributions obtained from the twist-stretch model of the DNA compared to the experimental results for different base pair separations between the spin labels.

We then compared our model distance distributions to that of the experiment and varied the label flexibility until we got a reasonable fit with the experiment. We observed a reasonable agreement of the experiment with the model for an additional label flexibility of ~1.6 Å, besides that of the DNA.

Twist-stretch model of linear Cu^{2+}-DPA-DNA to predict on backbone distance

In the twist-stretch model, the axial component increases linearly, and the radial component is oscillatory as the base pair separation increases. As the base pair separation between the labels increases, the axial becomes the more dominant component. Hence, the radial offset of 3.3 Å of Cu^{2+} from the backbone becomes negligible when obtaining the resultant distance, as shown in Figure S11-A. Furthermore, we illustrate our point by plotting the most probable distances from the twist-stretch model for DNA backbone and Cu^{2+}-DPA-DNA, shown in Figure S11-B. Clearly, for a base-pair separation of 4 or higher, the contribution of the radial component of our DPA label towards the resultant distance is small. As a result, the Cu^{2+}-DPA distance closely resembles the DNA backbone distance, specifically the C3' and the C4'. Conceptually, in our method, the spin label point towards each other into the helical axis rather than away from the helical axis. In addition, the size of the linker is much smaller than traditional labels which can be as much as a nm long. Together, these factors help reduce the effect of the offset.

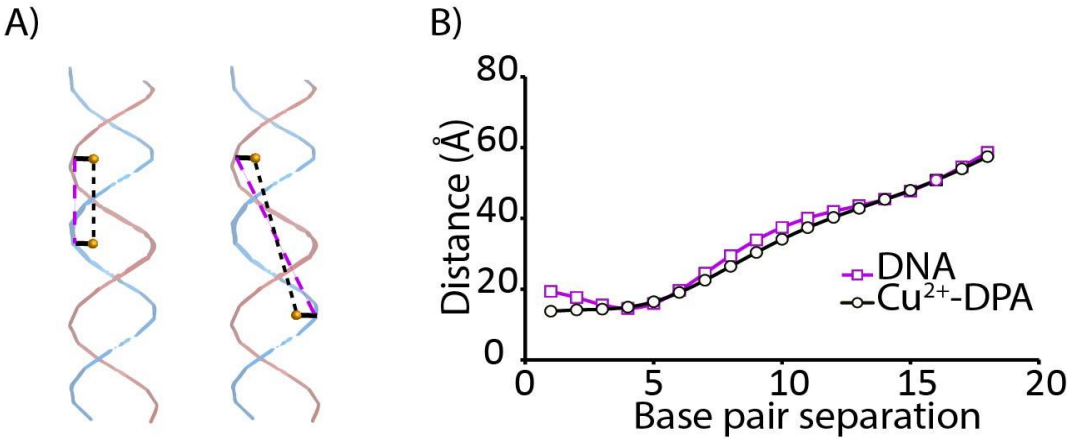

Figure S11: A) Two different models of the straight DNA where the Cu^{2+}-DPA motifs are a) on top of each other and b) opposing one another. The Cu^{2+} is modeled to be at 3.3 Å from the DNA backbone. In the first case, the Cu^{2+}-Cu^{2+} distance can be directly translated to DNA backbone distance as the 3.3 Å deviation does not play any role here. In the second case, the deviation, being significantly smaller than the axial distance, also does not contribute to the resultant distance. B) Plot of most probable distance of the DNA backbone and Cu^{2+}-DPA-DNA, obtained from the twist-stretch model, for different base-pair separations. We observe that for a base-pair separation of 4 and higher, the 3.3 Å between the Cu^{2+} and the DNA backbone has negligible contribution to the overall distance.

In-silico model of linear DNA and comparison with MD simulations

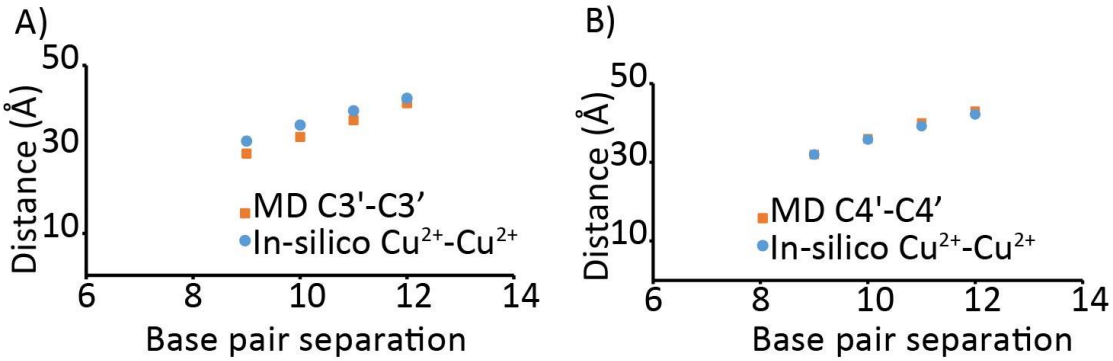

Figure S12: Comparison of distance from in-silico model of Cu^{2+}-DPA-DNA with most probable distance from MD simulations for A) C3'-C3' and B) C4'-C4'. We see a slightly better agreement of the Cu^{2+}-Cu^{2+} distance with the C4'-C4' backbone distance.

We performed in-silico modeling of the Cu^{2+}-DPA on the linear DNA, with same sequence as used in our experiments. We built a DNA duplex, and incorporated the Cu^{2+}-DPA at two specific sites using PyMOL software[9] and performed the following analysis for base-pair separations of 9-12. We measured the Cu^{2+}-Cu^{2+} distance from the in-silico model and compared with the most probable distances of the C3'-C3' and C4'-C4' distance distributions obtained from MD simulations. We observed a slightly better agreement of the Cu^{2+}-Cu^{2+} distances to the MD C4'-C4' distances than the C3'-C3', as shown in Figure S12. However, the C3'-C3' also agrees within 2-3 Å. Importantly, this agreement also shows that the offset of Cu^{2+} from the DNA backbone does not significantly affect the overall distance.

Analysis of experimental standard deviation:

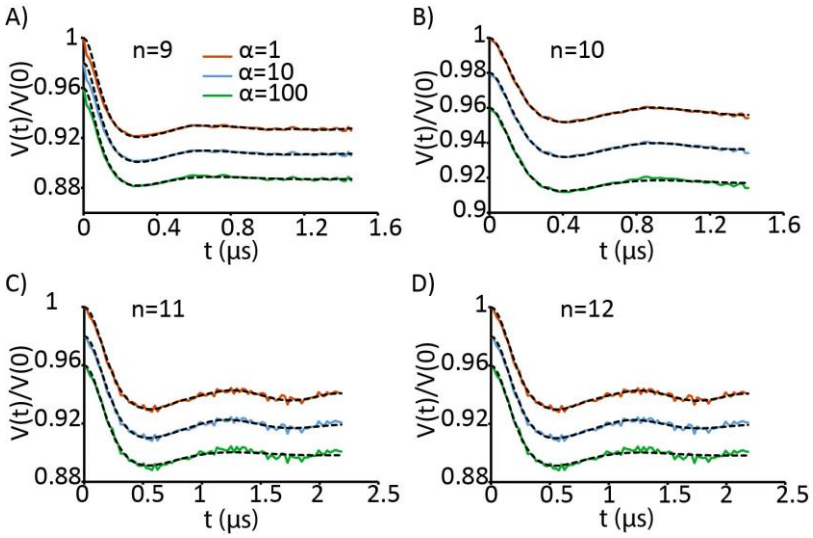

Figure S13: Background subtracted time-domain DEER data for the different DNA duplexes with variable α-values (1, 10 and 100). There is a good agreement of the experiment with the fit, over the different α-values used.

To estimate the flexibility of Cu^{2+}-DPA label, we further calculated the standard deviation of our experimental results and compared with the twist-stretch model. When analyzing distance distributions using Tikhonov regularization, the width of our distance distributions is strongly dependent on the regularization parameter, α. As a result, we varied the α-value between 1 to 100, making sure that we still obtain a good fit to our data. Figure S13 shows the variable α used to analyze each individual DEER data and the corresponding fits. We obtained the standard deviation for each of these fits and calculated the average (Figure S14-A). On the other hand, the distance distributions provided in the main text (Figure 6) were obtained using the optimum α-value. Accordingly, we calculated the standard deviations of the above distributions and compared with that of the average. We see a reasonable agreement between the two as well, as shown in Figure S14-B.

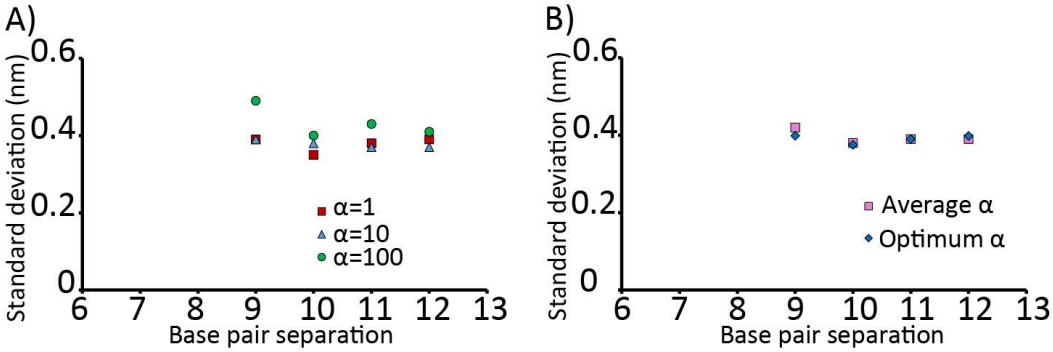

Figure S14: A) Comparison of standard deviations of distance distributions for varying α-values. B) Comparison of the average of the standard deviations for different α with that obtained from the optimum α-value.

In-silico model of non-linear DNA

To examine how the Cu^{2+}-DPA label would report on the backbone distances of a bent DNA, we took the following approach. We generated two PDB structures of the DNA using the same sequence as the control DNA (Figure 1 of main text), with a bent angle of 90° and 150°, respectively, as shown in Figure S15. The bent structures were created using the 3D-DART software[12]. We calculated the C3'-C3' and C4'-C4' distances for base pair separations of 9 through 12. To compare how Cu^{2+}-DPA reports on the bent DNA, we generated an in-silico model, where we incorporated the Cu^{2+}-DPA at two specific sites of the bent DNA using PyMOL software[9]. We measured the Cu^{2+}-Cu^{2+} distance, repeated this step for different base pair separations and compared the DNA backbone distances with the Cu^{2+}-DPA distance. Clearly, the Cu^{2+}-DPA distances agree within 2-3 Å, within the caveats of the model. Finally, we used a bent DNA PDB structure, reported in literature (PDB 1A73)[13]. This PDB structure comprises of an endonuclease bound DNA where the DNA is strongly bent for cleavage. We performed similar in-silico modeling on the PDB structure and measured the Cu^{2+}-Cu^{2+} distance along with the backbone distances.

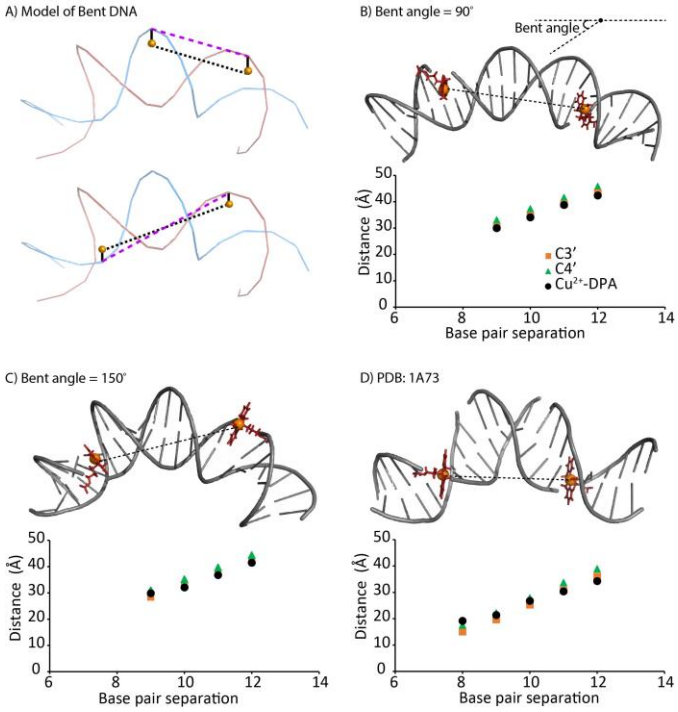

Figure S15: A) Model of a bent DNA showing the backbone distance and the Cu^{2+}-Cu^{2+} distance. In the first case, the position of the two labels being almost parallel causes the Cu^{2+}-Cu^{2+} distance to directly translate into the DNA backbone distance. In the second case, the labels are anti-parallel. However, the offset of Cu^{2+} from the backbone does not significantly contribute to the measured distance. B) and C) Modeled DNA structures with a bent angle of 90° and 150°, respectively. D) PDB of a bent DNA (PDB: 1A73). C3' (orange squares) and C4' (green triangles) DNA backbone distances were measured. In-silico modeling was performed incorporating Cu^{2+}-DPA phosphoramidite into the DNA duplex for different base-pair separations and the corresponding Cu^{2+}-Cu^{2+} distance was measured (black circles). All the three distances show reasonable agreement within the caveats of the model.

References

1 C. R. Cantor and I. Tinoco, Calculated optical properties of 64 trinucleoside diphosphates, *Biopolymers*, 1967, **5**, 821–835.

2 E. G. Richards, Use of tables in calculation of absorption, optical rotatory dispersion and circular dichroism of polyribonucleotides, *Handb. Biochem. Mol. Biol. 3rd Edn.*, 1975, **1**, 596–603.

3 M. Senior, R. A. Jones and K. J. Breslauer, Influence of dangling thymidine residues on the stability and structure of two DNA duplexes, *Biochemistry*, 1988, **27**, 3879–3885.

4 G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger and H. Jung, DeerAnalysis2006 - a Comprehensive Software Package for Analyzing Pulsed ELDOR Data, *Appl. Magn. Reson.*, 2006, **30**, 473–498.

5 I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case and M. Orozco, Parmbsc1: a refined force field for DNA simulations, *Nat. Methods*, 2016, **13**, 55–58.

6 D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, The Amber biomolecular simulation programs, *J. Comput. Chem.*, 2005, **26**, 1668–1688.

7 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926–935.

8 A. Marko, V. Denysenkov, D. Margraf, P. Cekan, O. Schiemann, S. T. Sigurdsson and T. F. Prisner, Conformational flexibility of DNA, *J. Am. Chem. Soc.*, 2011, **133**, 13375–13379.

9 W. L. Delano, *PyMOL: An Open-Source Molecular Graphics Tool*, 2002.

10 K. Y. Choi, H. Ryu, N. Do Sung and M. Suh, Synthesis, properties, and X-ray structure of [Cu(dpa)Cl2] (dpa = di-(2-picolyl)amine), *J. Chem. Crystallogr.*, 2003, **33**, 947–950.

11 R. S. Mathew-Fenn, R. Das and P. A. B. Harbury, Remeasuring the Double Helix, *Science*, 2008, **322**, 446–449.

12 M. van Dijk and A. M. J. J. Bonvin, 3D-DART: a DNA structure modelling server, *Nucleic Acids Res.*, 2009, **37**, W235–W239.

13 K. E. Flick, M. S. Jurica, R. J. Monnat and B. L. Stoddard, DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-Ppol, *Nature*, 1998, **394**, 96–101.