Supporting Information

High resolution measurement of long-range distances in RNA: pulse EPR spectroscopy with TEMPO-labeled nucleotides.

Karin Halbmair[a], Jan Seikowski[a], Igor Tkach[a], Claudia Höbartner[b,*], Deniz Sezer[c,*], Marina Bennati[a,b,*]

[a] Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

[b] Institute for Organic and Biomolecular Chemistry, Georg-August-University of Göttingen, 37077 Göttingen, Germany

[c] Faculty of Engineering and Natural Sciences, Sabanci University, 34956 Istanbul, Turkey.

* Corresponding authors

SI1: Sample characterization by HPLC and CW-EPR spectroscopy.
Figure S1: (a) Anion exchange traces of purified single-stranded double-C^T-labeled RNAs. (b) Spin concentration calibration curve. (c) RNA labelling efficiency.

SI2: Examination of orientation selection.

Figure S2: High- and low-power Q-band DEER on [13,28] RNA duplex.

Figure S3: High-power Q-band DEER on different RNA samples.

Figure S4: PELDOR/DEER traces with longer evolution time.

SI3: Effect of background subtraction on distances and distributions.

Figure S5: Analysis with four different background corrections of samples 1, 3, 6 and 7.

Figure S6: Comparison of PELDOR/DEER traces at higher RNA concentrations.

Table S1: Energies and dihedral angles of the six optimal conformations of C^T.

Figure S7: Values of the C^T dihedral angles during the 34 ns MD simulation.

SI4: Restrained MD simulations.

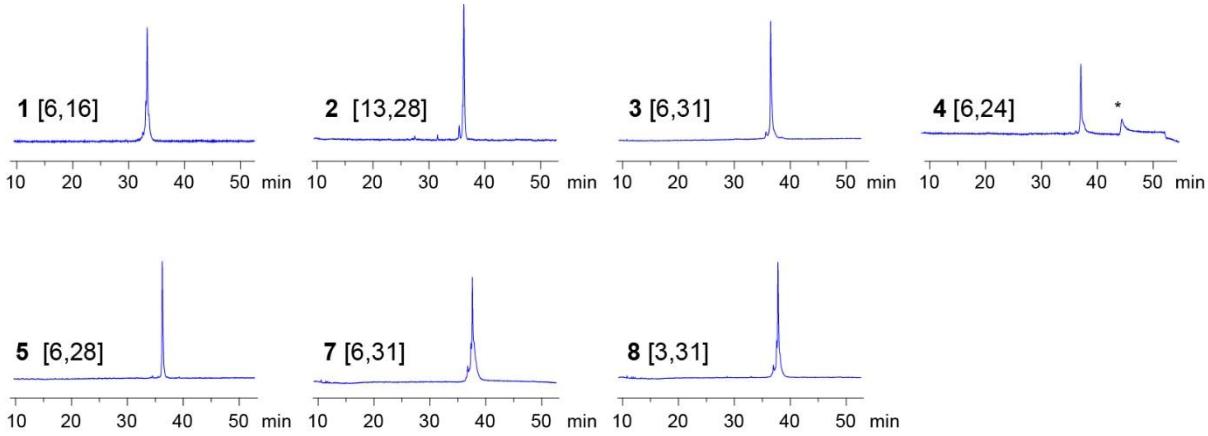
Figure S8: Histogramms of distance distributions calculated for 34bp RNA with spin labels at positions 6,28 and 6,16,31

Figure S9: View of the spin label at position 16 from the MD simulation of 6-16-28.

SI1: Sample characterization by HPLC and CW-EPR spectroscopy

Figure S1(a): Anion exchange traces of purified single-stranded double-C^T-labeled RNAs.

UV absorbance monitored at 260 nm, Dionex DNAPAc PA200, 4x250 mm, 25 mM Tris buffer, pH 8.0, 6 M urea, $NaClO_4$ gradient, 80°C. * is a gradient artefact due to shorter run time.



CW EPR spectroscopy at X-band frequencies was employed to characterize the label efficiency. Room temperature experiments were performed in a Bruker Elexsys E500 spectrometer equipped with a Bruker super high Q resonator ER4122SHQE. Glass capillaries of 1 mm inner diameter (ID) were filled with a sample volume 20 µl. Spin concentrations were calculated by doubly integrating the CW-EPR spectrum and comparing the intensity with a calibration curve recorded with 4-hydroxy TEMPO at concentrations between 50 and 200 µM as well as with the nominal RNA concentration determined by UV absorbance. Labelling efficiencies between 80 and 100 % were determined for all samples (Fig. S1b,c).

Figure S1 (b): Spin concentration calibration curve.
Determination of the effective spin concentration in the RNA samples by continuous wave (CW) EPR. CW-EPR spectra were recorded in RNA buffer at X-band microwave frequencies (9,7 GHz).

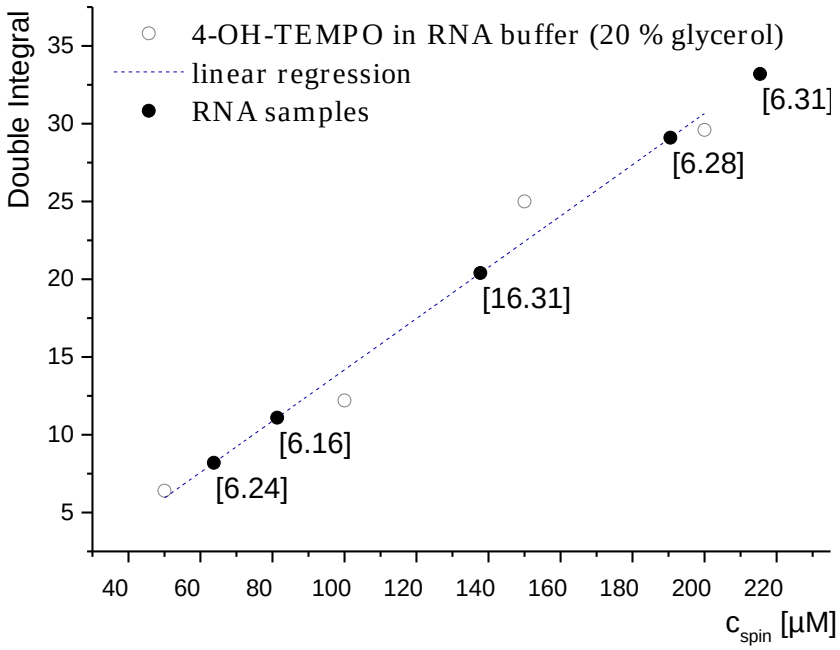
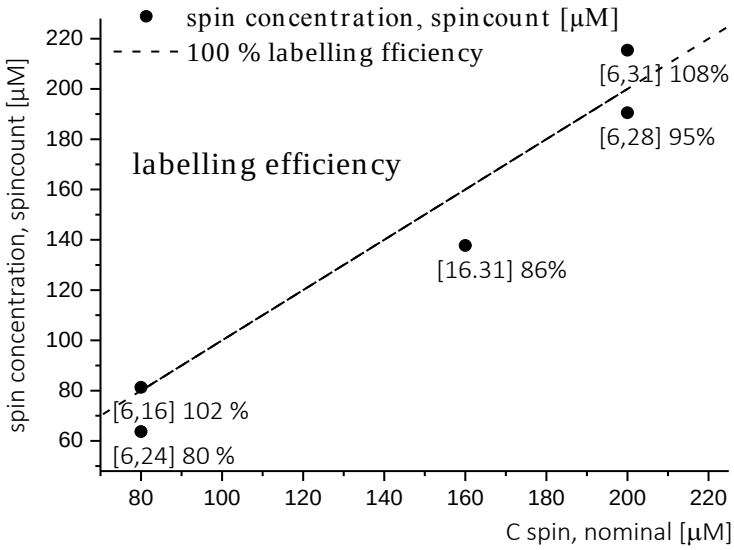


Figure S1(c): RNA labelling efficiency.
Comparison of the expected (nominal) spin concentration with the experimentally determined one and resulting RNA labelling efficiency for each sample.



SI2: Examination of orientation selection.

Figure S2: (a) high-power PELDOR/DEER traces recorded on sample dsRNA[13,28].
Upper part: Spectral positions of the pump and detection pulses according to color code. **Center:** Simulated time traces for four label conformations arising from the combinations of the two most populated conformations 1 and 2 of the label (Table S1). The four conformations are labeled [13/1, 28/1], [13/2, 28/2], [13/1, 28/2] and [13/2, 28/1]. Simulations were conducted using a home written program that takes into account orientation selection in PELDOR experiments.[1, 2] Experimental parameters (pulse lengths, frequency separation, EPR detection frequency, nitroxide EPR parameters) were considered in the simulation. Distance in the individual conformation was extracted from the ideal A-form RNA (each individual conformation has a slightly different distance, indicated in the top of the figures). Relative population of the four conformations was assumed as equal, for simplicity. **Bottom left:** experimental PELDOR/DEER traces, their Fourier transformations (Pake patterns) and corresponding distance distributions obtained with the program DEERanalysis on these traces. **Bottom right:** Sum of the simulated time traces of the four label pair conformations, their Fourier transformations (Pake patterns) and corresponding distance distributions from DEERanalysis. The magenta line displays for comparison a simulation without orientation selection (complete excitation of powder pattern) as sum of the contribution of the four different distances. There are no difference to the simulation using the experimental set up (red line in bottom right plot, $\Delta\nu$ = 90 MHz). The results best illustrate the suppression of orientation selective effects with broadband excitation.

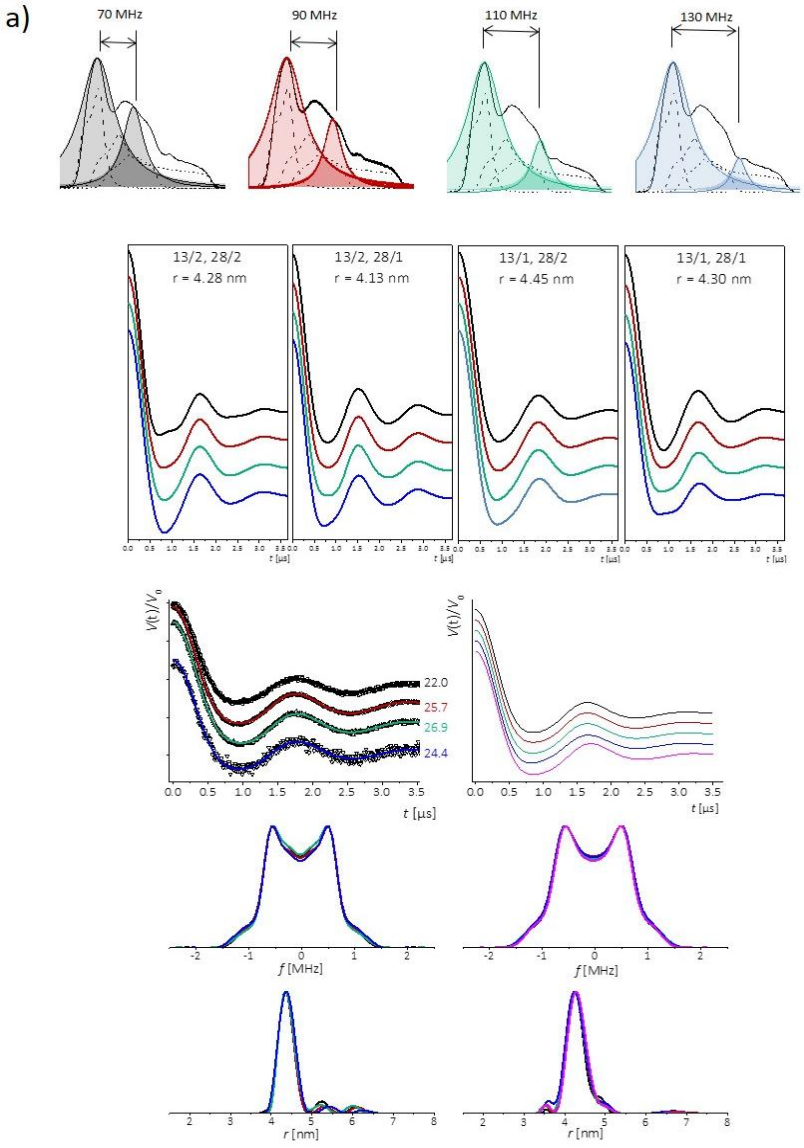


Figure S2 (b): Low-power (selective) PELDOR/DEER traces on sample dsRNA[13,28].

Upper part: Spectral positions of the pump and detection pulses for selective excitation according to color code. **Center:** Simulated time traces for four different label pairs arising from the combinations of the two most populated conformations 1 and 2 of the label (Table S1), using the experimental parameters and the distances corresponding to the label pair in the individual ideal A-form RNA. The four conformations are labeled [13/1, 28/1], [13/2, 28/2], [13/1, 28/2] and [13/2, 28/1]. Each individual conformation has a slightly different distance as indicated in the top of the figures.
Bottom left: Experimental PELDOR/DEER traces, their Fourier transformations (Pake patterns) and corresponding (artificial) distance distributions obtained with the program DEERanalysis on these traces. We note that the program DEERanalysis produces an artifact peak at lower distances as it does not consider orientation selection. However, the intensity of this artificial peak, arising from the dipolar frequency component $\nu_{\|}$, can best visualize the deviation from an ideal Pake pattern. **Bottom center:** Sum of simulated time traces of the four label pairs, their Fourier transformations (Pake patterns) and artificial distance distributions from DEERanalysis. **Bottom right:** Comparison with simulated time traces for the pair having both labels in conformation 2, their Fourier transformations (Pake patterns) and artificial distance distributions obtained with DEERanalysis.
The predicted orientation selection is overall in very good qualitative agreement with the shape of the traces. In the sum of the four contributions, the high field traces (pink) seems to overestimate the contribution of $\nu_{\|}$, which instead is well reproduced by conformation 2/2, suggesting that this is more populated. A further improvement in the simulation of the orientation selectivity would require a more detailed analysis of the individual contributions and more precise orientation selection at higher frequency,[1,2] which goes beyond the scope of this work.

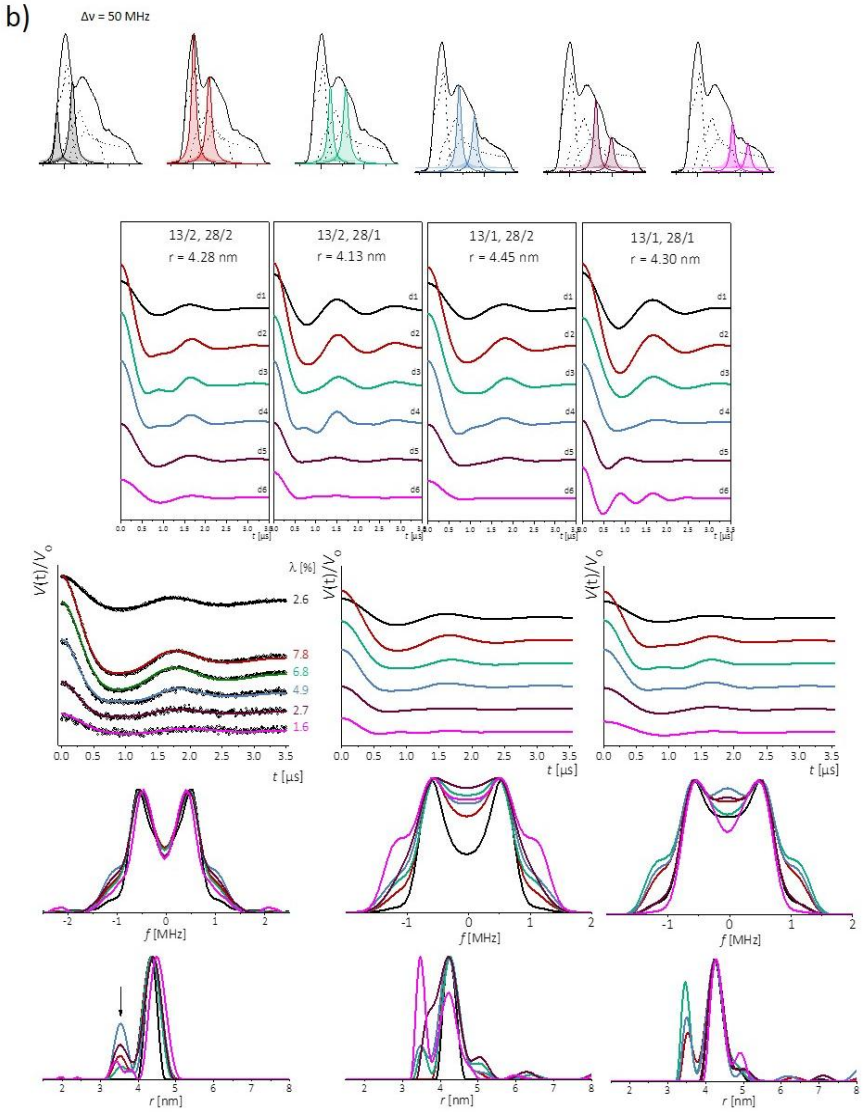


Figure S3: High-power Q-band DEER on different RNA samples.

(a) Background-corrected high-power Q-band DEER traces recorded on RNA samples **2, 3, 4, 5** (Figure 1) and their corresponding Pake patterns from fourier transformation. DEER modulation depths λ are given in %. Traces were recorded by pumping at the m_I = +1 HF transition and detecting at different positions over the EPR spectrum by varying the pump-detect frequency separation (90-190 MHz) as in (b). Pulses in DEER trace were for pump: t_π = 12; detection: $t_{\pi/2}$ and t_π = 12 and 24 ns, respectively.

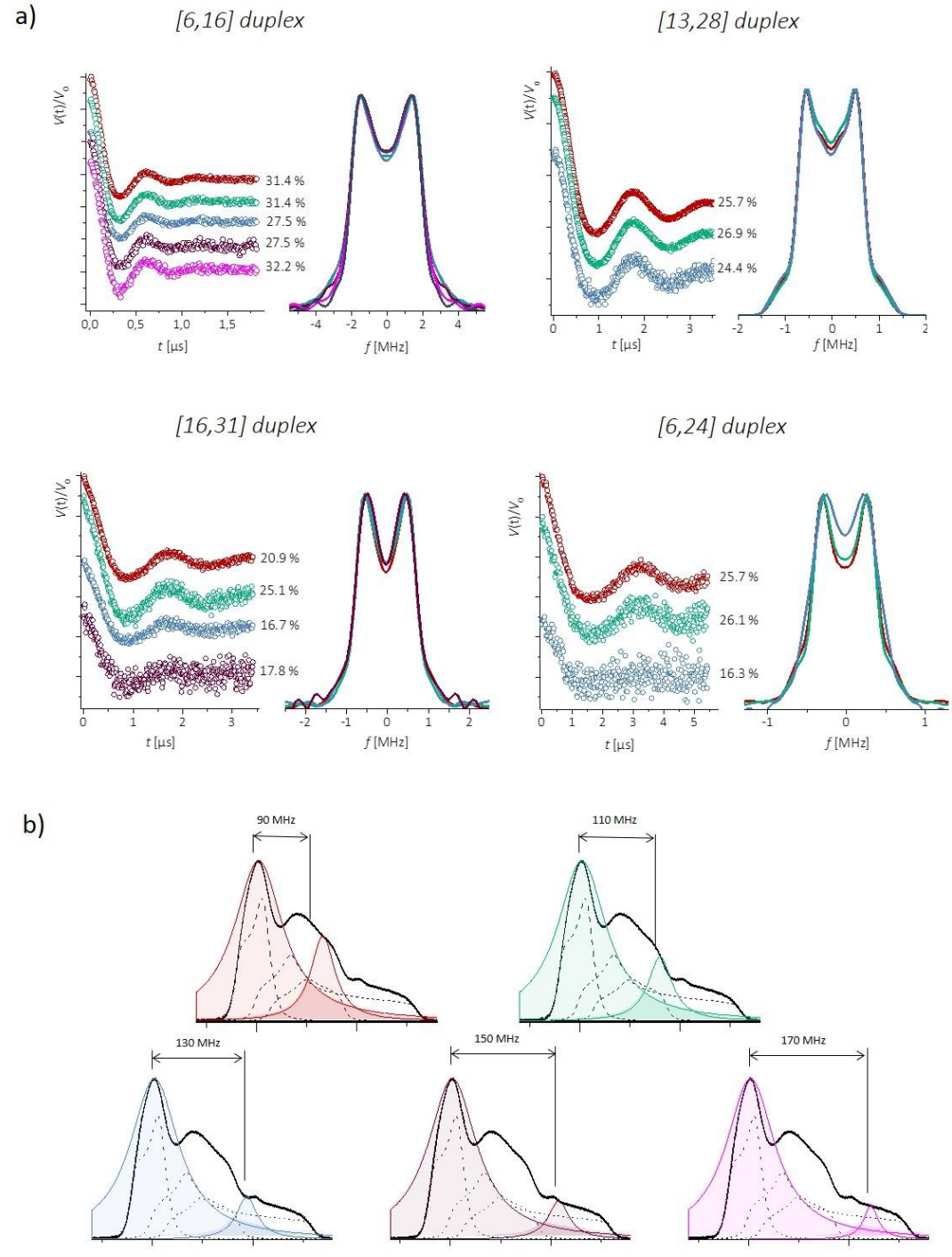
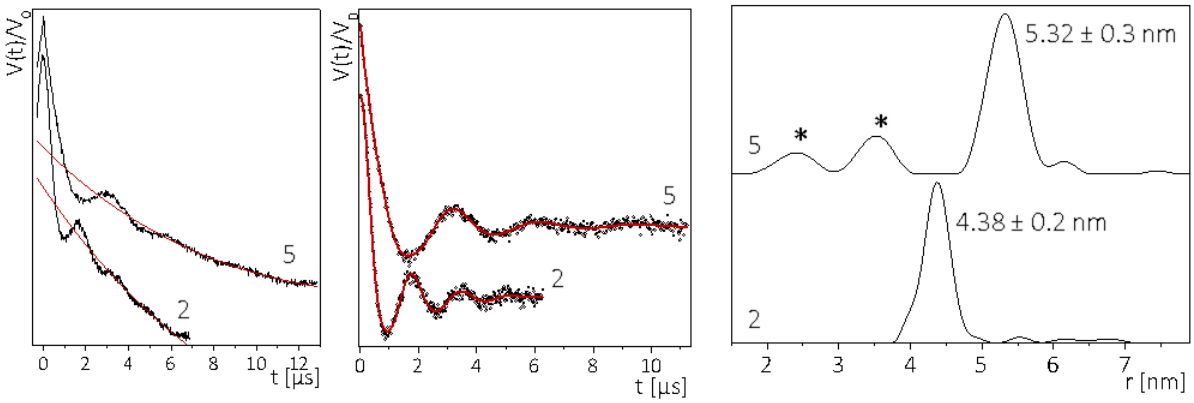


Figure S4: PELDOR/DEER time traces with longer evolution time for sample 2 and 5

Time traces with a longer evolution time T of 7 and 13 µs were recorded for samples **2** and **5** to evaluate the uncertainty in the distance distribution. A second order polynomial fitted to about the last 50 % of the traces was chosen for background correction (left). The distance distributions (right) obtained from DeerAnalysis fit (middle) did not show any difference in Δr when compared to experiments recorded with an evolution time covering 2.5 oscillations (figure 3, main text). The signal to noise however is decreased due to the larger acquisition window. Asterisks in the distance distribution (right) denote additional frequency components that likely arise from aggregates (compare Figure S6).



SI3: Effect of background subtraction on extracted distances and distributions

To evaluate the error in the main distances and their distribution caused by an uncertainty in background subtraction, we have performed the analysis using either a mono-exponential decay or a second order polynomial and choosing either the last 50 % or more than 50% of the traces as fitting region. The most reasonable background subtraction should result (after FT) in a Pake pattern with little distortions around the zero frequency. The α-values, which determine the width of the distribution were chosen by the L-curve criterion, which ensures maximum smoothing (maximum width of the distribution) with minimum r.m.s.d. of the fit.[3]

We present the analysis for four representative traces:

1) trace for which the oscillation has decayed to zero (sample **1**);

2) trace with 2.5 oscillation periods (sample **3**) ;

3) trace with only 1.5 oscillation periods and large S/N (sample **6**)

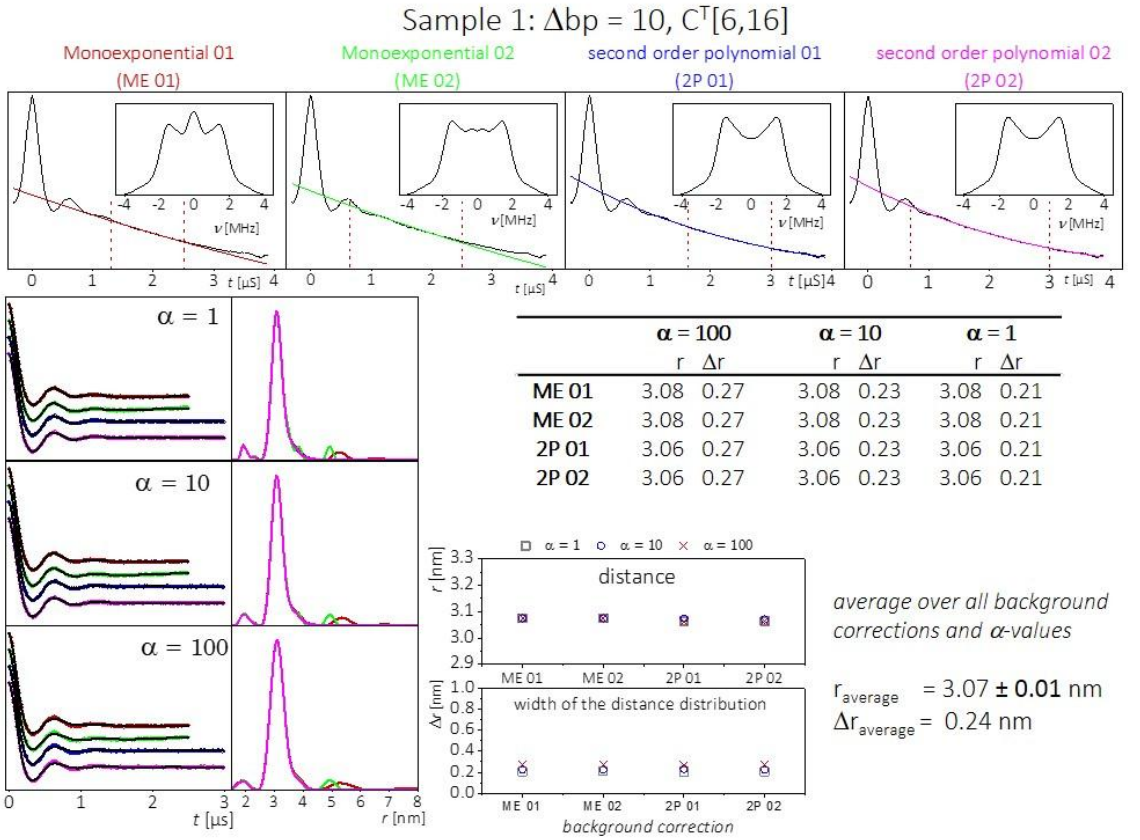
4) trace with only 1.5 oscillation periods and weaker S/N (sample **7**)

The results are summarized in the following table:

sample	*r [nm]	Δr [nm]
1	3.07	0.23
3	4.32	0.25
6	6.24	0.38
7	7.3	0.51

(*) The uncertainty of the peak distance was estimated from three different sets of experiments to be ± 0.05 nm if 2.5 oscillations are visible (sample **1-5**) up to ± 0.1 nm if only 1.5 oscillations are recorded (sample **6-8**). The uncertainty from the background correction in the peak distance is found less than the statistical error arising from the sample triplicates. Δr (distance distribution) was obtained from the analysis of the individual traces as illustrated in the following.

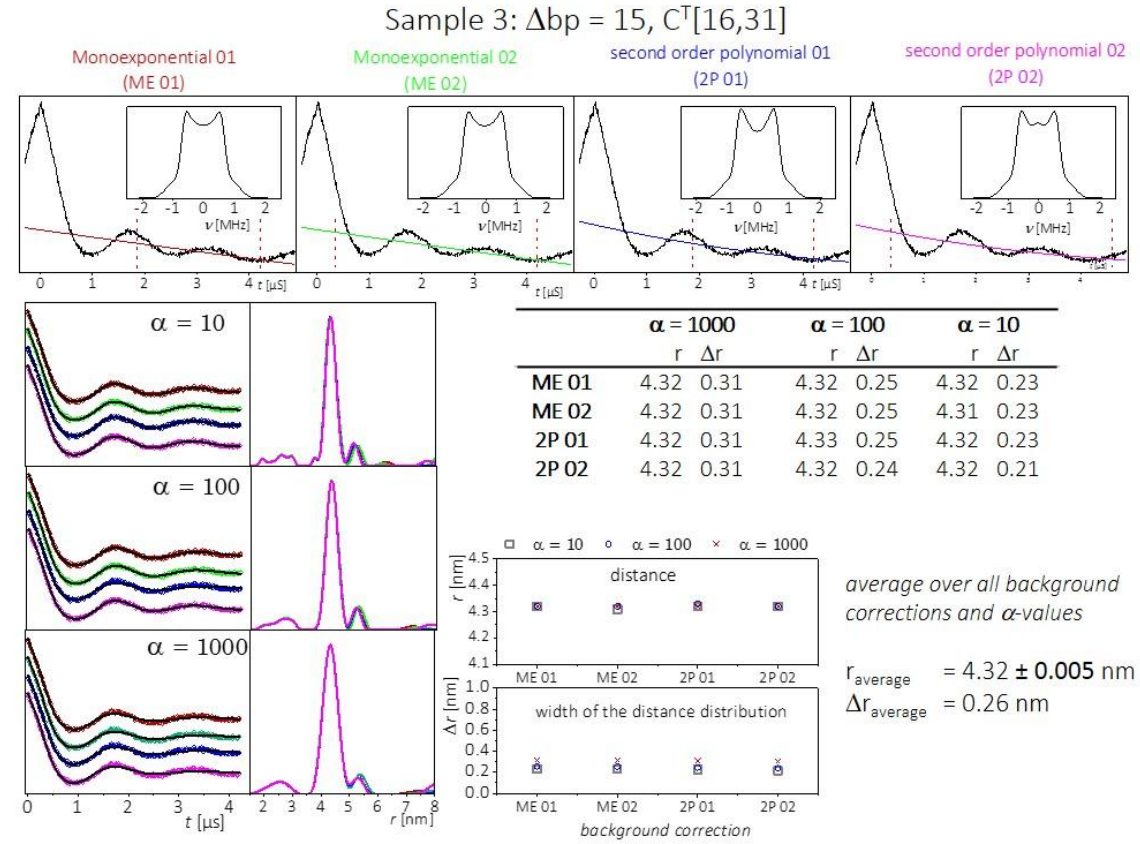
Figure S5: Analysis with four different background corrections of samples 1, 3, 6 and 7.



	α = 100		α = 10		α = 1	
	r	Δr	r	Δr	r	Δr
ME 01	3.08	0.27	3.08	0.23	3.08	0.21
ME 02	3.08	0.27	3.08	0.23	3.08	0.21
2P 01	3.06	0.27	3.06	0.23	3.06	0.21
2P 02	3.06	0.27	3.06	0.23	3.06	0.21

average over all background corrections and α-values

$r_{average}$ = 3.07 **± 0.01** nm
$\Delta r_{average}$ = 0.24 nm



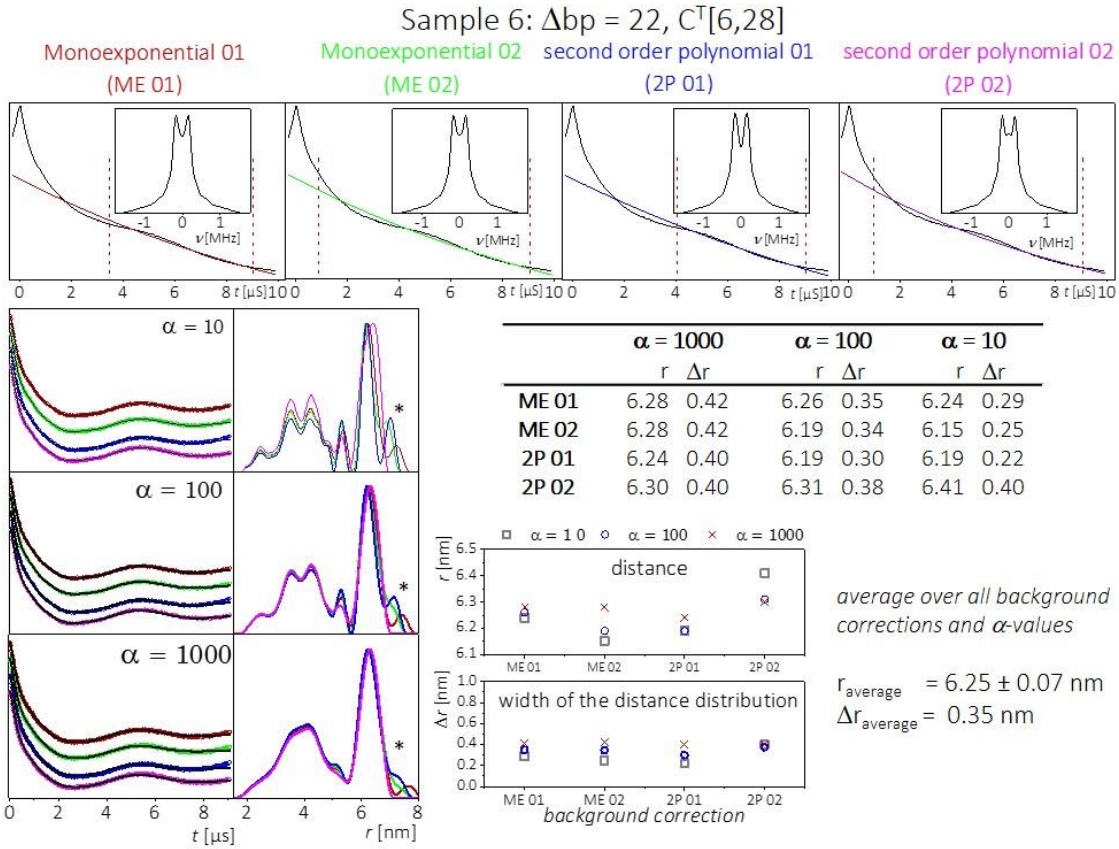
	α = 1000		α = 100		α = 10	
	r	Δr	r	Δr	r	Δr
ME 01	4.32	0.31	4.32	0.25	4.32	0.23
ME 02	4.32	0.31	4.32	0.25	4.31	0.23
2P 01	4.32	0.31	4.33	0.25	4.32	0.23
2P 02	4.32	0.31	4.32	0.24	4.32	0.21

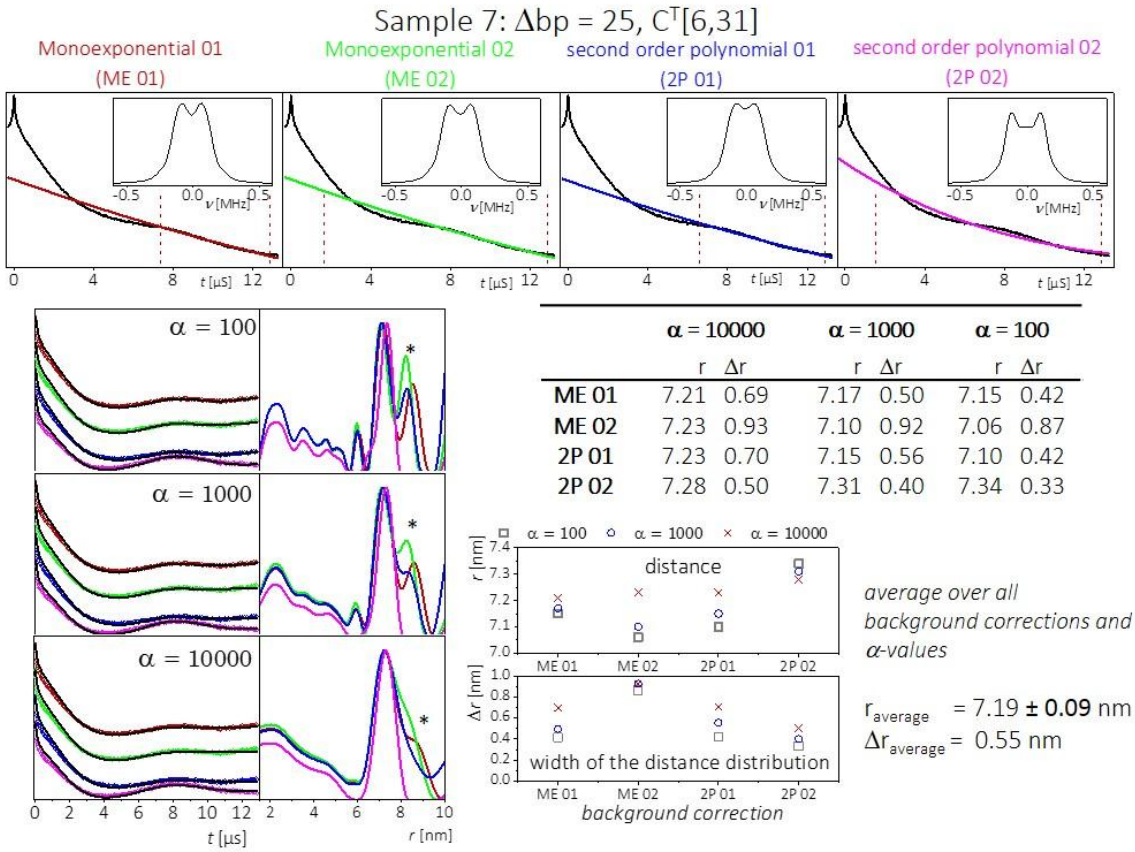
average over all background corrections and α-values

$r_{average}$ = 4.32 **± 0.005** nm
$\Delta r_{average}$ = 0.26 nm

sample	1		3	
	r [nm]	Δr [nm]	r [nm]	Δr [nm]
ME 01	3.08	0.23	4.32	0.25
ME 02	3.08	0.23	4.32	0.25
2P 01	3.07*	0.23*	4.32*	0.25*
2P 02	3.06	0.23	4.33	0.24
average	3.07	0.23	4.32	0.25
σ	± 0.01	± 0.00	± 0.01	± 0.01

For sample 1 and 3 it is visible that the peak distance varies not more than 0.02 nm (standard deviation $\sigma = \pm 0.01$) as a function of the background subtraction, which is less than the statistical error when repeating the experiment. Results from background correction 2P 01 (marked with an asterisk in the table) are displayed in the main text (figure 3).





sample	6		7	
	r [nm]	Δr [nm]	r [nm]	Δr [nm]
ME 01	6.26	0.35	7.21	0.69
ME 02	6.19	0.34	7.23	0.93
2P 01	6.19	0.31	7.23	0.70
2P 02	6.32*	0.38*	7.28*	0.51*
average	6.24	0.35		
σ	± 0.06	± 0.03		

For sample 6 and 7 the peak distance varies up to 0.12 nm ($\sigma = \pm 0.06$) with the chosen background subtraction. The background model marked with asterisk was chosen for analysis given in the main text (figure 3). The width of the distribution for sample 7 however strongly depends on the background correction, which is due to the artefacts marked by the asterisk in the distribution. As only the second order polynomial fitted to more than 50 % of the trace (2P 02) resulted in a reasonable Pake pattern, only this analysis was considered. For sample 6 the largest Δr was chosen as upper limit.

Figure S6: PELDOR/DEER time traces of samples 6 - 8 at duplex concentrations of 100 µM.

In the traces with longer inter spin distances an additional high frequency component is visible, which is particularly strong in sample **6**. Analysis of trace of **6** displays a second distance at around 3.5 nm. Extraction of the Δbp from our ruler (Fig. 4) leads to a base pair separation of about 12. This separation is well consistent with an aggregate, in which the two duplexes add to form a longer one, as shown in the picture below. This form of aggregation can explain the fact that the artefact is preferably observed when the labels are inserted at nucleotides close to the end of the sequence. In this case, the intermolecular distance becomes shorter and better visible.

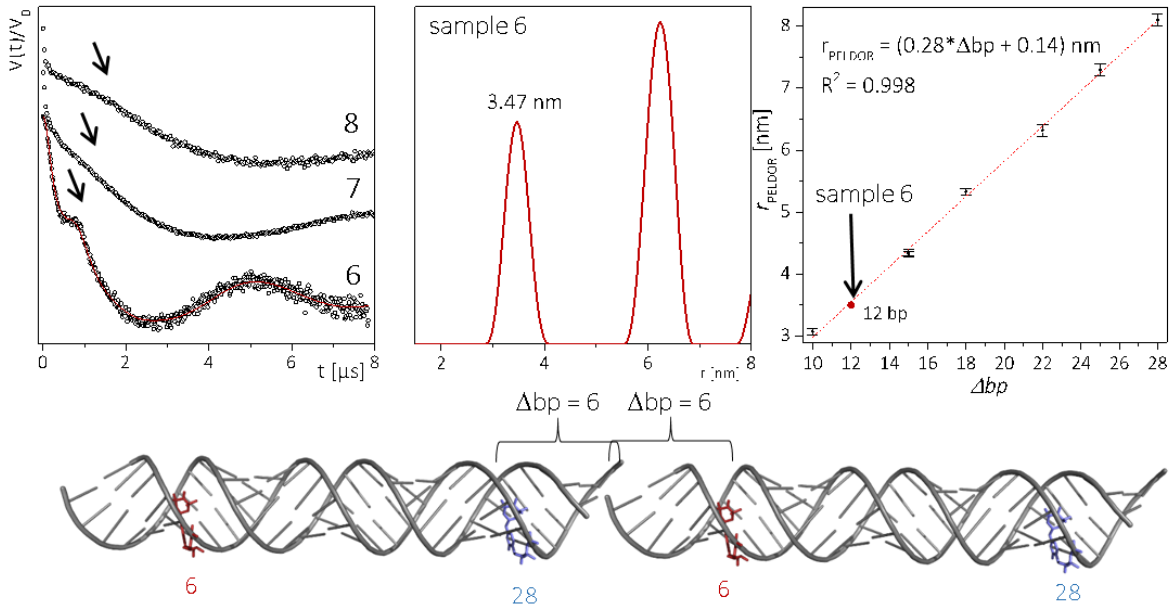
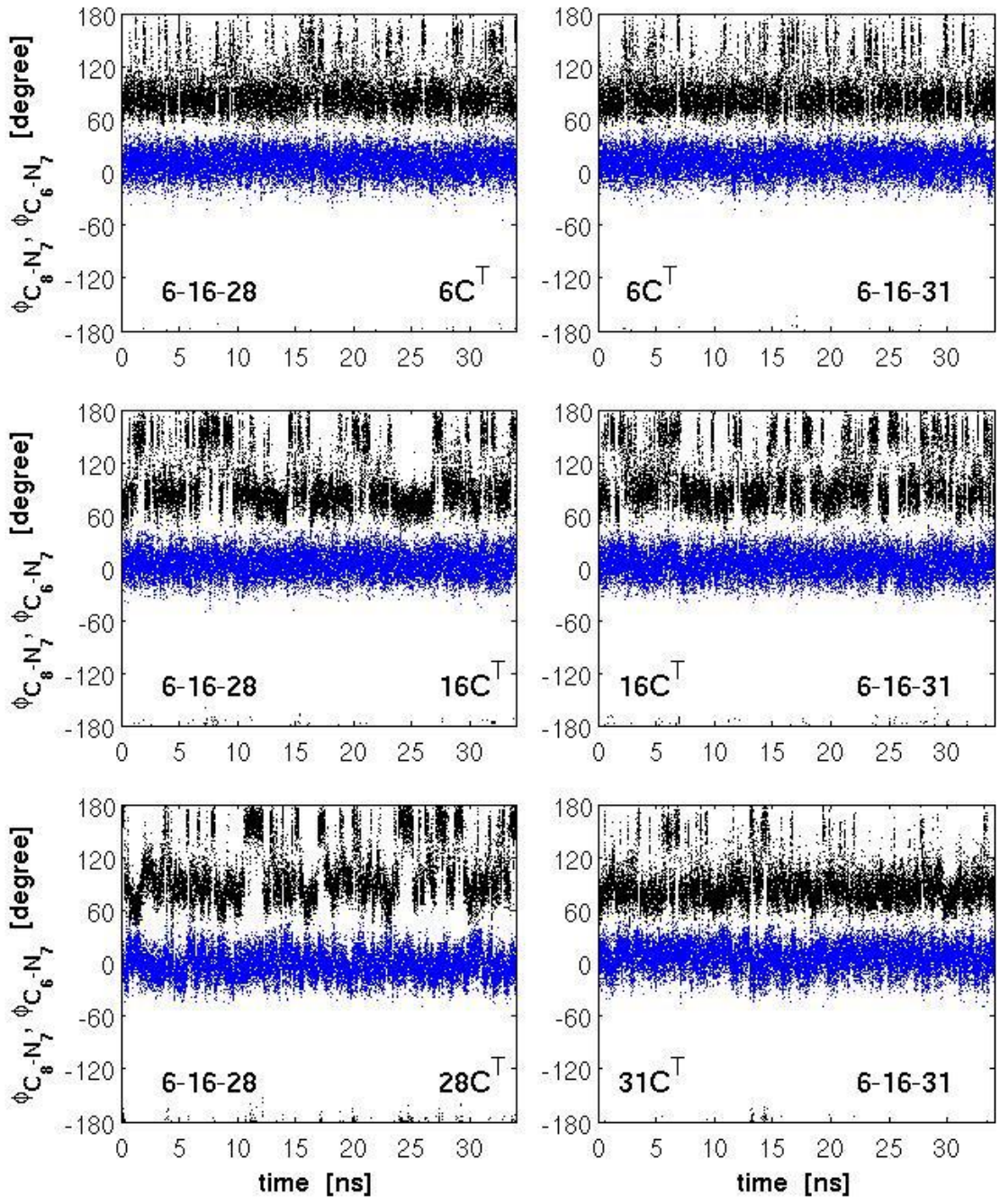


Table S1: Energies and dihedral angles of the six optimal conformations of C^T shown in Fig. 5 of the main text.

Mimimum #	1	2	3	4	5	6
Energy (kcal/mol)	1.96	1.96	5.49	0.00	0.00	2.55
Dihedral ϕ_1 (degree)	162	77	–63	156	83	–63
Dihedral ϕ_2 (degree)	–12	12	0	175	–175	180

S7: Values of the C^T dihedral angles during the 34 ns MD simulation.

Values of the C^T dihedral angles C_8-N_7 (black) and C_6-N_7 (blue) during the 34 ns MD simulation of the A-RNA helices labeled at position 6-16-28 (left) and 6-16-31 (right). From the two rotatable bonds connecting TEMPO to the cytosine base, only the one closest to the TEMPO ring (C_8-N_7) isomerizes. Isomerizations around the other rotatable bond (C_4-N_7), which is closest to the cytosine ring, do not occur since those would interfere with the Watson-Crick base pairing of C^T with the G on the complementary RNA strand (see inset of Fig. 1).



SI4: Restrained MD simulations

In an effort to assess the relative contributions of local and global flexibility to the width of the distance distribution between two spin labels, each construct was simulated for 7 ns by harmonically restraining all the P and C1' atoms of the RNA to the positions of the ideal A-RNA helix with a spring constant of stiffness 0.2 kcal/mol/Å^2. All the other atoms were not restrained. Pinning down the phosphorus atoms prevents any significant distortions of the RNA helix from the initial reference structure. Pinning down the C1' atoms, which connect the bases to the sugar moiety, allows only for small fluctuations of the base in the vicinity of the starting positions. Thus, the restrained simulations should reflect local contributions to the inter-spin distance distribution due to the conformational freedom of the spin label and the thermal vibrations of the base to which it is attached. Any additional broadening of the distance distribution should be ascribed to the larger-scale distortions of the bases and the global flexibility of the RNA helix.

Figure S8: Histograms of the distance distributions obtained from the 7 ns MD simulations of the constructs 6-16-28 (a) and 6-16-31 (b) with harmonically restrained P and C1' atoms.

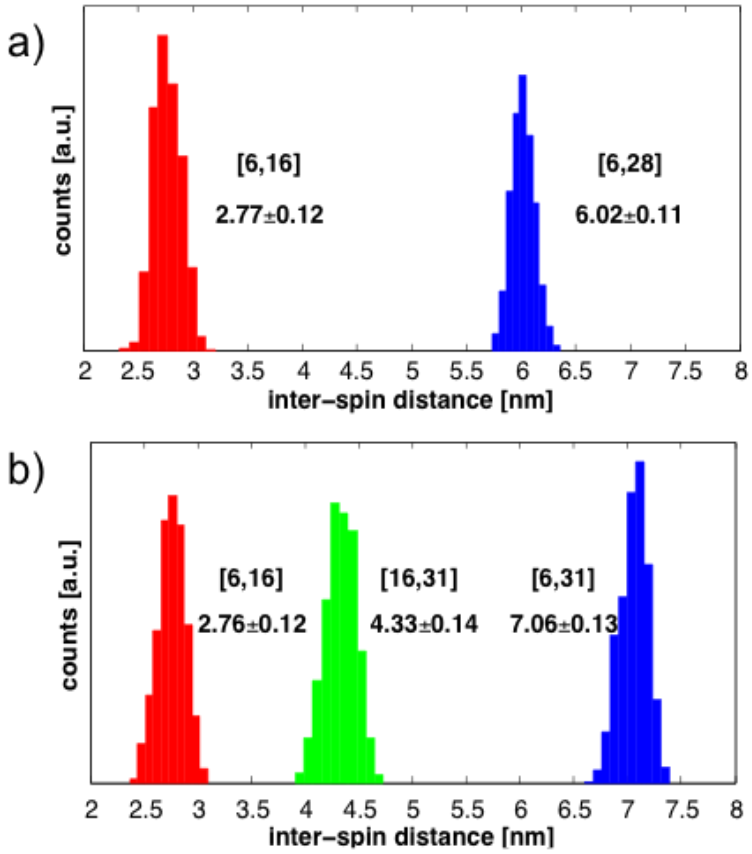
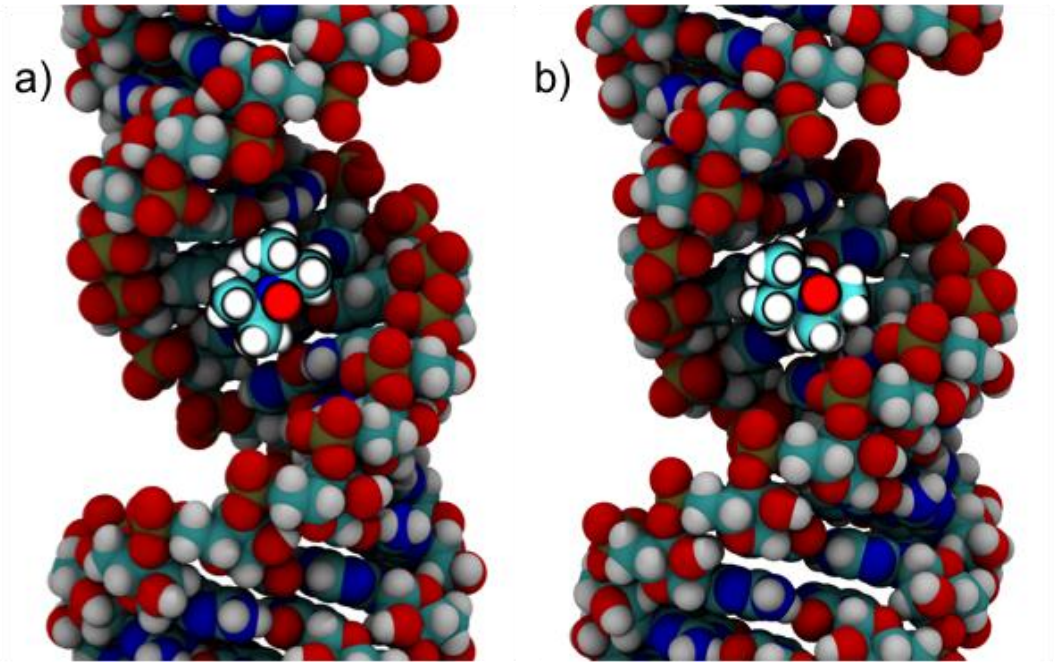


Fig. S9: View of the spin label at position 16 from the MD simulation of 6-16-28.

The position of the TEMPO oxygen with respect to the labeled base is almost unchanged between conformation 2 (a) and conformation 1 (b) of C^T. (Conformation numbers are defined in Fig. 5 of the main text.)



References

1. V. P. Denysenkov, T. F. Prisner, J. Stubbe and M. Bennati, *Proceedings of the National Academy of Sciences*, 2006, **103**, 13386-13390.
2. I. Tkach, S. Pornsuwan, C. Hobartner, F. Wachowius, S. T. Sigurdsson, T. Y. Baranova, U. Diederichsen, G. Sicoli and M. Bennati, *PCCP*, 2013, **15**, 3433-3437.
3. G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger and H. Jung, *Appl. Magn. Reson.*, 2006, **30**, 473-498.