Global Structure of a Three-Way Junction in a Phi29 Packaging RNA Dimer Determined Using Site-Directed Spin Labeling

Xiaojun Zhang[1], Chang-Shung Tung[2], Glenna Z. Sowa[1], Ma'mon M. Hatmal[3], Ian S. Haworth[3,4], and Peter Z Qin[1*]

[1] Department of Chemistry, University of Southern California, Los Angeles, California, 90089, United States

[2] MS K710, T-10, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States

[3] Department of Biochemistry, University of Southern California, Los Angeles, California, 90089, United States

[4] Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, Los Angeles, California 90089, United States

*To whom correspondence should be addressed: LJS-251, 840 Downey Way, Los Angeles, CA 90089-0744; Tel: (213) 821-2461; Fax: (213) 740-0930; Email: pzq@usc.edu.

Supporting Information

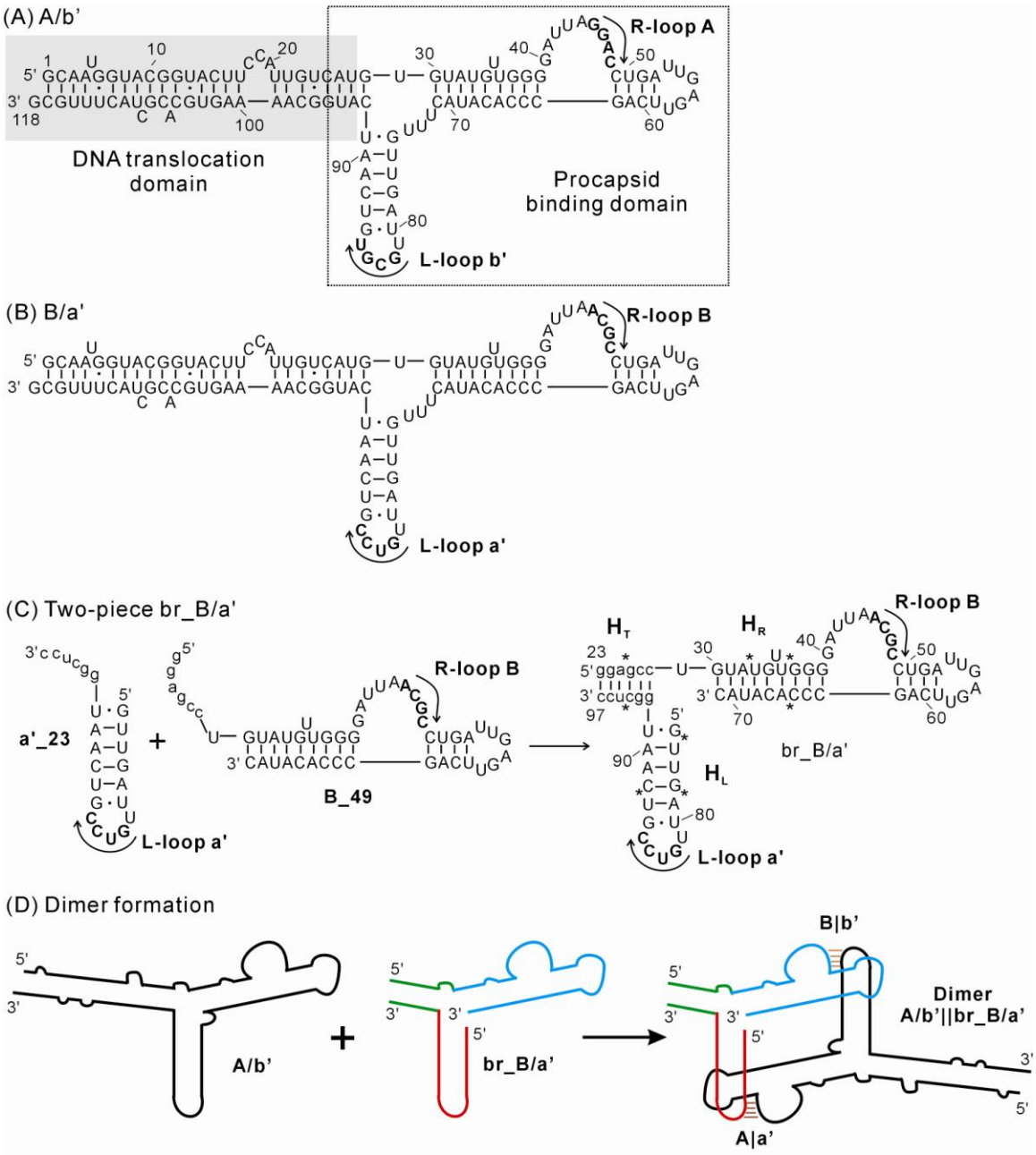


Figure S1: Sequences and secondary structures of pRNA constructs. (A) A/b'. The sequence is numbered as shown, with the R-loop sequence (5'-GGAC-3', designated as A) and the L-loop sequence (5'-GCGU-3', designated as b') shown in bold. Approximate boundaries of the DNA translocation domain and procapsid binding domain are marked. (B) B/a'. R-loop sequence (5'-ACGC-3', designated as B) and L-loop sequence (5'-GUCC-3', designated as a') are shown. (C) br_B/a'. RNA strands a'_23 and B_49 are held together by the H_T helix (shown in lower case letters) to form br_B/a'. "*" indicates R5 labeling sites as shown in main text Figure 1A. (D) Schematic representation of dimer formation. The three sub-modules of br_B/a' are colored in green, red, and blue, respectively. Intermolecular base-pairings between the R- and L-loop (A||a' and B||b') are marked.

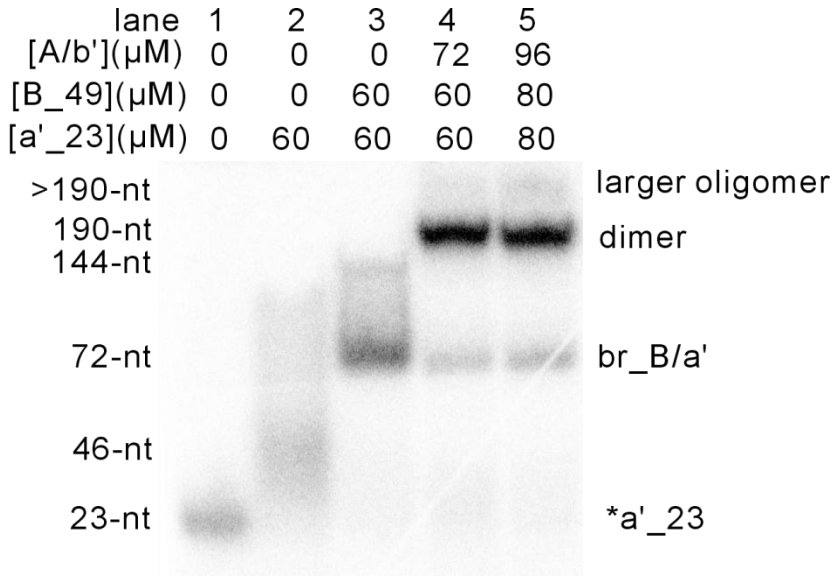


Figure S2: Assembly of pRNAs. Sample in each lane contained 50 mM Tris-HCl (pH 7.6), 3 mM $MgCl_2$, 50% (v/v of the final solution) glycerol, a trace amount of ^{32}P labeled a'_23 (*a'_23), and various amount of unlabeled RNAs as indicated. Lane 1: individual *a'_23 (23-nt) observed at low RNA concentration; Lane 2: homodimer of a'_23 (46-nt) formed in the absence of the B_49 partner at RNA concentration presented under the EPR condition (i.e., 60 μM); Lane 3: formation of the br_B/a' monomer (72-nt) under EPR condition, with a small fraction of br_B/a' homodimer (144-nt) formed in the absence of the A/b' partner; Lane 4 and Lane 5: formation of the A/b'||br_B/a' dimer (190-nt). In lanes 4 and 5, the dimer accounts for > 80% of the total population, while small fraction of br_B/a' monomer (72-nt) and larger pRNA oligomers (> 190-nt) can be detected. Pulsed EPR samples were assembled using RNA concentrations shown in lane 4, as they afford the best balance between the desired pRNA dimer and the undesired monomer and higher oligomers.

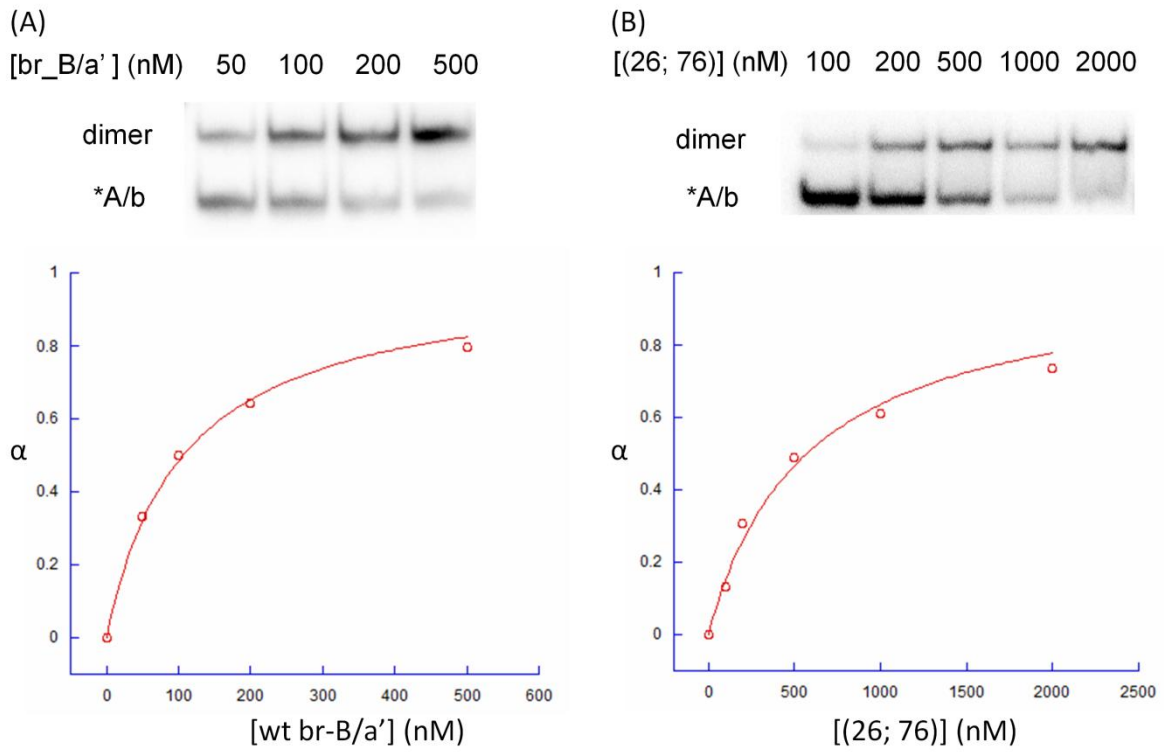


Figure S3: Measurements of pRNA dimer dissociation constant (K_d) for unmodified br_B/a' (panel A) and spin-labeled br_B/a' (26; 76) (panel B). Top: autoradiograph with signal derived from trace amount of radio-labeled *A/b'. As concentrations of br_B/a' increased, monomeric A/b' was observed to shift into dimer. Bottom: Quantitative binding analysis. The fraction of dimer α, were calculated and plotted against the concentration of corresponding br_B/a'. Fitting the shown data to equation 1 in main text (red curves) yielded a K_d value of 106 nM for unmodified br_B/a' and 580 nM for labeled (26; 76).

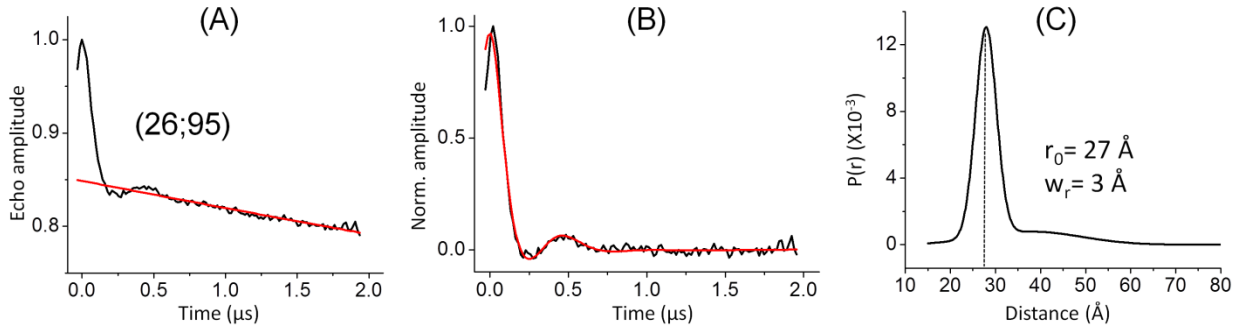


Figure S4: DEER data for a pair of R5's attached to sites 26 and 95, which are located across from each other in the H_T helix (see Figure S1C or Figure 1A in main text). (A) Original echo decay data, with measured echo decay shown in black and background decay in red. (B) Dipolar evolution data, with the normalized background-corrected dipolar evolution data shown in black and the best-fit curve shown in red. (C) Distance distribution profile, with the most probable distance (r_0) and the width of distribution (w_r) indicated. The measured distance of 27 Å is consistent with the expected value from an A-form RNA duplex (25.5 Å), thus demonstrating assembly of the 2-piece br_B/a'.

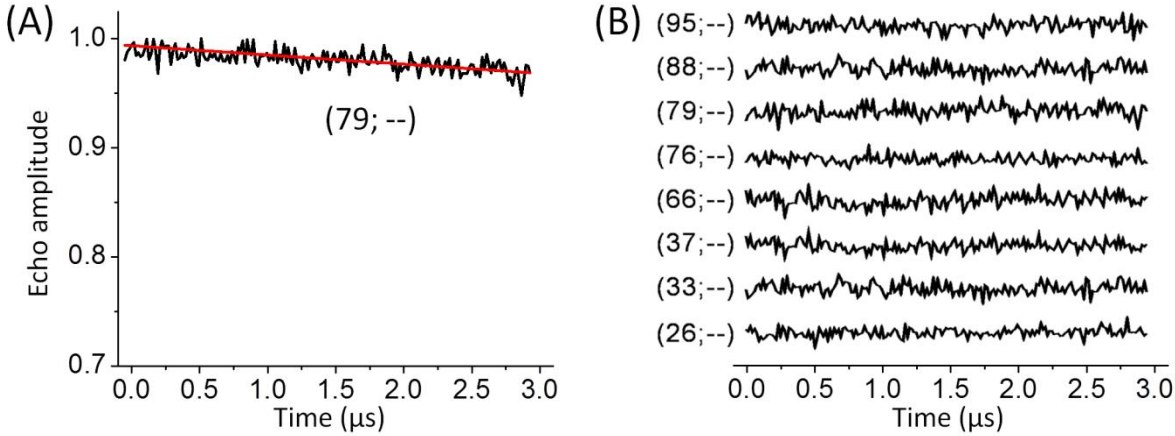
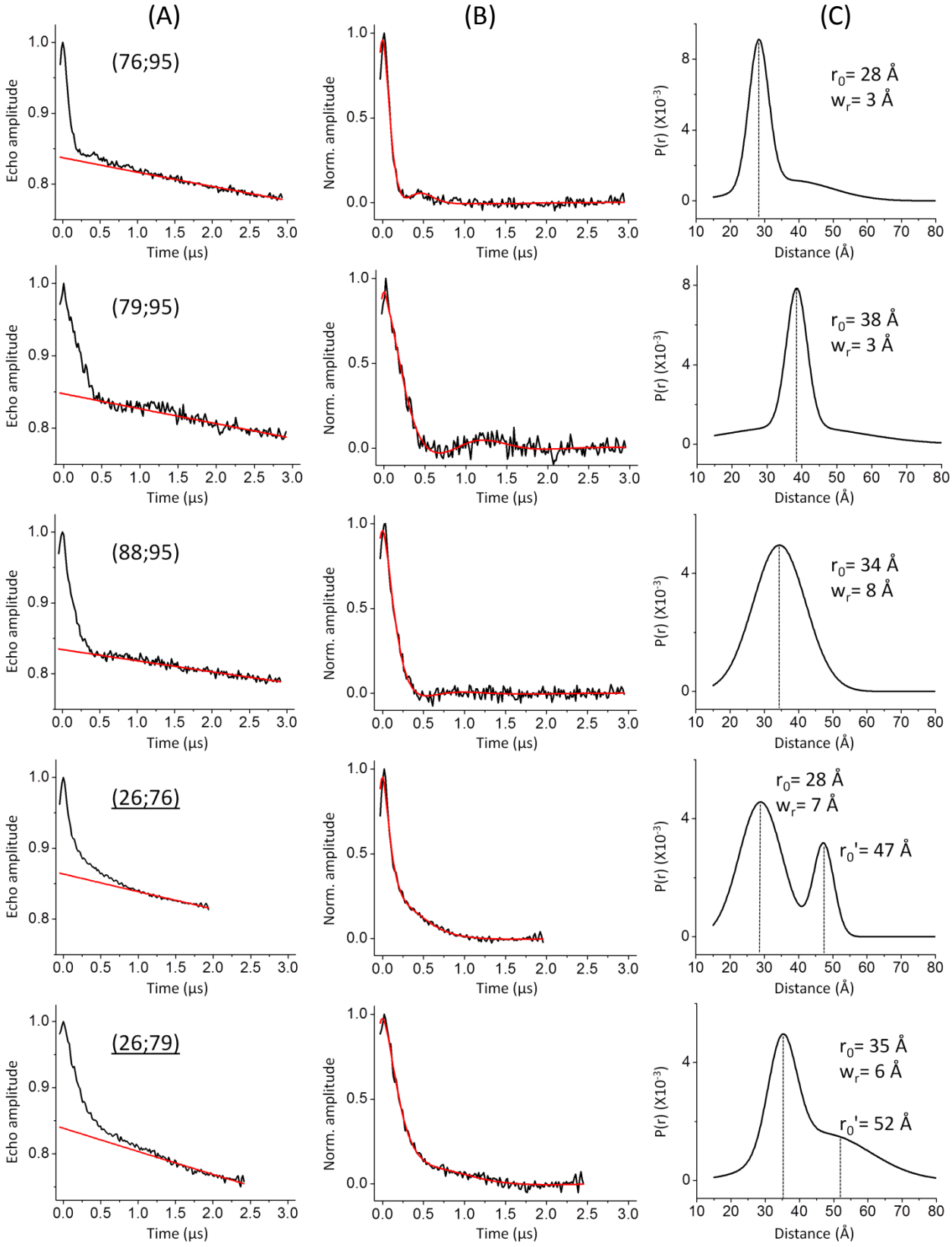
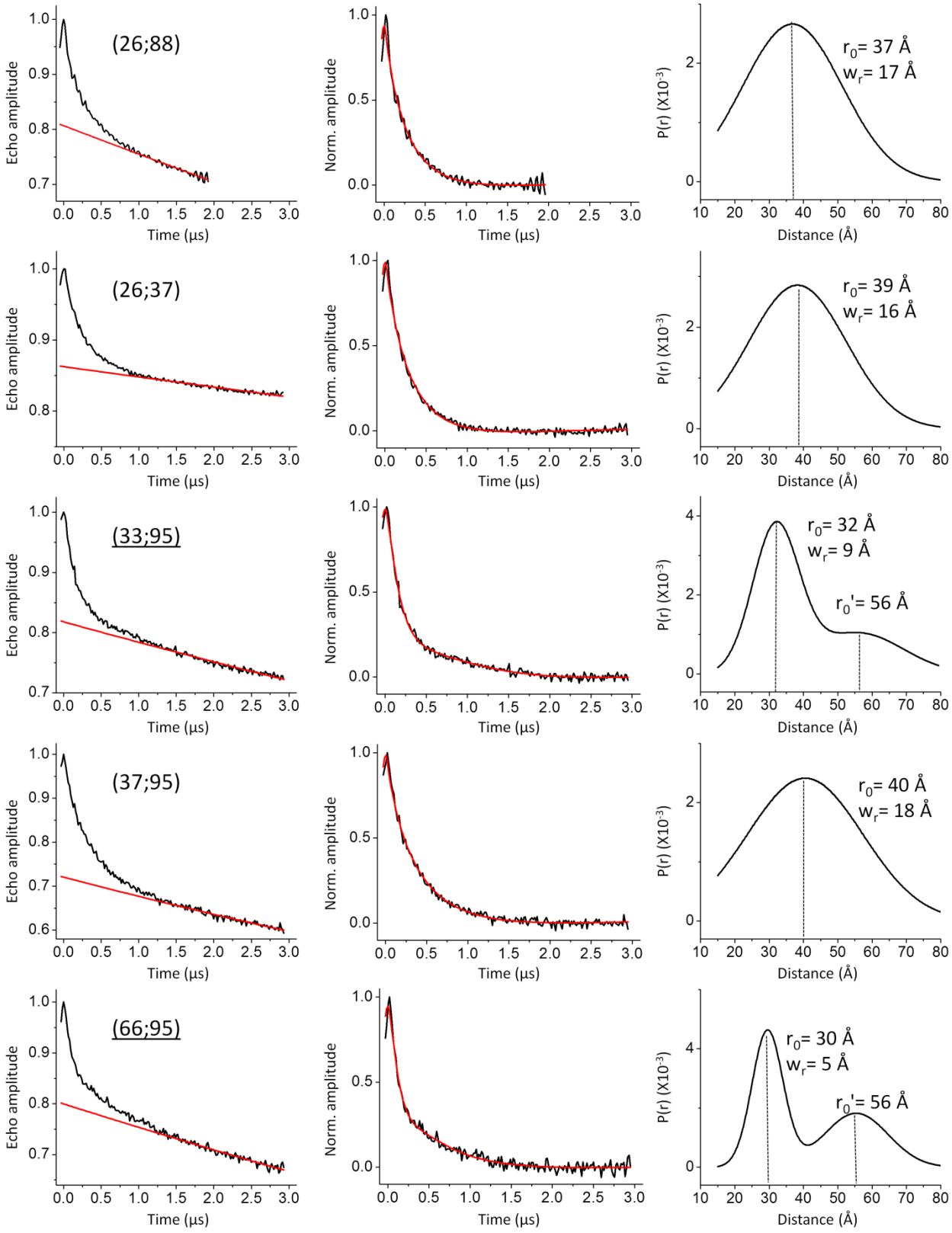
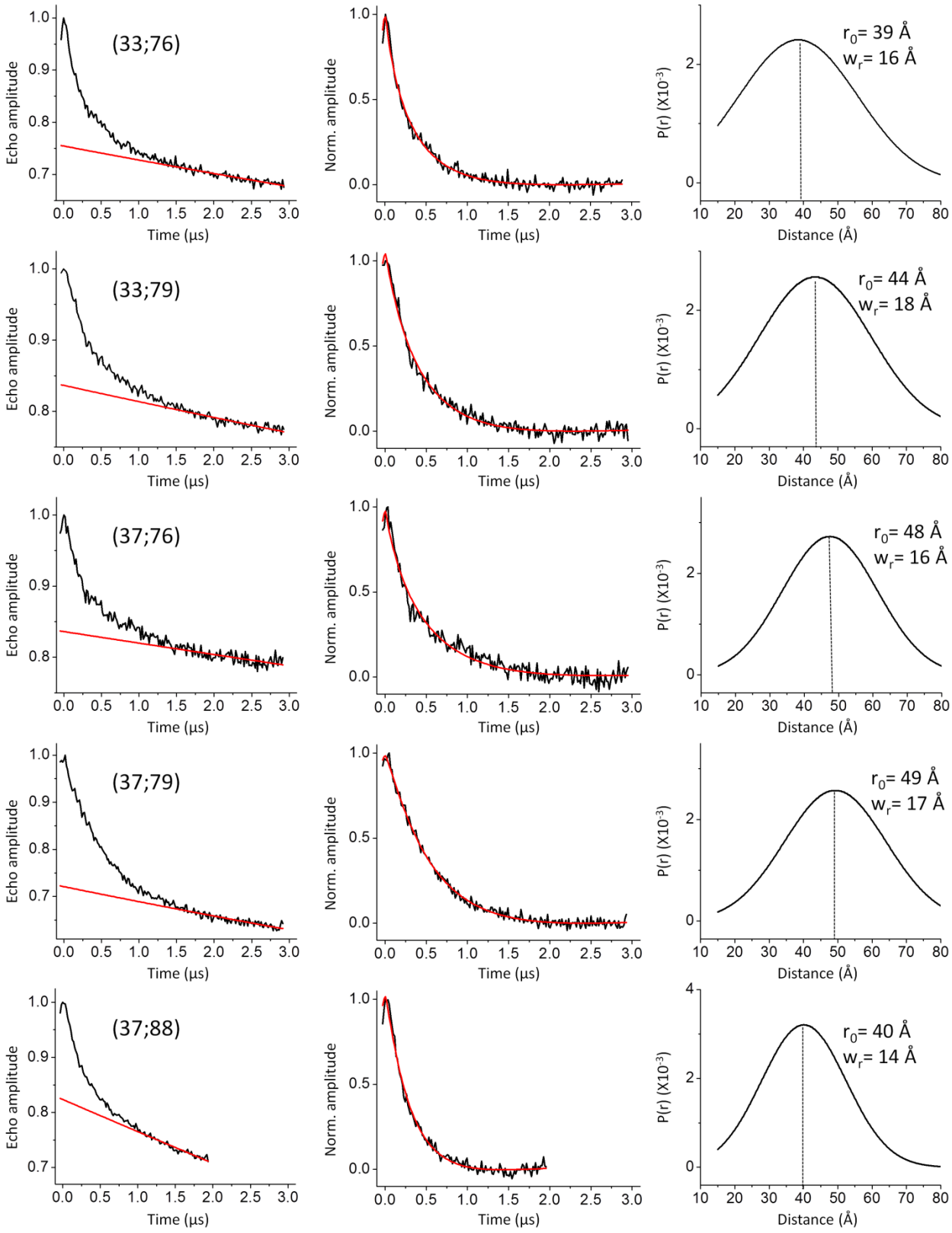


Figure S5: (A) An example of DEER measurement on a single-labeled pRNA dimer. The spin label was attached to site 79, and the sample was designated as (79; --) (see Figure S1C or Figure 1A in main text). The black trace represents the measured echo decay, and the red trace represents the background computed by fitting an exponential decay corresponding to a homogeneous 3-dimentional distribution of electron spin to the last half of the data (see Methods). (B) Background-corrected dipolar evolution curve for all single-labeled pRNA dimer samples used in the study. No decay or oscillation is observed after background correction, indicating a lack of random aggregation of pRNA or spin-spin interactions between different pRNA dimers.


(A)
(B)
(C)
(76;95)
(79;95)
(88;95)
(26;76)
(26;79)
Echo amplitude
Norm. amplitude
P(r) (X10-3)
Time (µs)
Distance (Å)
r0= 28 Å
wr= 3 Å
r0= 38 Å
wr= 3 Å
r0= 34 Å
wr= 8 Å
r0= 28 Å
wr= 7 Å
r0'= 47 Å
r0= 35 Å
wr= 6 Å
r0'= 52 Å


(26;88)
Echo amplitude
Time (μs)
Norm. amplitude
P(r) (X10-3)
r0= 37 Å
wr= 17 Å
Distance (Å)
(26;37)
r0= 39 Å
wr= 16 Å
(33;95)
r0= 32 Å
wr= 9 Å
r0'= 56 Å
(37;95)
r0= 40 Å
wr= 18 Å
(66;95)
r0= 30 Å
wr= 5 Å
r0'= 56 Å


(33;76)
Echo amplitude
Time (µs)
Norm. amplitude
P(r) (X10-3)
r0= 39 Å
wr= 16 Å
Distance (Å)
(33;79)
r0= 44 Å
wr= 18 Å
(37;76)
r0= 48 Å
wr= 16 Å
(37;79)
r0= 49 Å
wr= 17 Å
(37;88)
r0= 40 Å
wr= 14 Å

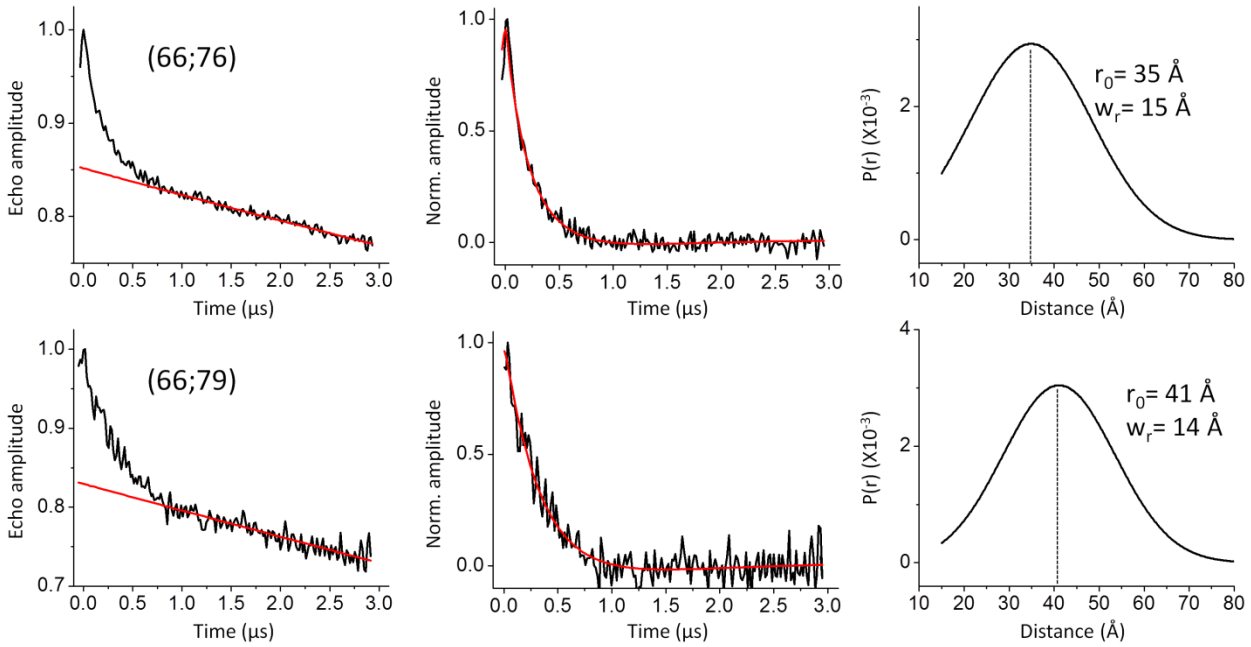


Figure S6: DEER data and analyses. Column (A) is the original echo decay data, with measured echo decay shown in black and background decay in red. Column (B) is the dipolar evolution data, with the normalized background-corrected dipolar evolution data shown in black and the best-fit curves shown in red. Column (C) is the distance distribution profiles computed using Defit, with the most probable distance (r_0) and the width of distribution (w_r) indicated. Each distance set is designated by the labeling sites (see Figure 1A in main text). The 4 data sets yielding distance distribution profiles with two populations are underlined.

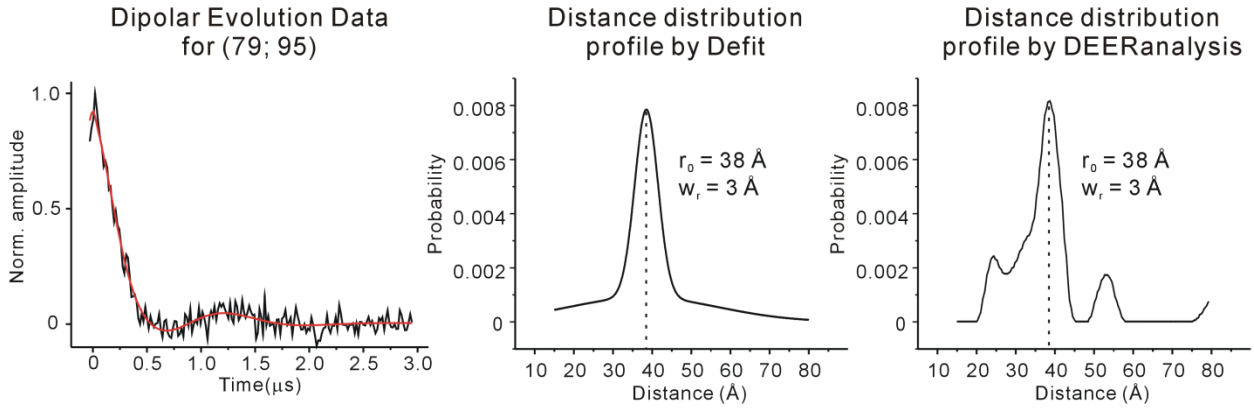


Figure S7: Analysis of data set (79; 95) with either Defit or DeerAnalysis. The resulting distance distribution profiles were consistent. The w_r value for the distance distribution profile by DeerAnalysis shown here is the half-width at half-maximum for the central major peak.

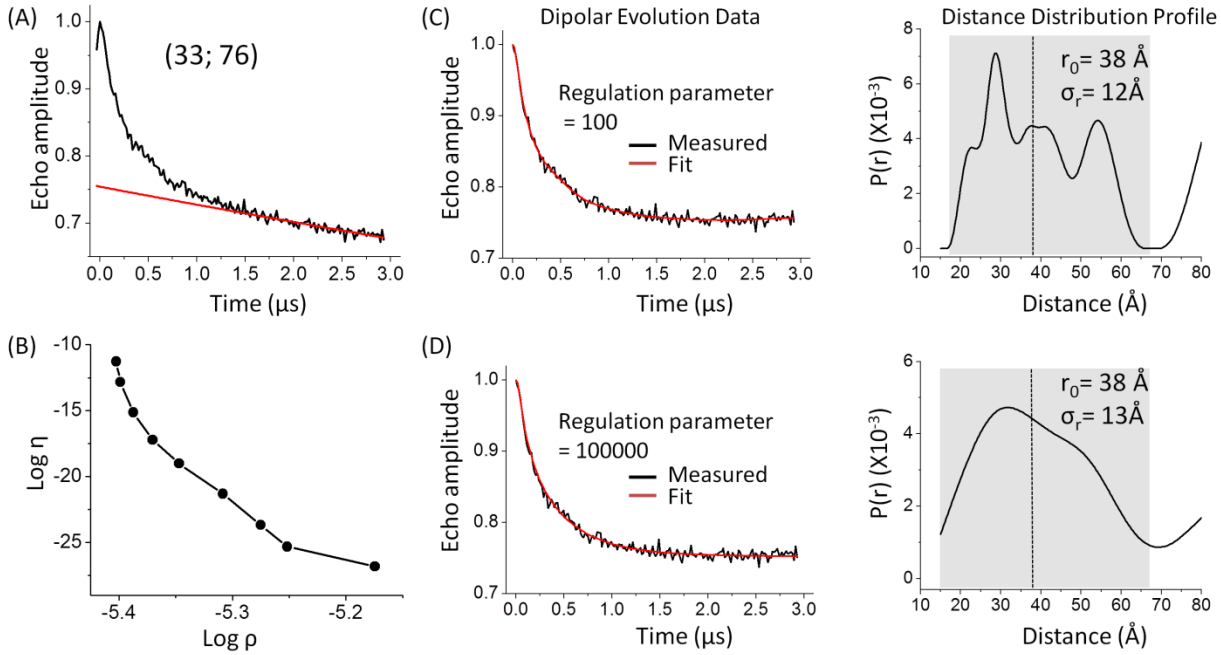


Figure S8: Analysis of data set (33; 76) with DeerAnalysis. (A) Normalized experimental echo evolution data (black) and background signal (red). (B) The L-curve generated by the Tikhonov fit. The lack of a clear turning point indicates that an optimized regulation parameter cannot be defined. (C) The best-fit trace (red) and the distance distribution P(r) obtained using a regulation parameter of 100. The mean distance (r_0) and standard deviation (σ_r) were calculated as described (Cai, Q., *et. al*., *Nucleic Acids Res* 2006, 34, 4722-30.) using lower and upper boundary defined by the shaded area. (D) The best-fit trace (red) and the distance distribution P(r) obtained using a regulation parameter of 100000.

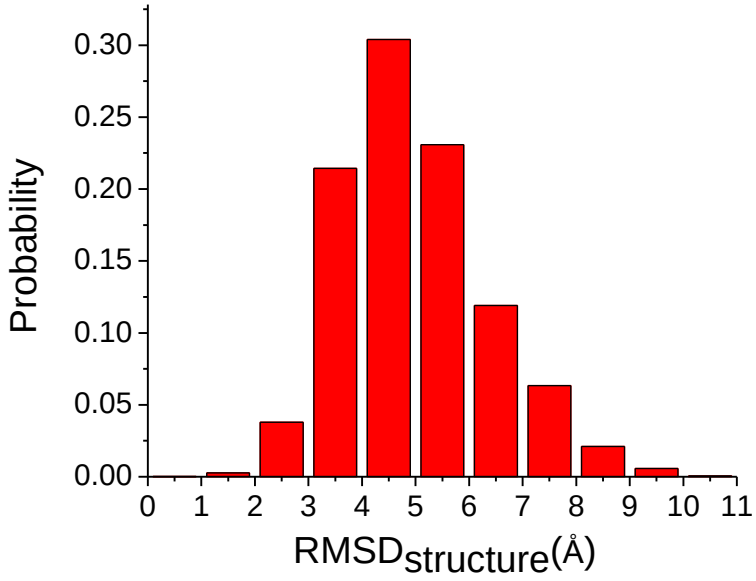


Figure S9: Distribution of heavy atom root-mean-square-deviations ($RMSD_{structure}$) between the best-fit model (Figure 3A, main text) and each viable model obtained by the grid search (see Methods).

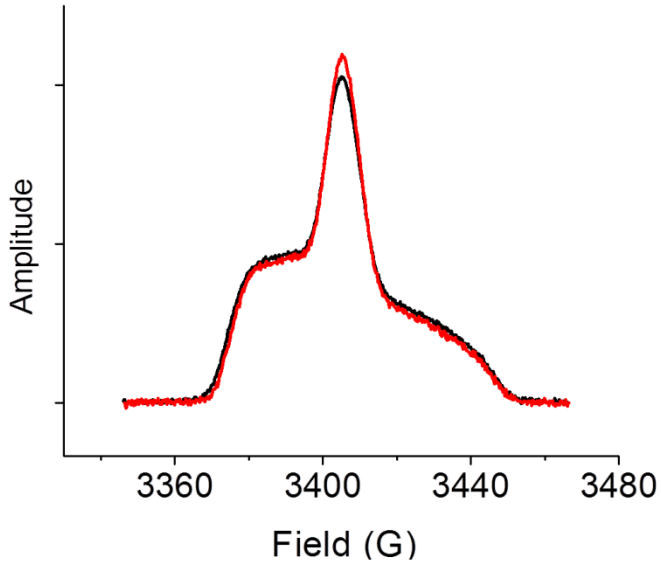


Figure S10: Normalized absorption spectra (field sweep) of the double-labeled (26; 88) (black) and the sum of its single-labeled controls (red). The two spectra show minimal difference and thus there is no short distance present (i.e., 10.7 Å as predicted based on the pRNA tetramer crystal structure).

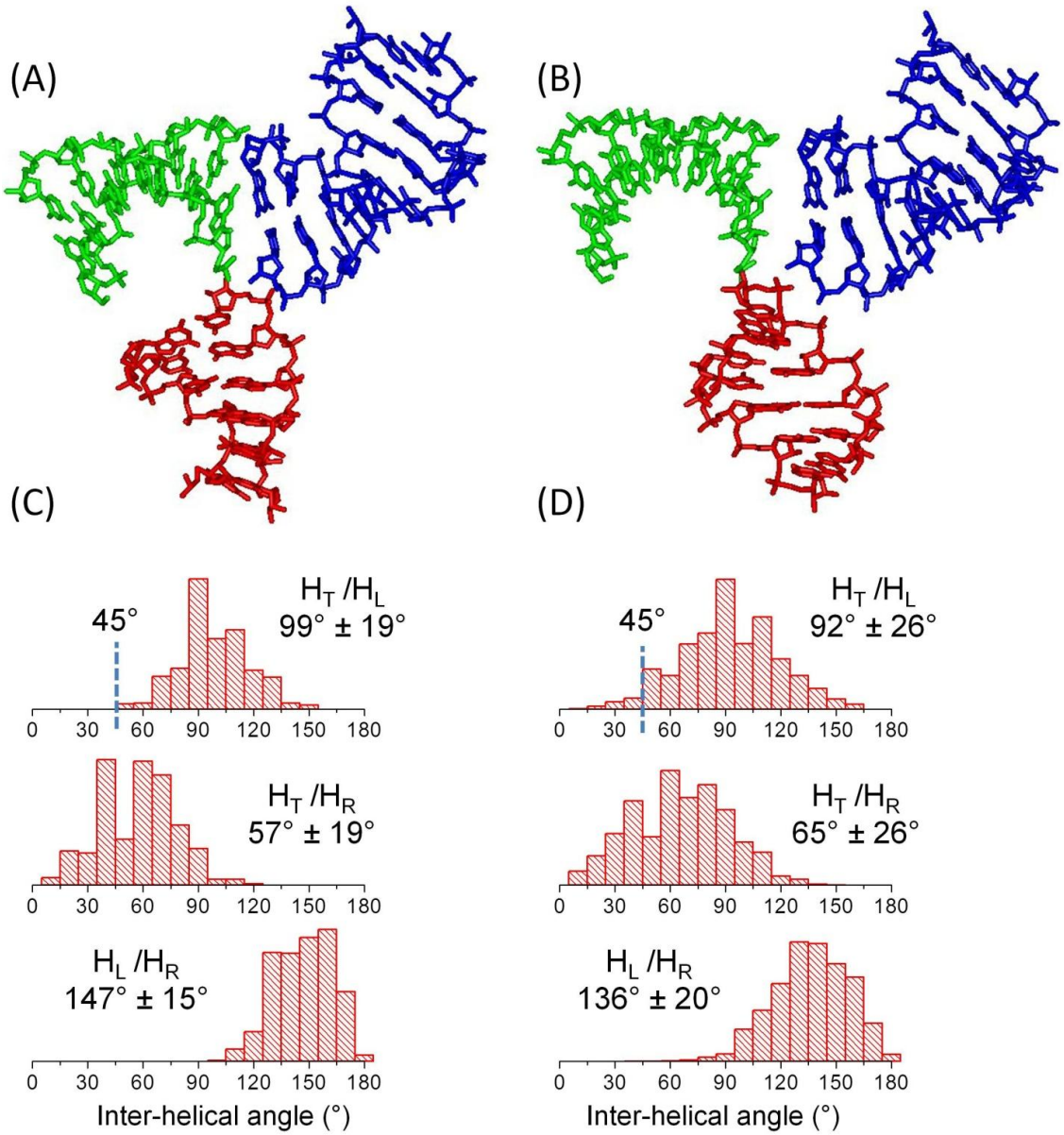


Figure S11: Comparison of models obtained using different number of DEER distances. Panels (A) and (C) show, respectively, the top-ranked model and the inter-helical angle distributions of the viable models obtained from grid searches using all 17 sets of DEER distances. Panels (B) and (D) show the corresponding results obtained with 12 of the 17 distances while omitting data sets (26; 37), (26; 76); (26; 79), (26, 88) and (37; 88). Comparisons between panels (A) and (B) reveal that the top-ranked models are very similar: inter-helical angles ($\theta_{T,L}$, $\theta_{T,R}$, $\theta_{L,R}$) are (93°, 57°, 147°) for (A) and (94°, 65°, 143°) for (B); and the heavy atom root-mean-square-deviation between the two models is 3.5 Å. Comparisons between (C) and (D) show the average inter-helical angles are very similar, although searches using fewer DEER distances (i.e., panel (D)) give broader distributions. Particularly, panel (D) shows 3.6% of the populations have $\theta_{T,L} < 45°$, while these more linear H_T/H_L configurations are absent in panel (C).

Table S1: Widths of the distance distribution obtained from Defit analyses.

Data Set		Distance Distribution Width w_r (Å)
H_T *vs.* H_L	(76; 95)	3
	(79; 95)	3
	(88; 95)	8
	(26; 76)	7
	(26; 79)	6
	(26; 88)	17
H_T *vs.* H_R	(26; 37)	16
	(33; 95)	9
	(37; 95)	18
	(66; 95)	5
H_R *vs.* H_L	(33; 76)	16
	(33; 79)	18
	(37; 76)	16
	(37; 79)	17
	(37; 88)	14
	(66; 76)	15
	(66; 79)	14