Supplementary Information:

Conformational dynamics in extended RGD-containing peptides

William R. Lindemann, Alexander J. Mijalis, José Luis Alonso, Peter P. Borbat, Jack H. Freed, M. Amin Arnaout, Bradley L. Pentelute, Julia H. Ortony

Table of Contents

I. Peptide Synthesis

LC-MS (liquid chromatography–mass spectrometry) elution profiles of each set of peptides, as well as the peaks used for identification, are displayed in Figures S1-3.

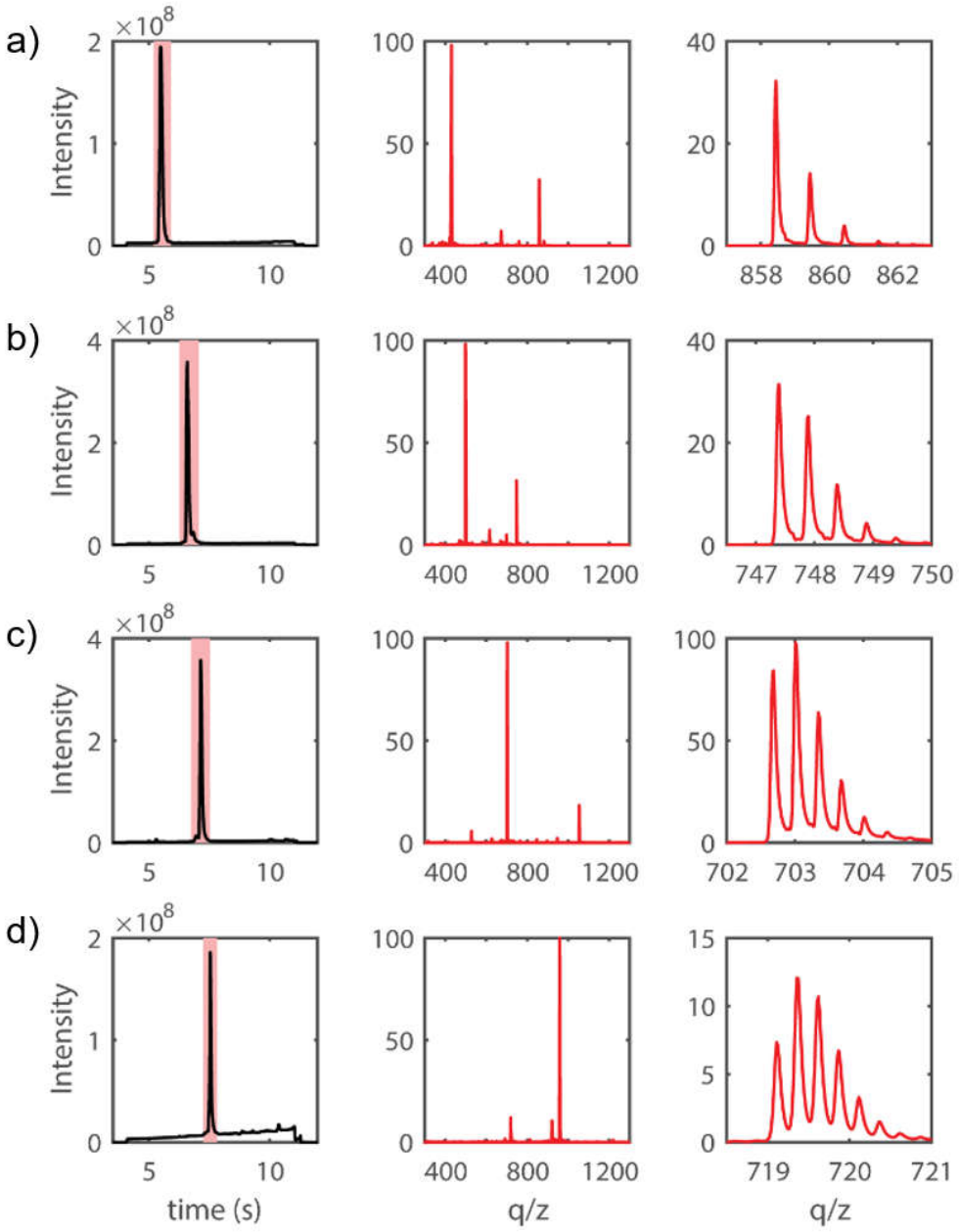

Figure S1. LC-MS elution profiles (left) and mass-to-charge (m/z) data integrated over the highlighted region (middle, right) for a) FMP9, b) FMP15, c) FMP21, and d) FMP27. The specific peaks used for mass determination were the FMP9 $[M+H]^{1+}$ peak-set, the FMP15 $[M+2H]^{2+}$ peak-set, the FMP21 $[M+3H]^{3+}$ peak-set and the FMP27 $[M+4H]^{4+}$ peak-set. These are shown in high-resolution on the right.

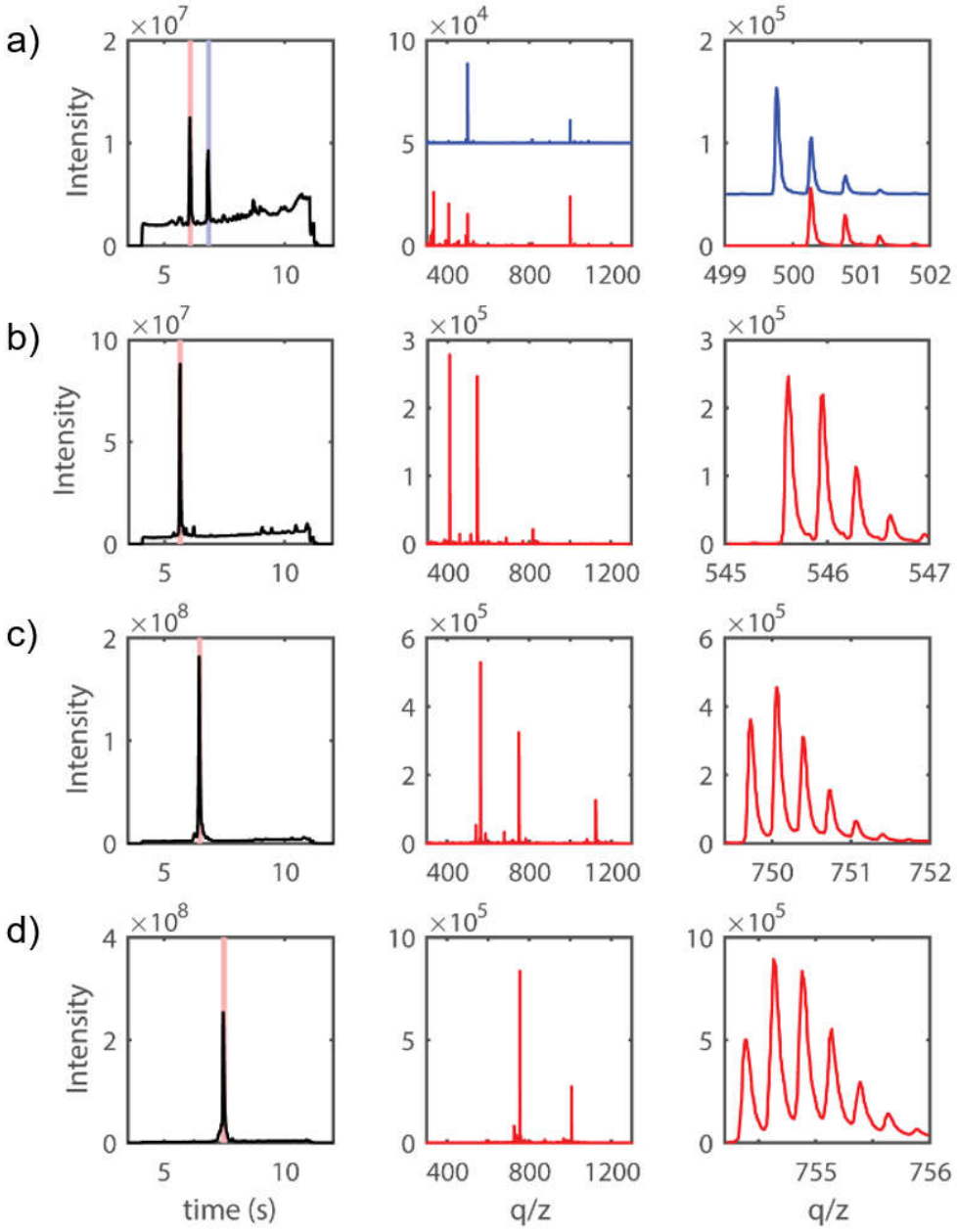

Figure S2. LC-MS elution profiles (left) and mass-to-charge (m/z) data integrated over the highlighted region (middle, right) for a) FMP9-J5, b) FMP15-J8, c) FMP21-J11, and d) FMP27-J14. The specific peaks used for mass determination were the FMP9-J5 $[M+2H]^{2+}$ peak-set, the FMP15-J8 $[M+3H]^{3+}$ peak-set, the FMP21-J11 $[M+3H]^{3+}$ peak-set and the FMP27-J14 $[M+4H]^{4+}$ peak-set. These are shown in high-resolution on the right. The observed masses typically deviate by 1 hydrogen from the mass of the TOAC peptides because TOACs reduce to hydroxylamines (+1 H atom) under the acidic conditions on the column. They were re-oxidized during experiments, so this did not affect other measurements. In FMP9-J5 both forms were observed as separate peaks. The nitroxide peak is plotted/highlighted in blue, while the hydroxylamine peak is plotted/highlighted in red.

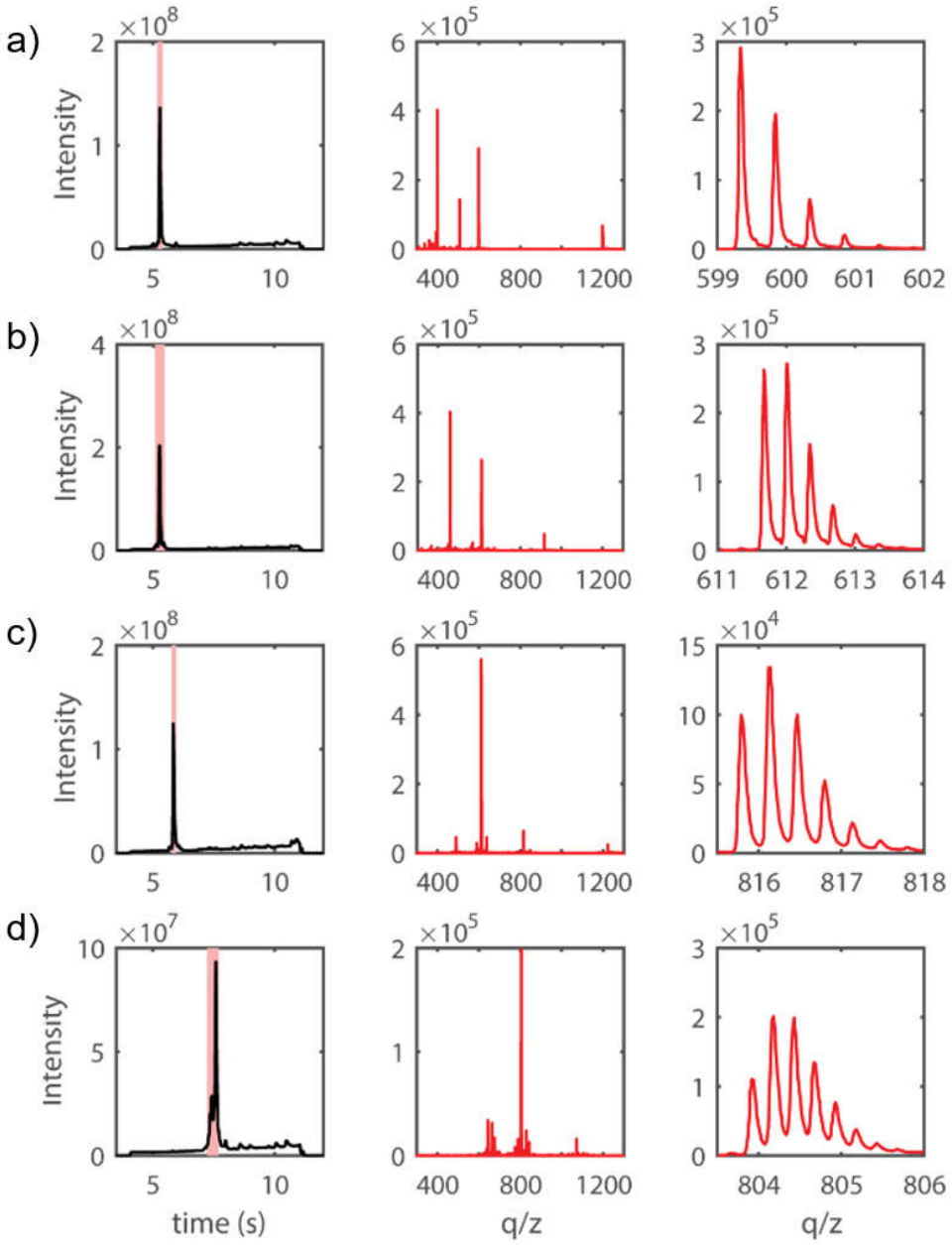

Figure S3. LC-MS elution profiles (left) and mass-to-charge (m/z) data integrated over the highlighted region (middle, right) for a) FMP9-J0J5, b) FMP15-J0J8, c) FMP21-J0J11, and d) FMP27-J0J14. The specific peaks used for mass determination were the FMP9-J0J5 $[M+2H]^{2+}$ peak-set, the FMP15-J0J8 $[M+3H]^{3+}$ peak-set, the FMP21-J0J11 $[M+3H]^{3+}$ peak-set and the FMP27-J0J14 $[M+4H]^{4+}$ peak-set. These are shown in high-resolution on the right. The observed masses deviate by 2 hydrogens from the mass of the TOAC peptides because TOACs reduce to hydroxylamines (+1 H atom) under the acidic conditions on the column. They were re-oxidized during experiments, so this did not affect other measurements.

II. Peptide Structure Simulations

Conformation of the RGD site of each peptide was determined by 200 ns MD simulations. Conformation of the corresponding RGD fragment of fibronectin, determined by brief MD simulation (25 ns), are shown for comparison.

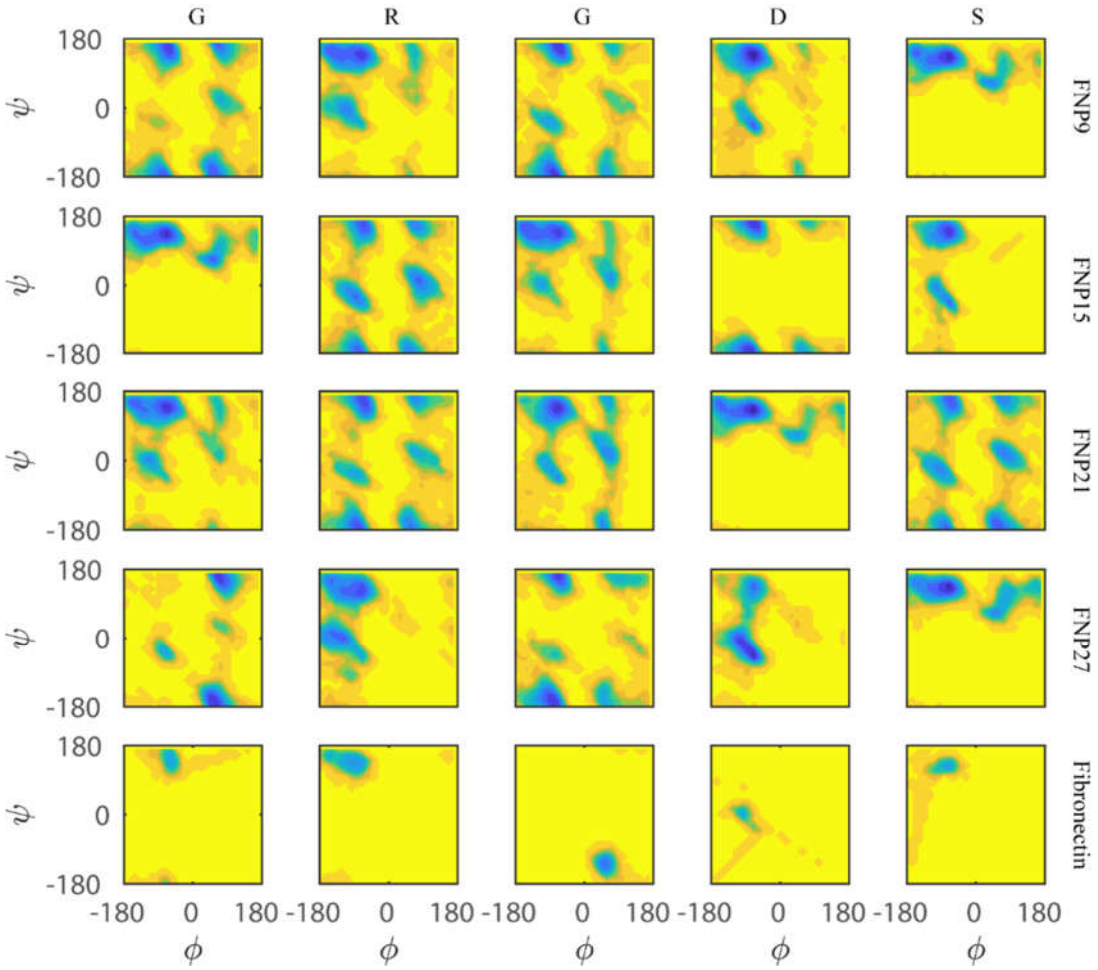

Figure S4. Ramachandran plots of the RGD site in 200 ns peptide simulations and a 25 ns (brief) simulation of Fibronectin, labeled by amino acid (top) and by sequence name (right). The x- and y-axes represent the dihedral angles, while the color axis represents the energy associated with a given conformation within a residue. These plots demonstrate the conformational flexibility of the RGD-loop in all four of the peptides (since many conformations are stable), but affirm that the correct binding conformation (observed in the fibronectin row) is likely accessed by all four of the shorter peptides. Fibronectin likely has other stable conformations, but these are not reflected in the 1FNF crystal structure that we studied.

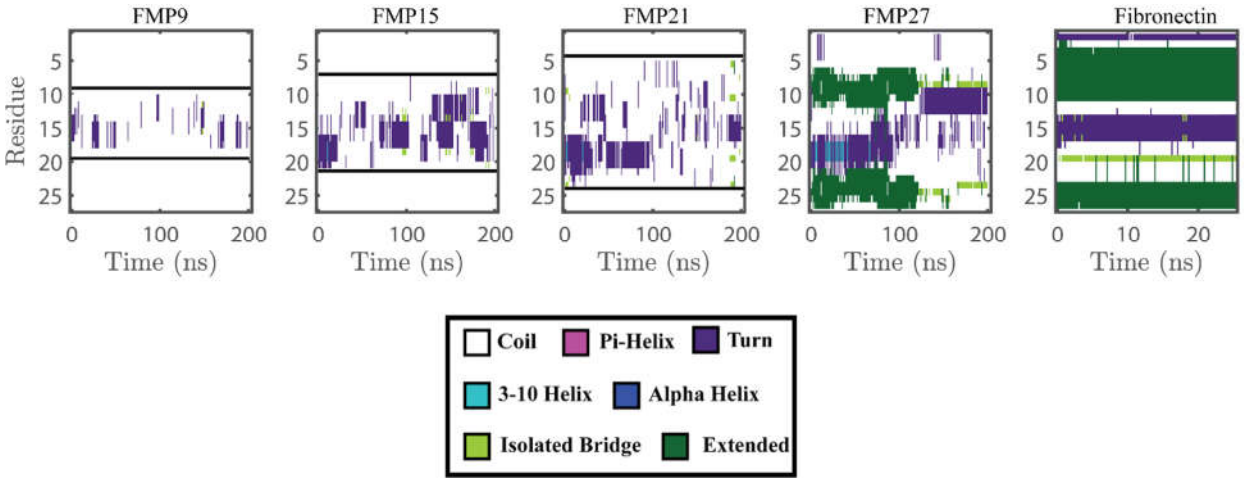

Figure S5. Computed secondary structure of peptides as a function of time. In the shortest peptide, no long-term conformations are stable beyond random coils. However, in the longer 3 peptides, secondary structure emerges which is stable on the order of 5-10 ns. These figures suggest that the longest peptide, FMP27, probably doesn't adequately explore its conformation space within the brief span of our simulations, but the distance distribution extracted from this simulation still compares favorably with the DEER distribution, suggesting that the behavior observed in this simulation is probably fairly representative of the overall behavior of the peptide. We attempted to confirm the presence of an extended region by circular dichroism (CD) analysis, but we didn't observe the presence of any notable β-sheet like peaks. Calculations using DichroCalc[1] suggest that the extended region of FMP27 isn't long enough to significantly change the signal in the 190-250nm region, and attempts to analyze this by FTIR were unsuccessful due to aggregation of FMP27 at the prerequisite concentration.

III. Fibronectin Disorder Prediction

We predicted the disorder associated with the RGD site in fibronectin using the 1FNF sequence and the PrDOS prediction tool.[2]

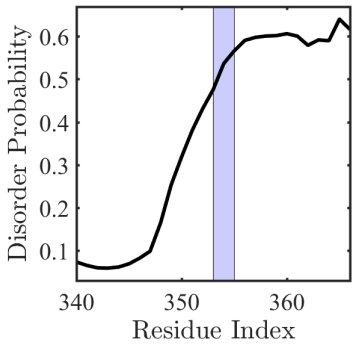

Figure S6. Predicted disorder probability for each amino-acid residue in human fibronectin, featuring the RGD site (blue), computed by PrDOS[2]. Disorder probability ranges from 0 to 1, so the algorithm predicts that the RGD region is likely disordered.

IV. EPR Fits

EPR spectra were collected with a TOAC (2,2,6,6-tetramethyl-N-oxyl-4-amino-4-carboxylic acid) spin label positioned at the glycine of the RGD site. These EPR spectra and corresponding fits used to calculate dynamics of each sample are shown in Fig. S7.

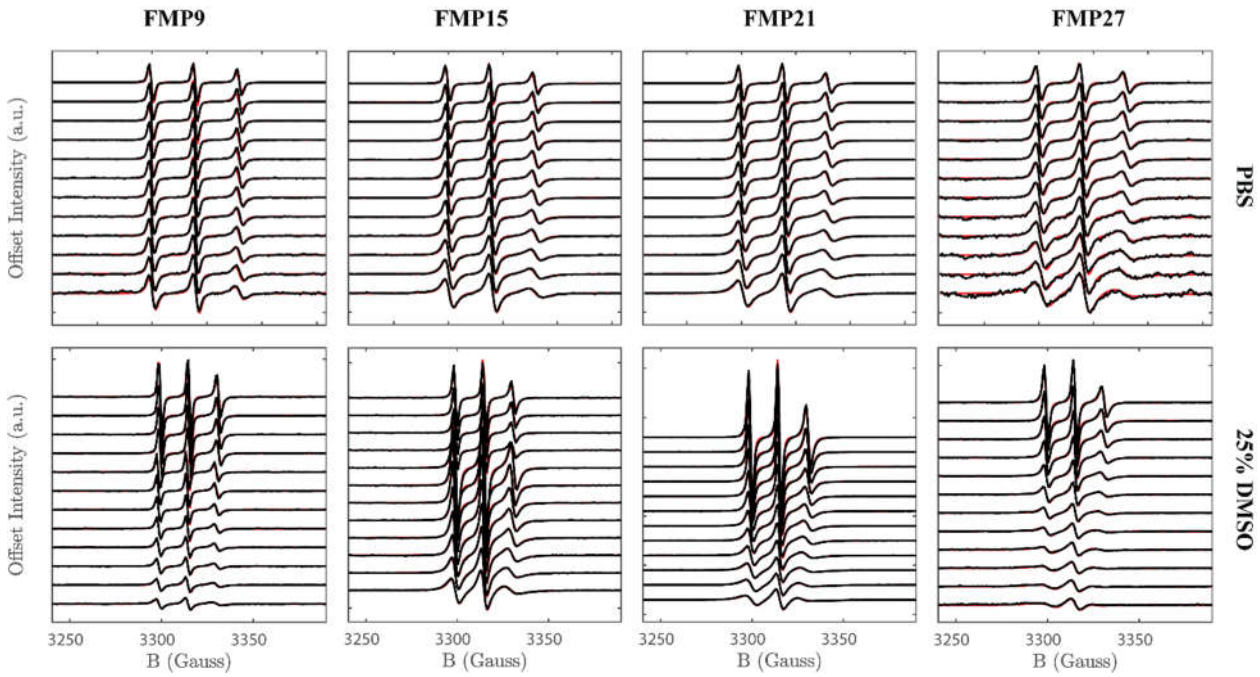

Figure S7. CW-EPR data (black) superimposed over spectral best-fits (red) for each sample. Spectra are vertically offset based on their temperature, so that spectra collected at higher temperatures have a higher vertical offset. For associated dynamic data, refer to Figure 2.

V. Double electron-electron resonance (DEER)

According to the protocol outlined in the Materials and Methods section, we computed distance distributions from the time-domain DEER data presented in Figure S8. We also present an overlay of the DEER distance distributions from Figure 4, to allow easy visual comparison. DEER experiments were performed using a $\pi/2$-t_1-π -t_2-π pulse sequence with $\pi/2$- and π-pulses having widths of 16 and 32 ns, which was applied at the low-field side of the nitroxide spectrum.

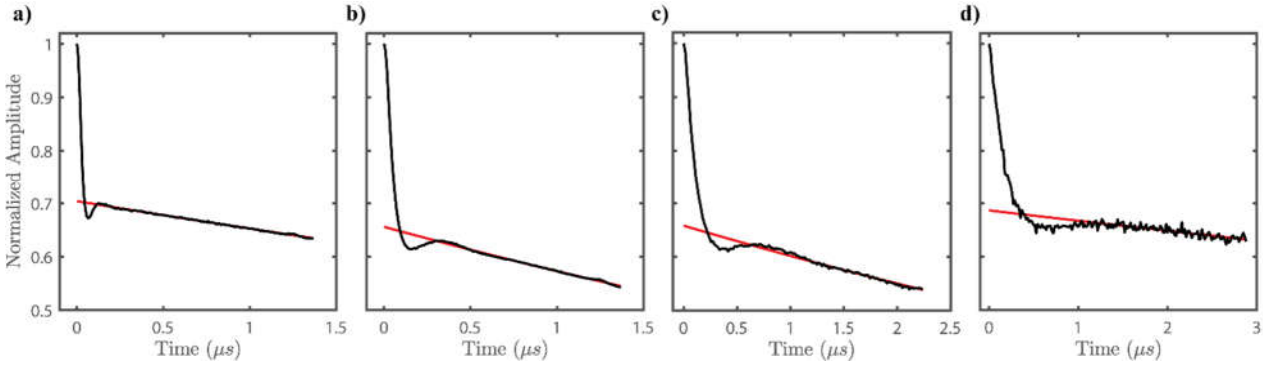

Figure S8. Raw time-domain data (black) and fitted backgrounds (red) for DEER experiments on a) FMP9-J0J5, b) FMP15-J0J8, c) FMP21-J0J11, and d) FMP27-J0J14. Used to compute the observed distance distributions presented in Figure 4. The Y axis is the signal amplitude, normalized to the amplitude at the first timepoint.

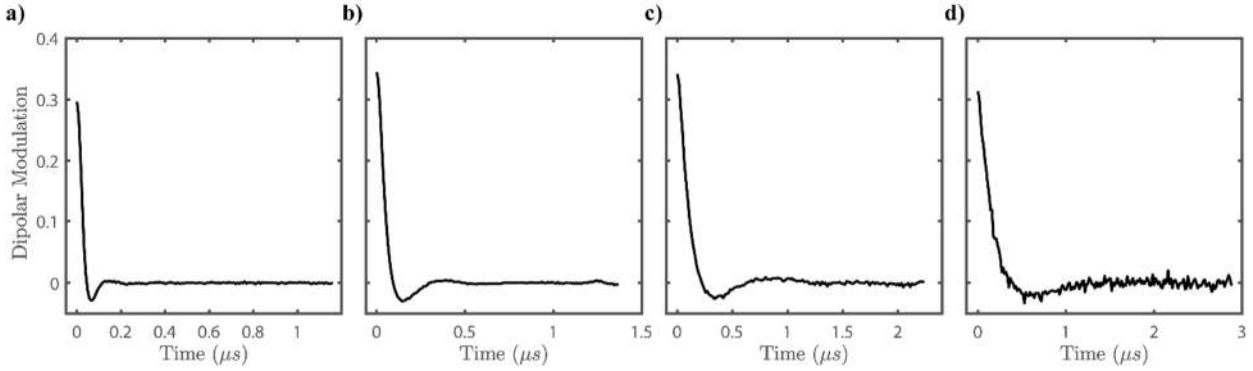

Figure S9. Background-subtracted data from DEER experiments on a) FMP9-J0J5, b) FMP15-J0J8, c) FMP21-J0J11, and d) FMP27-J0J14. Used to compute the observed distance distributions presented in Figure 4. After background removal from Figure S8, remaining signal corresponds precisely to the dipolar modulation observed in the sample.

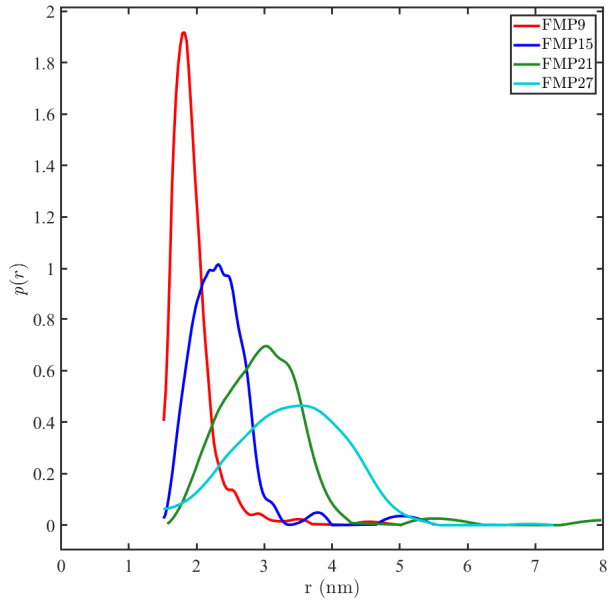

Figure S10. Distance distributions for FMPs extracted by DEER. These are identical to the black curves reported in Figure 4, but are visually overlaid to demonstrate the change in inter-probe distance with length. $p(r)$ is the probability density function obtained by fitting the dipolar modulation observed in Figure S9 using L-curve Tikhonov regularization.

VI. References

1. Bulheller, B. M. & Hirst, J. D. DichroCalc — circular and linear dichroism online. *Bioinformatics* **25**, 539–540 (2009).

2. Ishida, T. & Kinoshita, K. PrDOS: Prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* **35**, 460–464 (2007).