Supporting Information

Pulsed ESR dipolar spectroscopy for distance measurements in immobilized spin labeled proteins in liquid solution.

Zhongyu Yang[1], Yangping Liu[2], Peter Borbat[3], Jay L. Zweier[2], Jack H. Freed[3] and Wayne L. Hubbell[1]

[1]Jules Stein Eye Institute and Departments of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, United States.

[2]Center for Biomedical EPR Spectroscopy and Imaging, The Davis Heart and Lung Research Institute, the Division of Cardiovascular Medicine, Department of Internal Medicine, The Ohio State University, Columbus, Ohio 43210, United States.

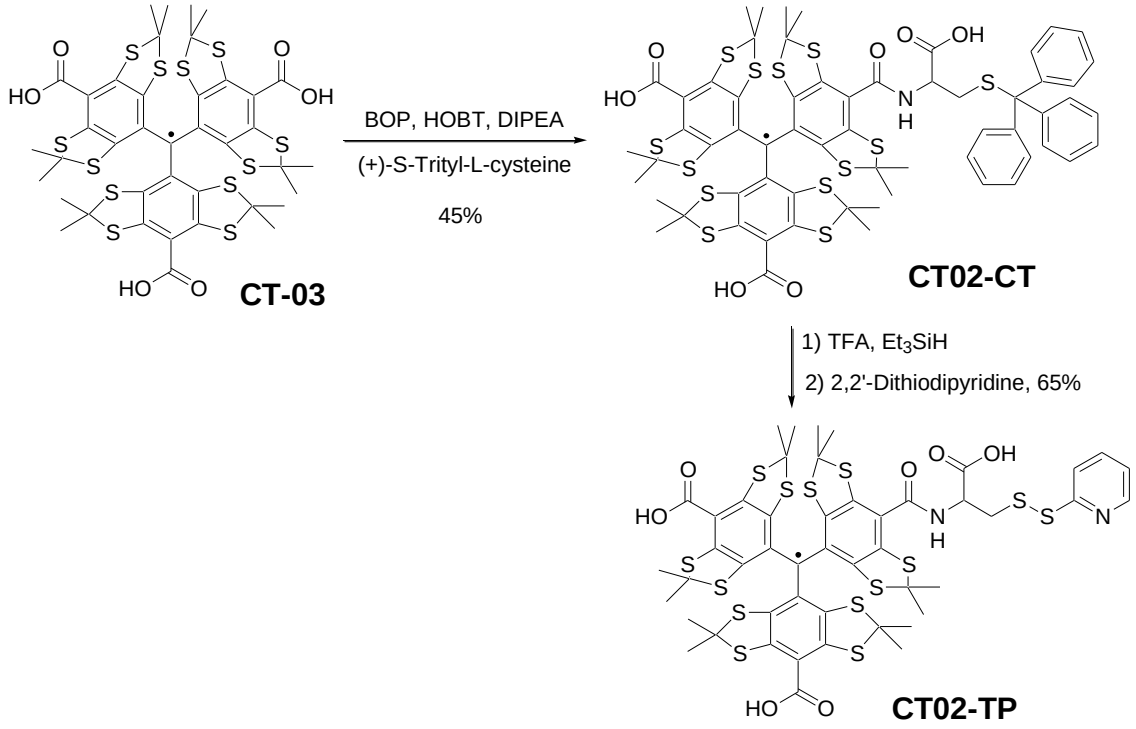
[3]National Biomedical Center for Advanced ESR Technology, Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, United States.

Corresponding to Wayne L. Hubbell, Jay L. Zweier and Jack H. Freed

Email: hubbellw@jsei.ucla.edu; jay.zweier@osumc.edu; jhf3@cornell.edu.

1. Details of TAM synthesis

To the solution of CT-03 (50.0 mg, 50 μmol), 1-hydroxybenzotriazole (HOBt, 20.3 mg, 150 μmol), and (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluoro-phosphate (BOP, 23.2 mg, 52.5 μmol) in dry DMF (10 mL) was added N,N-diisopropylethylamine (DIPEA, 100 μL) under N_2. The reaction mixture was stirred at room temperature for 20 min, and then (+)-S-trityl-L-cysteine (19.1 mg, 52.5 μmol) in 4 mL of DMF was added dropwise. The resulting mixture was continuously stirred for 18 h at room temperature. Solvent was removed under vacuum, and the residue was dissolved in phosphate buffer (0.1 M, pH 7.4) and purified by column chromatography on reversed-phase C-18 using water followed by 0-15% acetonitrile in water as eluants. The second fraction was collected and concentrated to give the compound CT02-CT as a green solid (31.8 mg, 45%). Thereafter, CT02-CT was dissolved in DCM (2 mL), TFA (2 mL) and 20 μL of triethylsilane. The reaction mixture was stirred for 3 h at room temperature and evaporated to dryness under vacuum. The residue was redissolved in DMF (5 mL) under N_2 and 2,2'-dithiodipyridine (7.7 mg, 35 μmol) was then added. The reaction mixture was stirred overnight and concentrated under vacuum. The residue was purified by column chromatography on reversed-phase C-18 using water followed by 0-15% acetonitrile in water as eluants to give the trityl spin label CT02-TP (18.7 mg, 65%). HRMS [MALDI-TOF, dihydroxybenzoic acid as the matrix] m/z calcd. for $C_{48}H_{47}N_2O_7S_{14}^{\bullet}$ ($[M]^+$) 1210.947, found 1211.070.



Scheme S1. Details of the synthesis of CT02-TP.

2. Labeling protein with CT02-TP

Double mutants of 65C/76C and 65C/80C were generated by QuikChange site-directed mutagenesis of the pET11a-T4L genetic construct.[1,2] Mutations were verified by DNA sequencing. Both mutants contain the pseudo-wild-type mutations C54T and C97A.[3] Cysteine mutants of T4L were expressed, purified, and then desalted (to remove DTT) into a buffer suitable for spin labeling (the "spin labeling buffer", which contains 50 mM MOPS and 25 mM NaCl) using previously reported procedure.[1] To protect the sulfhydryl groups, a 10 fold molar excess of *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate (MTSL, a generous gift from Prof. Kalman Hideg) was added and left to react at 4 ℃ overnight. Excess MTSL was removed using an Amicon spin concentrator (Millipore, 10,000 MWCO, 50 ml). The reaction with MTSL generates the nitroxide side chain R1as a protecting group that can be monitored by ESR.[4]

3. Attaching T4L to Sepharose

CNBr-activated sepharose beads were obtained from Sigma-Aldrich. Typical bead volume for each sample was ~ 50 μl suspended in ~ 1 ml of spin labeling buffer. Approximately 0.5 mg of protected protein from step 2 above (25 nmol) was added to the beads/buffer mixture ***(total protein concentration of ~500 μM).*** After incubating 2-3 hrs at room temperature, samples were centrifuged at ***13000 ×g*** for 1 min. The supernatant was removed and concentrated using the spin concentrator (Millipore, 10,000 MWCO, 500 μl). The unbound protein in the supernatant was determined by the absorbance at 280nm (ε_{280} = 24,750 $cm^{-1}M^{-1}$); essentially all protein was coupled, leading to a final protein concentration on beads of 500 μM.

The beads were resuspended using 1 ml of the spin labeling buffer. Dithiothreitol was added to a final concentration of 10 mM and allowed to incubate for 3 hrs at room temperature to remove the R1 protecting group. DTT and products were removed by washing with spin labeling buffer 6 times (for each wash, the beads were resuspended in 1 ml buffer, centrifuged at ***13000 × g*** for 1 min and the supernatant removed). CW ESR spectra were used to confirm that there was no R1 spin label in the bead sample. The CT02-TP reagent (cf. Scheme 1 and main text) was then added to each sample with a molar ratio of about 3:1 (CT02-TP to free cysteine) and allowed to react for 12-16 hrs at 4 ℃. Unreacted CT02-TP was removed by washing 3 times with the spin labeling buffer.

4. Labeling efficiency.

The labeling efficiency was determined by using the 4-Pyridine Disulfide (4-PyDS) Assay as described previously.[5] 4-PyDS reacts with cysteines to yield 4-Thiopyridone, which has a UV absorption at 324 nm with a relatively high extinction coefficient (1.98×10^4 M cm^{-1}).[6] Using this absorption, the content of free thiol can be calculated. The 4-PyDS reagent was obtained from Sigma-Aldrich (sold as Aldritiol™-4). For each experiment, fresh 4-PyDS stock solution was prepared with a concentration of 100 mM. To determine the reaction yield of the protein on beads with CTO2-TP, 4-PyDS was added to a suspension of the beads+protein at a final concentration of 0.5 mM. After approximately 5 minutes incubation at room temperature, the mixture was centrifuged for 1 minute at ***13000 × g*** and the UV absorbance of the supernatant at 324 nm was used to determine the free thiol concentration. Comparison of this value to that of a reference sample treated in the same way but prior to

reaction with CTO2-TP showed that approximately 65% of the cysteines in the protein in each case had reacted with CTO2-TP.

5. Continuous Wave-ESR of the TAM-labeled protein on CNBr-activated sepharose

The CW ESR spectra were collected on a Varian E-109 X-band spectrometer equipped with a loop-gap resonator at room temperature.[7] ***Typical sample size for a CW experiment is 5 µL.*** Optimized ESR parameters are: 0.1 mW detection power, 0.1 G modulation amplitude, 20 G scan range and 60 seconds single scan time. After being attached to the CNBr-activated sepharose, the line width was approximately 2.2 Gauss (6.2 MHz). ***Weak line broadening due to spin-spin dipolar interaction was observed as shown in Figures S1, C and D. The weak broadening in both 65/80 TAM and 65/76 TAM compared to the expected broadening of ≈2-3 Gauss is the result of a substantial amount of singly-labeled protein, for which the spins are not interacting. Good fits to the dipolar broadened spectra according to Altenbach et al[8] were obtained based on inter-spin distances derived from the DQC data and ≈ 39% and ≈27%, respectively, of interacting spin pairs were determined; the remainder is due to singly labeled protein. These values are not unreasonable when compared to the 65% estimate for the total labeled cysteines, which would statistically give about 42% of spin pairs. This effect highlights an important feature of DQC, namely that isolating the dipolar coupling from the other sources of signal with a double-quantum filter enables accurate distance measurements to be made even in the presence of significant amounts of singly labeled protein, which is a distinct advantage over line broadening measurements. This advantage is more substantial for dilute samples, where intermolecular effects are minimal and CW spectra noisier.***

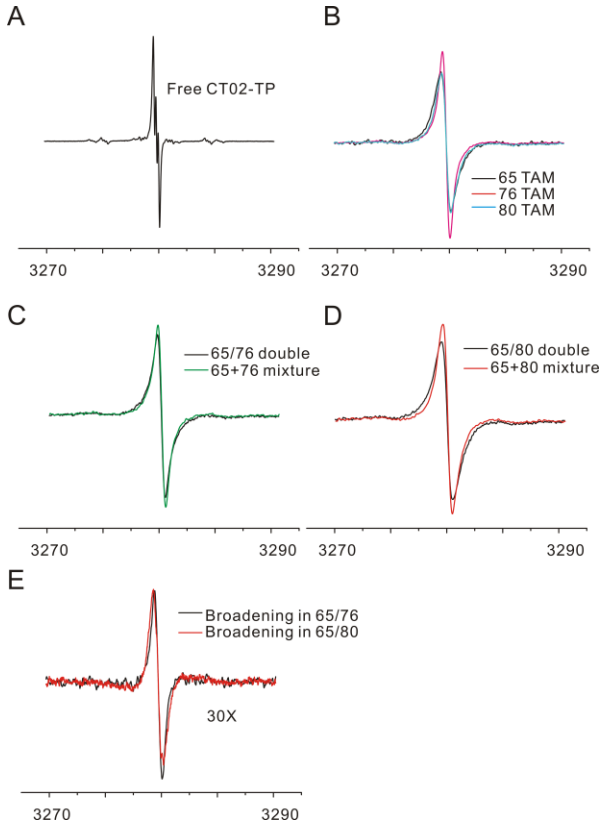


Figure S1. (A), CW-ESR spectrum of the free CT02-TP in solution at 20 ℃. (B), CW-ESR spectra of three singly TAM labeled T4L mutants, 65C (dark), 76C (purple), and 80C (blue). (C), CW-ESR spectra of doubly labeled 65/76C mutant (dark) and a mixture of singly labeled 65C and 76C mutants (green). (D), CW-ESR spectra of doubly labeled 65/80C mutant (dark) and a mixture of singly labeled 65C and 80C mutants (red). (E), A comparison of the broadening effects (residual spectrum after subtracting the CW spectrum of mixture of singly labeled sample from that of doubly labeled sample) in the 65/76 sample and the 65/80 sample. Spectra in (B), (C) and (D) are collected from T4L mutants covalently attached to CNBr-activated sepharose. CW-ESR parameters used in these measurements are listed in the text.

The broadening effect observed in 65/80 TAM appears to be similar to that from the 65/76 TAM (Figure S1, E), whereas one would expect that latter to be broader due to a shorter interacting distance. This is likely due to the lower percentage of spin pairs in the 65/76 mutant compared to the 65/80C mutant as determined above. The lower population of interacting pairs in the 65/76 mutant is also supported by relatively larger intermolecular background in the DQC data for 65/76 TAM compared to 65/80 TAM (Figure 2 and S3) (the background is linear in spin labeling efficiency, while concentration of pairs is quadratic (Ref 9: Borbat and Freed, 2007)).

6. DQC pulse sequence

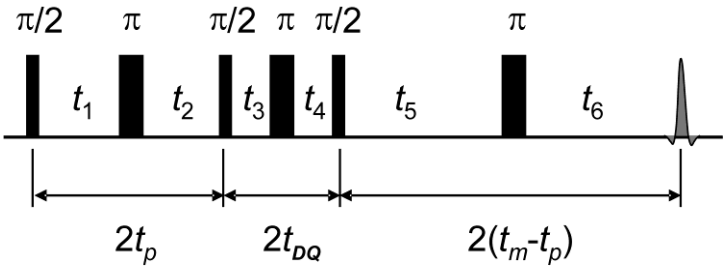


Figure S2. DQC 6-pulse sequence uses three π/2 and three π pulses separated by $t_1 = t_2 = t_p$; $t_3 = t_4 = t_{DQ}$; and $t_5 = t_m - t_p$. The 6-pulse echo is formed at time $t_6 = t_5$ after the last pulse. The echo is recorded as a function of $t_\xi \equiv t_m - 2t_p$ by varying t_p and t_5 in steps of Δt, such that, the sum $t_p + t_5 = t_m$ is constant as is t_{DQ}. Varying intervals in this manner does not shift the position of the 6-pulse echo, but t_ξ ranges from $-t_m$ to $+t_m$. The DQC modulation of the echo amplitude is isolated from the basic echo signal by applying phase cycling. The resulting signal is symmetric with respect to $t_\xi = 0$, therefore t_ξ is usually varied in the range $(0, t_m)$.

The 6-pulse DQC sequence used in all measurements is shown in Fig. S2. ***All DQC measurements were performed at 17.2 GHz. Typical sample size for a DQC experiment is 10 μL with a sample volume contributing to the signal that was estimated as 8±2 μL.*** The pulse lengths used were 10 and 20 ns, for π/2 and π pulses, respectively. A step size Δt of 4 ns was used in all measurements, giving 8 ns steps in t_ξ., and t_{DQ} was fixed at 44 ns. A 64-step phase cycle[10], which is a subset of the 256-step phase cycle ***was applied to isolate the DQ signal. Since the receiver phase was different from that shown in the basic 64-line table on page 72 in reference #10, the simplest way to describe the phase cycle is as follows: take the first 32 lines of the phase cycle as is, then discard the last 32 lines in the printed version and replace them with a new set of 32 lines, which are obtained from the first 32 lines by the sixth pulse phase changed to y and the sign of the receiver phase is inverted. No CYCLOPS steps, prescribed for constructing 128- and 256-step phase tables, were used. The phase cycle constructed in this way is a quarter of the 256-step phase cycle. The raw data are shown in Fig. S3.***

7. Raw DQC data

The number of data points, collected for the two doubly labeled samples (50 and 75 points) and for a control sample of a singly TAM labeled 65C and 76C mixture (75 points), shown in Fig. S3, depend on t_m, which was 400 and 600 ns, respectively. Five extra points for negative t_ξ were recorded to better display the maximum at zero t_ξ. The pulse repetition rate was limited to 4 kHz to avoid heating of the sample. Signal averaging time for each sample was 3 to 6 hours. ***All signals were normalized to unity at zero time and plotted as for DEER, but neither the shape of the individual signals nor the difference between the control and doubly-labeled data should be interpreted in terms of modulation depth, because the background for DQC has a different origin and is not even present, or is very small, in dilute samples.***

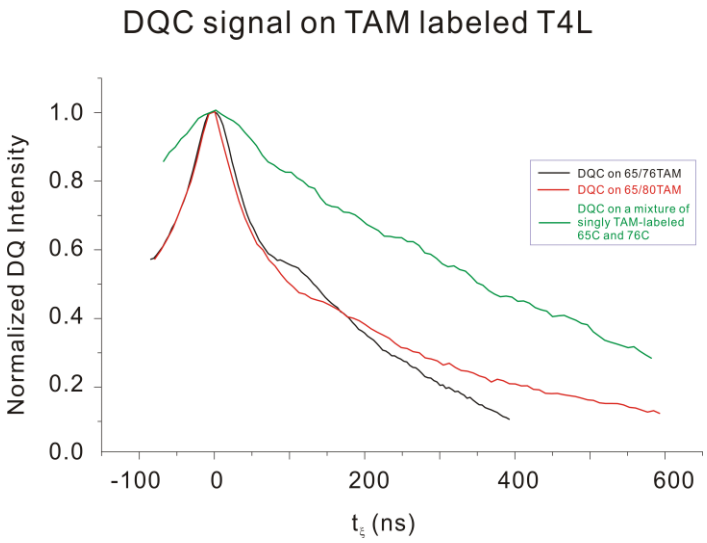


Figure S3. ***Raw DQC data for the TAM labeled 65/76C (dark curve) and 65/80C (red curve) mutants of T4L. The (scaled up) DQC signal on a mixture of singly TAM labeled 65C and 76C is also shown (green curve). The signals go to zero as t_ξ reach t_m, but for obvious technical reasons, one requires that the minimum value of t_p is always greater than zero (so $t_\xi < t_m$). The minimal t_p was set to 50 ns (65/76) and 200 ns for the other cases.***

8. Fitting the DQC signal

A Gaussian function was assumed to represent the distance distribution between TAM spin labels. The average distance and the standard deviation of the Gaussian function were varied until a good fit was obtained (judged by finding the lowest χ^2 value). The equation used for the fitting, based on the literature,[9] is

$$V(t_\xi)/V(0) = \int_{r_{min}}^{r_{max}} \int_0^{\pi/2} P(r)(\cos at_m - \cos at_\xi)\sin\theta d\theta dr \cong \int_{r_{min}}^{r_{max}} \int_0^1 P(r)\cos at_\xi du dr\,, \qquad \text{(S1)}$$

where $P(r)$ is the Gaussian function with r the spin-spin distance, θ is the angle between the spin-spin vector and the external magnetic field, $u \equiv \cos\theta$, and $a = \gamma_e^2\hbar(3\cos^2\theta - 1)/r^3$. Note that in Eq. S1, written in the weak-coupling limit, the first term in parentheses (in the first equality) introduces a small constant effect, which is unimportant in these experiments. ***That is, at low concentrations, (corresponding to an isolated spin pair) only a very small background signal relative to the main signal is introduced by this term, seeing that*** $\langle \cos at_m \rangle_a \to 0$ ***as*** $t_m \to \infty$. ***This holds when*** $t_m >> \langle \delta a^2 \rangle^{-1/2}$, ***which is the case when t_m is***

sufficient for the distance measured. In the second equality this term was, thus, omitted, leaving the pure (oscillating about zero) dipolar signal. In comparison, the signal in DEER can be written in the form

$$V(t)/V(0) = (1-p) + p\int_{r_{min}}^{r_{max}}\int_{0}^{1} P(r)\cos at\,du\,dr\,, \qquad \textbf{(S2)}$$

with a relatively large constant background (1 – p), otherwise the oscillating terms in Eqs. S1-S2 are similar. The amplitude of the second term is referred to as the "modulation depth", but DQC does not have the large constant term, so the scaling is arbitrary so that modulation depth is undefined. At higher concentrations intermolecular dipole-dipole interactions may contribute a more significant background, which becomes especially pronounced in the case of incomplete labeling (cf. Fig. S3). In such cases it has to be removed, for example, by fitting it to a polynomial, as was done in this work. The baseline of each data set was approximated by a second-degree polynomial. Only baseline corrected data were used in the Gaussian function-based analysis as shown in the main text.

The Tikhonov regularization was also used to fit the experimental data, and the results are shown in Figure S4. The most probable distances and widths of the distance distributions are close to those obtained from the Gaussian model in both cases. For the 65/80C sample features at 1.5 nm are also observed, so we cannot entirely rule out a population at this distance range. However, it is not unusual in pulsed dipolar ESR spectroscopy to obtain spurious content for distances below the main peak if signal distortions are present.

The average distances between TAM spin labels are shorter by about 5 Å compared to nitroxide-nitroxide distances measured using DQC at low temperatures (reference #11 of the S.I.). This can be rationalized by the fact that the two nitroxide spin labels are actually pointing away from each other (Reference #11 of the S.I.). The widths of the distance distributions between TAM spin labels are narrower compared to nitroxide-nitroxide distance distributions.

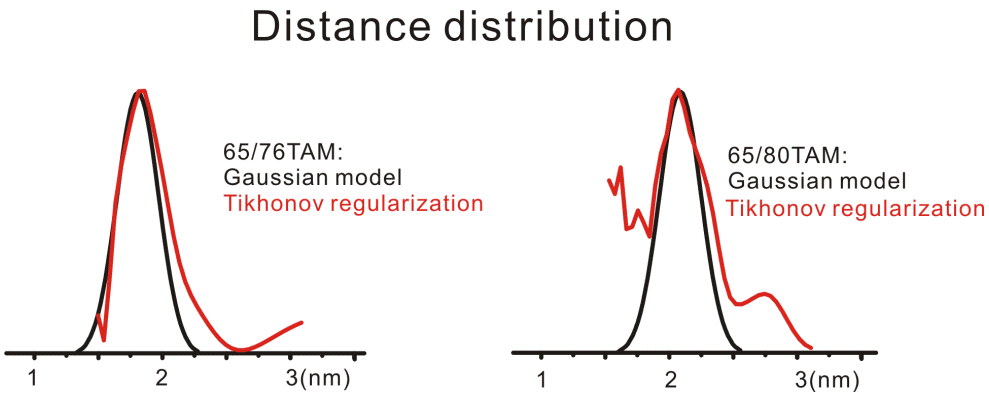


Figure S4. Distance distribution extracted from DQC data for the TAM labeled 65/76C (left) and 65/80C (right) mutants of T4L using the Gaussian model and Tikhonov regularization.

9. <u>Conformational freedom of the TAM label.</u>

The TAM label has, in principle, 8 points of internal flexibility, corresponding to the 8 dihedral angles of the bonds, numbering from the C_α-C_β bond as X_1 to the C=O – aromatic bond as X_8. The TAM label employs the same disulfide linkage used widely in the R1

nitroxide label, for which we have more than 20 crystal structures. The lessons from these structures are thus applicable to the TAM label up to X_3. These are summarized in Fleissner et al [12] and Warshaviak et al [13] and are: (1) The X_1 dihedral has the same value as the native side chain at the selected site and is either in the {t} configuration (nominally 180^o) or the {m} configuration (nominally -60^o); (2) X_2 is fixed by a ubiquitous S_δ---HC_α interaction found in all structures, and predicted by QM calculations[13]; for X_1={t}, X_2 is either {p} ($+60^o$) or {m}. The S_δ---HC_α interaction "locks" the atom group C_α-C_β-S-S in a fixed configuration; (3) X_3 is nominally $\pm 90^o$, the choice determined by local steric interactions. In the TAM radical, X_7 and X_8 are determined the requirement of a planar peptide group, and planarity of that group with the aromatic ring as a result of pi-conjugation. That leaves X_4 –X_6 to be determined by internal interactions within the side chain. Rotations about X_4 and X_6 cause only minor changes in the position of the TAM radical itself since the axes of these bonds pass nearly through the central carbon. The dihedral X_5 can adopt the canonical values of $\approx \pm 60, 180$, but only 180 is permitted due to extreme steric crowding. When the TAM radical is modeled at 65 and 76 and at 65 and 80 according to these considerations, the distances between the central carbons of the TAM radicals are very close to the distances determined from the DQC data. In summary, the linkage between the protein and the TAM radical is fairly constrained by the sulfur-backbone interactions, by the planarity of the peptide-aromatic group, and by steric constraints due to the large size of the TAM radical itself. The good agreement of the experimental and modeled distance, together with the relatively narrow distance distribution, suggests that at least the major population of TAM radicals is relatively well localized.

10. ***TAM Dynamics***

Even when T4L is immobilized, one expects the TAM labels to engage in some restricted range of reorientational motion leading to only partial averaging of the small magnetic anisotropies defined by its hf and g tensors. A sensible model for each TAM side-chain would be the MOMD (microscopic order macroscopic disorder) model frequently used in interpreting such a class of systems as implied in the previous section. Using singly TAM labeled T4L at 17.2 GHz we measured the full width at half height in the frozen sample of ~11.5 MHz, and at 0 °C it is ~7.5 MHz, so the side-chain motion only partially averages the magnetic anisotropy. This spectral line is still inhomogeneously broadened as expected for a MOMD spectrum, given that its homogeneous T_2 is ~700 ns, corresponding to a homogeneous Lorentzian full width of ~0.5 MHz, which is at least an order of magnitude below the CW width; results on bi-labeled TAM are confusing because they are a mixture of singly and doubly labeled T4L. We now present the results of a simple but instructive model for the effect of fluctuations in r.

Using simple estimates for the TAM side-chains distributions assuming the similarity to MTSSL [14] (e.g. $\Delta r \cong 2$Å per side-chain or ~6.5 Å distribution width in the static case) as well as a scaled up value from that of T4L [14]due to the increased size (by about a factor of 7) of TAM over MTSSL, we estimate τ_c for this process of ca. 8 ns, then for r_{12}=20Å ($\gamma_e^2 h/2\pi r^3$=6.5 MHz) we find orientation dependent broadenings from such dipolar fluctuations ranging from 0.2 MHz to 0.8 MHz for 90° and 0°, respectively. Thus the dipolar spectrum from DQC produced by all orientations is still close to Pake doublet, where the rotameric conformations are averaged out. That is for any given T4L orientation

the dipolar line is homogeneous. Thus, a rigorous treatment of the DQC experiment may need to include orientation-dependent relaxation for the distances in this range. The distance distributions could thus represent conformational heterogeneity of the protein, rather than a collection of spin-label rotamers. Slow backbone dynamics (<10 kHz) appears static on the time-scale of the DQC pulse sequence and may contribute to the width of distance distributions. For somewhat shorter distances electron spin relaxation becomes more substantial, and the "foot" of the Pake doublet may disappear due to a factor of 4 faster relaxation for this orientation. For a distance of 15 Å the line width is a factor of 5.6 greater compared to 20Å, i.e. ranging from ca. 1.1 MHz (90°) to 4.5 MHz (0°). Consequently, short distances (<15 Å) could become unobservable by DQC at ambient temperatures due to fast relaxation rates introduced by such broadening. This could be circumvented by designing more rigid side-chains.

Reference

(1) López, C. J.; Fleissner, M. R.; Guo, Z.; Kusnetzow, A. K.; Hubbell, W. L. *Protein Science* **2009**, *18*, 1637.

(2) Fleissner, M. R.; Brustad, E. M.; Kálai, T.; Altenbach, C.; Cascio, D.; Peters, F. B.; Hideg, K.; Peuker, S.; Schultz, P. G.; Hubbell, W. L. *Proceedings of the National Academy of Sciences* **2009**, *106*, 21637.

(3) Matsumura, M.; Matthews, B. W. *Science* **1989**, *243*, 792.

(4) Langen, R.; Cai, K.; Altenbach, C.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1999**, *38*, 7918.

(5) Fleissner, M. R.; Bridges, M. D.; Brooks, E. K.; Cascio, D.; Kálai, T.; Hideg, K.; Hubbell, W. L. *Proceedings of the National Academy of Sciences* **2011**, *108*, 16241.

(6) Grassetti, D. R.; Murray Jr, J. F. *Archives of Biochemistry and Biophysics* **1967**, *119*, 41.

(7) Hubbell, W. L.; Froncisz, W.; Hyde, J. S. *Review of Scientific Instruments* **1987**, *58*, 1879.

(8) **Altenbach, C., Hubbell, W.L.** ***Abstract from biophysical Society Poster, Long Beach, 2008, Abstract number 2467-Pos***

(9) Borbat, P. P.; Freed, J. H., In Methods in Enzymology, Editor: Melvin I. Simon, B. R. C. a. A. C., Academic Press: 2007; 423, 52.

(10) Borbat, P. P.; Freed, J. H. "Double-Quantum ESR and Distance Measurements" In Chapter 9 of Biological Magnetic Resonance; Kluwer Academics/Plenum Publishers: New York, 2000; Vol. 19.

(11) Borbat, P. P.; Mchaourab, H. S.; Freed, J. H. J. Am. Chem. Soc. 2002, 124, 5304.

(12) Fleissner, M.R.; Cascio, D.; Hubbell, W.L. Protein Sci. 2009, 18, 893.

(13) Warshaviak, D.T.; Serbulea, L.; Houk, K.N.; Hubbell, W.L. J.Phys. Chem. B 2011, 115, 397.

(14) Zhang, Z. W.; Fleissner, M. R.; Tipikin, D. S.; Liang, Z. C.; Moscicki, J. K.; Earle, K. A.; Hubbell, W. L.; Freed, J. H. J. Phys. Chem. B 2010, 114, 5503.