Supporting Information

Supplemental Methods The DEER Signal and Details on the modeling of dipolar signals

Table S1 Intermolecular distances between CheA:CheW and Tm14

Table S2 Summary of disulfide cross-linking between Cys-engineered Tm14 and CheA:CheW

Figure S1 Modeling of distance distributions from intermolecular spin-spin dipolar coupling

Figure S2 Simulations of CheW to Tm14 dipolar spin interactions

Figure S3 Simulations of P3 residue 318 to Tm14 (100 and 111) dipolar spin interactions

Figure S4 Simulations of P3 residue 331 to Tm14 (100 and 111) dipolar spin interactions

Figure S5 Simulation of P3 residue 301 to Tm14 residue 149 spin interactions

Figure S6 Simulation of P5 resdiue 545 to Tm14 residue 111 spin interactions

Supplemental Methods

The DEER Signal

In an isotropic medium the DEER signal from a pair of nitroxide electron spins separated by a distance r is given as

$$V(r,t) \propto 1 - pu(r,t) \qquad (2)$$

where $u(r,t) = 1 - \int_0^1 \cos[\omega_d(1-3u^2)t]du$, with t the evolution time, and p the fraction of spins flipped by the pump pulse (*1-4*). For large $t >> 2\pi/\omega_d$, $u \rightarrow 0$ and the signal reaches the constant background with magnitude of (1–p). The time varying part of V is referred to here as the dipolar signal. The amplitude of the dipolar signal is thus p in the ideal case, which we will refer to as the nominal dipolar signal amplitude in the sequel.

For a range of distances, the signal is obtained by integrating $V(r,t)$ in Eq. 2 with the distribution of spins over distance, $P(r)$. In writing Eq. 2 we have omitted any effects due to orientational selectivity (*5*) since such effects are often not significant for the flexible MTSSL side-chains used and are difficult to model in the absence of specific structural information.

The signal in Eq.2 is further modified by the intermolecular contributions produced by dipolar couplings with the electron spins residing on nearby molecules, i.e. $V = V_{intra}V_{inter}$, where V_{intra} is the signal given by Eq.2 and V_{inter} in an isotropic medium is given by the exponential expression $V_{inter} = \exp(-kpCt)$, (*1, 4*) with C being the average local spin concentration (i.e. within the distance range of several hundred Angstroms in the sample, p is defined above, and the constant $k \approx 1.03\times10^{-3}$ $\mu M^{-1}\mu s^{-1}$ (*2*). V_{inter} can be more complex due to any non-uniformity of the local distribution of surrounding electron spins but also due to "excluded volume effects", which originate from the finite size of the molecules.

The three-spin signal in DEER in the simplest form is given by the sum:

$$V_3 = V_{1,23} + V_{2,13} + V_{3,12} = A_{1,23}(1-pu_{12})(1-pu_{13}) + A_{2,13}(1-pu_{12})(1-pu_{23}) + A_{3,12}(1-pu_{13})(1-pu_{23}) \quad (3)$$

where $V_{i,kl}$ is the dipolar signal detected on spin i, when pumping on spins k and l. The amplitudes are given by $A_{i,jk} = C_i R_i x_i x_j x_k (1-p_a u_{ik}(t_m))(1-p_a u_{il}(t_m))$ with C_i representing relative fractions, x_m is spin labeling efficiency, R_i accounts for phase relaxation, t_m is the maximum evolution time, and p_a is the probability to flip spin A. Leaving out for clarity potential complications that could arise due to orientation selection (*6*), the first of the three similar signal components is:

$$V_{1,23} \propto (1-pu_{12})(1-pu_{13}) = (1-p(1-\cos\omega_{12}t)(1-p(1-\cos\omega_{13}t)) = 1+p^2-p(1-p)(u_{12}+u_{13})-p^2(u_-+u_+)/2 \quad (4)$$

with $u_\pm \equiv \cos(\omega_{12} \pm \omega_{13})t$. Thus, these nonlinear effects lead to harmonics with separation of $p/(2\text{-}2p)$ which is ~0.2 for a typical p value of 0.3 used in this work. Consequently, the distance distribution peak reconstructed by using standard algorithms will produce spurious peaks or sidebands for broad peaks, which in reality do not represent real distances. In the case of CheA, incomplete spin labeling on CheA would yield only a slight improvement, since the SNR will drop as x^2, therefore x cannot be made very small. Also p cannot be set very small, because in the present case of weak affinity interactions (K_d ~ 20-50 μM) an excess of spin-labeled receptor was used, leading to reduced ratios of the dipolar signal to the background. This amplifies undesirable effects of the background, such as those caused by concentration heterogeneity in the sample (clustering etc.), excluded volume effects, baseline artifacts of instrumental origin, relaxation, and electron-nuclear coherence effects. An unbound receptor module that has a shorter relaxation time is of benefit here, since its more rapid decay aids observation of the dipolar signal from CheA/CheW with a longer relaxation time interacting with the bound receptor (i.e. the latter is the pump spin). Using a complex of two heterodimers would provide a solution, with the additional benefits of yielding only inter-domain signals; however, an efficient procedure for engineering CheA heterodimers has not as yet been developed, and this approach would not be applicable to exchangeable CheW. A possible approach is to use two values of p (e.g. 0.2 and 0.4) and combine the respective time domain-signals in appropriate weights; however this method has not been attempted because the major source of error has

originated from the uncertain baseline, which for data analysis was taken to be linear on a log scale (*4*).

Details on Modeling of Dipolar Signals

A more detailed modeling of the domain flexibility ultimately should result in even better agreement with the experimental data. It should also be noted that the entire signal from the experiment where only CheA:CheW is labeled can be utilized to extract (deconvolute) the dipolar signal originating from coupling from these spins to the spin on the receptor when all components are labeled (to the accuracy limited by non-linear contributions; cf. Eq. 4 in Supplement). For the model parameters we also used appropriate concentrations of proteins, K_d, spin labeling efficiencies and relaxation times. Whereas spin labeling efficiencies of CheA and CheW can be determined accurately from the modulation depth (i.e. the magnitude of apparent p) of their respective DEER signal, they are less certain for the receptor. Phase relaxation times, i.e. T_2's (or more precisely the echo decay time) can be accurately determined for all components, including free and bound states. Therefore, the modeling required just the few adjustable parameters needed to fit mainly the background and the signal modulation depth. The residual between the simulated signal and the experiment gives an estimate to the uncertainty of the proposed structure. The ratio of relaxation parameters and relative occupancies of labeled sites weights the contributions to the dipolar signal. First, one needs to accurately reproduce the dipolar signal of the control experiment 2 (or else to use a smoothing spline approximation to the experimental signal), i.e. between the spins on CheA or CheW domains; then the variables such as labeling efficiencies, concentrations, relaxation, and estimates accounting for clustering, can be easily fit to minimize the residual.

The simulations coded using MATLAB (MathWorks, Inc.) were based on using (Eqs. 2 and 3) which included all the details required to account properly for the amplitudes. Then the distance distributions for the structure were generated by the Monte-Carlo method, and the two- and three-spin calculated time-domain signals ($V_{2\mathrm{i},k}$, $V_{3i,jk}$) were added and adjusted for the intermolecular background signal as described above.

References

1. Borbat, P. P., and Freed, J.H. (2007) Measuring distances by pulsed dipolar ESR spectroscopy:spin-labeled histidine kinases, *Meth. Enzymol. 423*, 52-116.

2. Jeschke, G., and Polyhach, Y. (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance, *Physical Chemistry Chemical Physics 9*, 1895-1910.

3. Schiemann, O., and Prisner, T. F. (2007) Long-range distance determinations in biomacromolecules by EPR spectroscopy, *Q. Rev. Biophys. 40*, 1-53.

4. Milov, A. D., Maryasov, A. G., and Tsvetkov, Y. D. (1998) Pulsed electron double resonance (PELDOR) and its applications in free-radicals research, *Applied Magnetic Resonance 15*, 107-143.

5. Margraf, D., Bode, B. E., Marko, A., Schiemann, O., and Prisner, T. F. (2007) Conformational flexibility of nitroxide biradicals determined by X-band PELDOR experiments, *Mol. Phys. 105*, 2153-2160.

6. Jeschke, G., Sajid, M., Schulte, M., and Godt, A. (2009) Three-spin correlations in double electron-electron resonance, *Phys. Chem Chem Phys 11*, 6580-6591.

Table S1. Intermolecular distances between Tm14 and CheA/CheW

	Spin label sites on single chain Tm14				
	100	111	149	160	167
CheW					
9	NA	NA	***20-30 Å***	*20-35Å	Inconclusive
31	NA	*50-60 Å	NA	NA	†22-35 Å
80	Inconclusive	NA	* 20-30 Å	NA	***20-30 Å***
137	NA	NA	NA	*35-45 Å	NA
139	NA	NA	#20-30 Å	Inconclusive	NA
P3					
301	* >60 Å	Inconclusive	†30-45 Å	NA	Inconclusive
318	***45-70 Å***	***45-65 Å***	NA	NA	NA
331	***35-70 Å***	***35-70 Å***	Inconclusive	NA	NA
P4					
371	Inconclusive	NA	Inconclusive	Inconclusive	NA
387	NA	NA	Inconclusive	*45-65 Å	Inconclusive
P5					
545	***45-75 Å***	***45-60 Å***	NA	†35-50 Å	Inconclusive
568	Inconclusive	NA	NA	NA	NA
634	***50-80 Å***	***40-70 Å***	NA	NA	NA
639	Inconclusive	*20-30 Å And **40-75 Å**	NA	NA	NA
646	Inconclusive	NA	NA	NA	NA

Distances in Bold indicate results of Tikhonov regularization.

***** Distance estimated qualitatively from observing the change in lineshape of dipolar signal on addition of spin labeled receptor when compared to dipolar signals measured in presence of unlabeled receptor.

† Distance determined by fitting of time-domain data only. See Figs. S2-S6.

Detection of distinct but weak amplitude intermolecular signal. Such signals may arise from oligomerization of single chain receptor itself.

Inconclusive: No evidence of intermolecular distances. It is possible that they fall within the same range as intramolecular distances and hence are difficult to resolve.

Table S2. Disulphide cross-linking between cysteine substituted sites on TM14, CheW and CheAΔ289.

	Cysteine sites on TM14					
	100	111	125	149	160	167
CheW						
9	X	X	X	strong*	-	X
15	X	X	X	X	-	X
31	X	X	X	X	X	X
80	-	-	-	X	-	X
100	-	-	-	-	X*	X*
101	weak	X	X	X	X*	X*
102	X	X	X	X	X*	-
139	-	-	-	X	-	X
P5						
545	X	X	-	-	-	X
553	X	X	X	-	-	-
568	-	-	-	-	-	
634	-	-	-	-	-	X
639	X	X	X	-	-	-
646	X	X	X	-	-	-
P4						
371	X	X	X	X	X	-

387	-	-	X	-	X	-

Table 5.2 continued.

401	-	-	X	-	-	-
458	-	-	X	-	-	-
496	X	X	strong *#	-	X	X
508	-	-	X	-	-	-
522	-	-	X	-	-	-

All are single chain receptors except site 125.

X : Negative cross-linking

* : Cysteine substituted sites CheW or CheAΔ289 tested for cross-linking in presence of wild type CheAΔ289 and CheW respectively

: Site 496 on CheAΔ289 did not cross-link with 125 site on single chain receptor

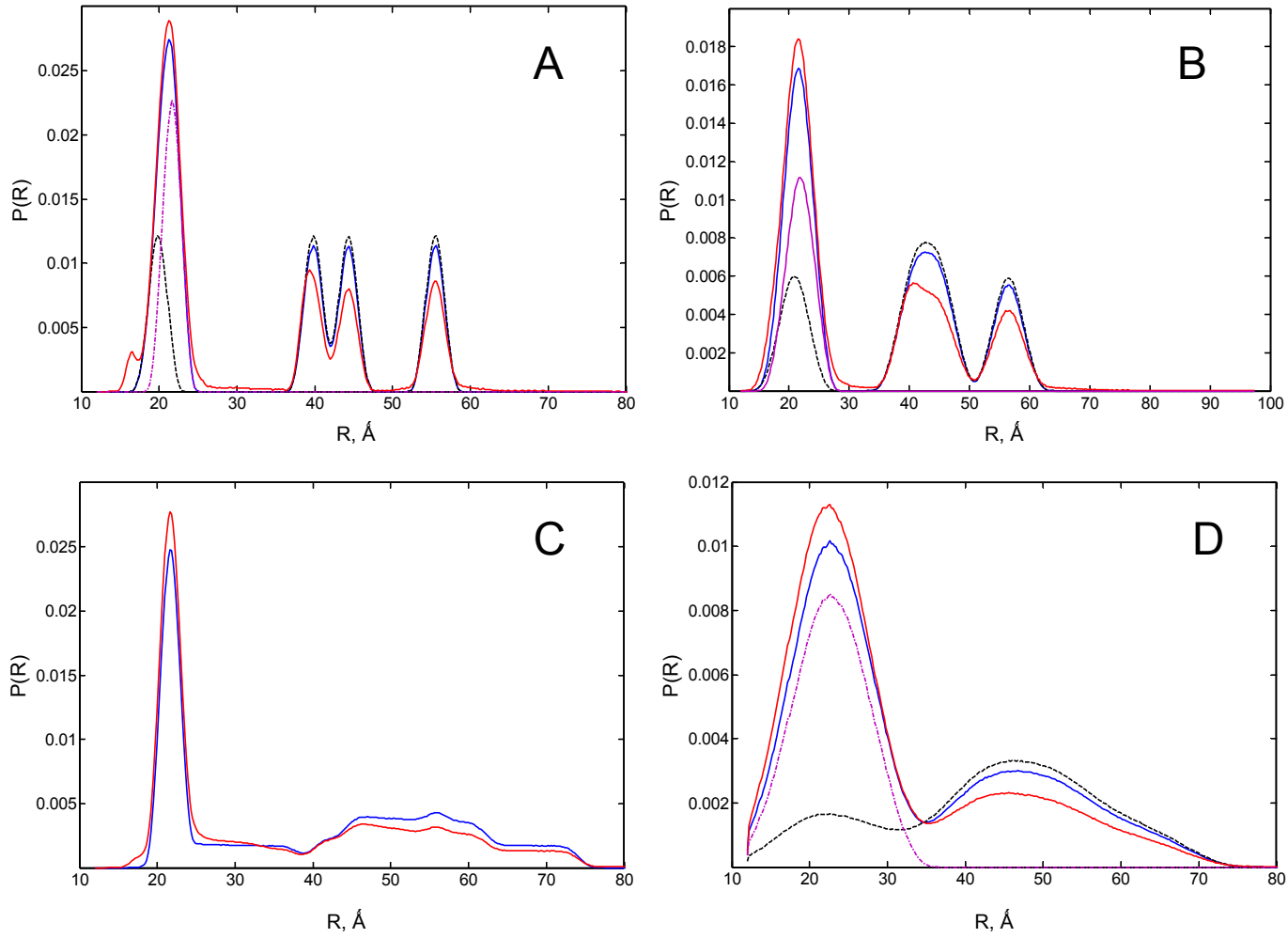


Fig. S1 An illustration of modeling distance distributions using as an example site 331 on P3 and site 100 on the single-chain receptor. The modeling in (A) used sufficiently confined nitroxide moieties to yield narrow distributions, whereas that in (B) used typical MTS spin-label flexibility, mimicked by restricting the nitroxide moiety to a spherical volume of radius ~3.5 A. Spin labeling efficiency was set to 100% for all sites, binding efficiency was 100%, T_2's have been set equal. The larger peak (magenta) corresponds to the distance between two 301 sites. Four smaller peaks (dashed black line) represent four distances between 301 sites and site 100 for the receptor in two possible binding orientations. The complete distance distribution based on the five pair distance distributions are plotted in blue. Accounting for three-spin case corrections produces modified distance distributions, which are plotted in red.

Panel (C) differs from (A) in using realistic spin-labeling and binding efficiencies, T_2's, and appropriate range of receptor positions. Finally, panel (D) shows distance distributions, which provided a good fit to the experimental data. (D) relates to (B) as (C) does to (A).

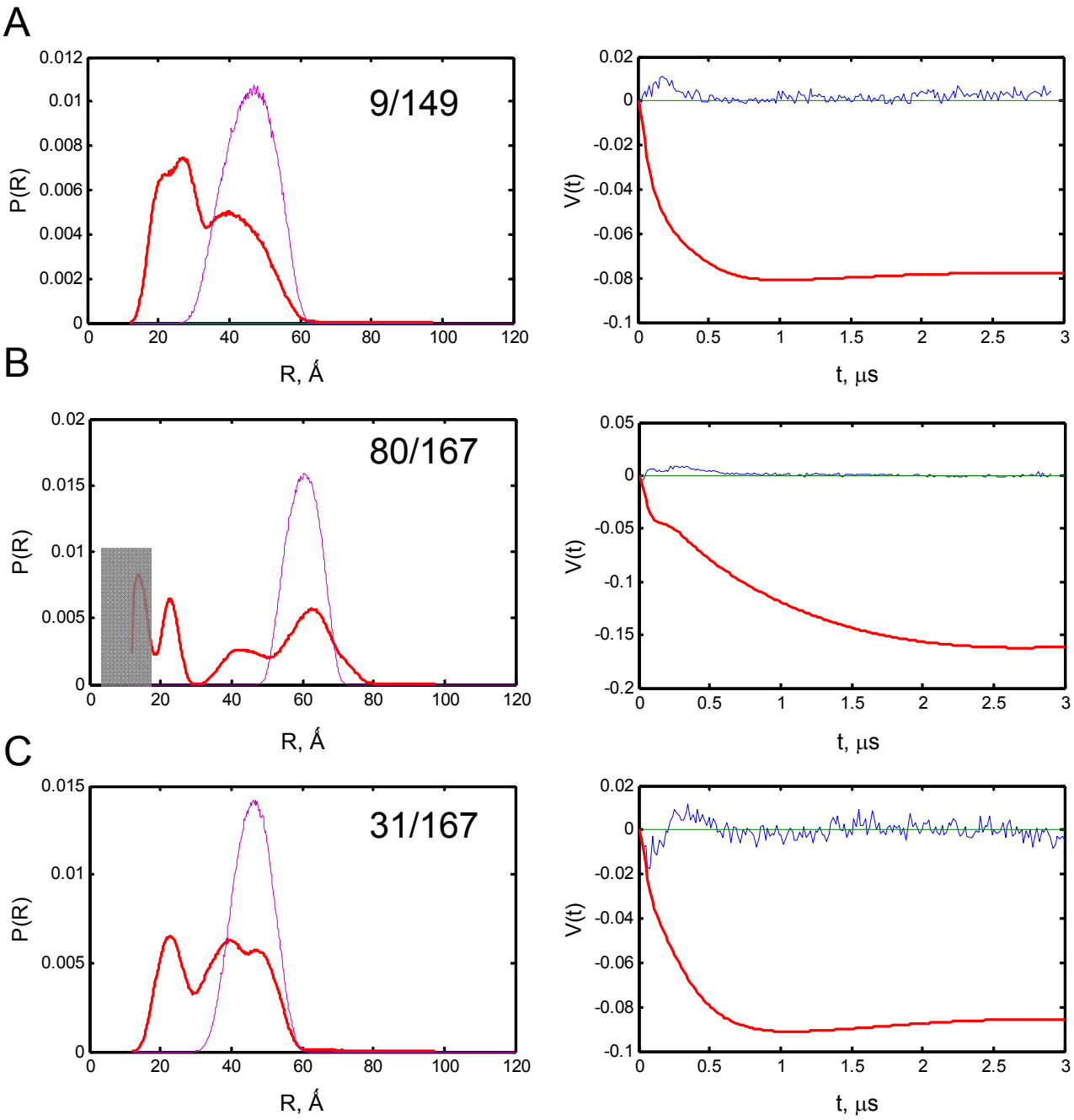


Fig. S2 Left panels show simulated distance distributions for three sets of spin labeled sites on CheW (9, 80, 31) and receptor (149, 167). Distance distributions representing control measurements with unlabeled receptor are plotted in magenta. Models A and B produced very similar results, with model A being slightly better for 80/167. A shaded box in (B) masks that part of the distance distributions, whose contribution to the time-domain signal is nearly suppressed in the DEER experiment. Right panels show time domain signals (in red) corresponding to distance distributions in the left panels. The residuals between the raw experimental and simulated data are plotted in blue.

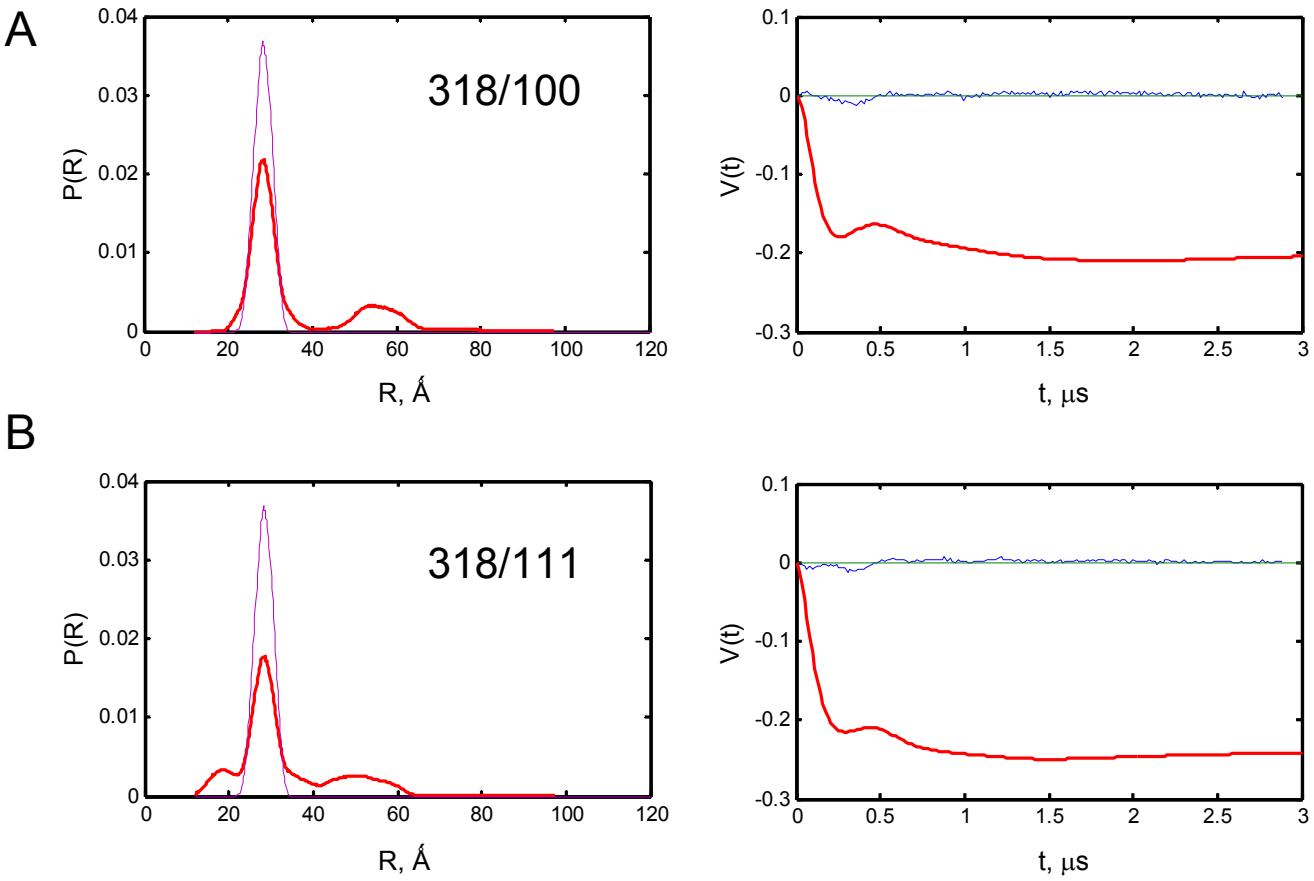
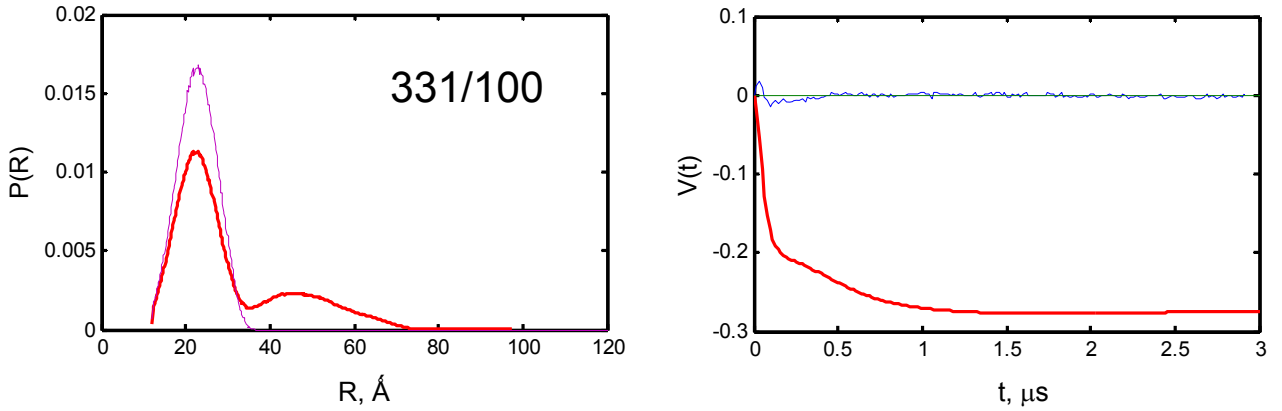


Fig. S3 Left panels shows simulated distance distributions between spin labeled site 318 on P3 and sites 100 and 111 on receptor. The distance distributions for the case of unlabeled receptor are plotted in magenta. Model B produced substantially better representation of the experimental data after an appropriate range of receptor positions was introduced.

Right panels – Time domain signals (in red) corresponding to distance distributions in the left panels. The residuals between the raw experimental and simulated data are plotted in blue. Small imperfections originate from simplified modeling of MTSL side chain at position 318.

C



D

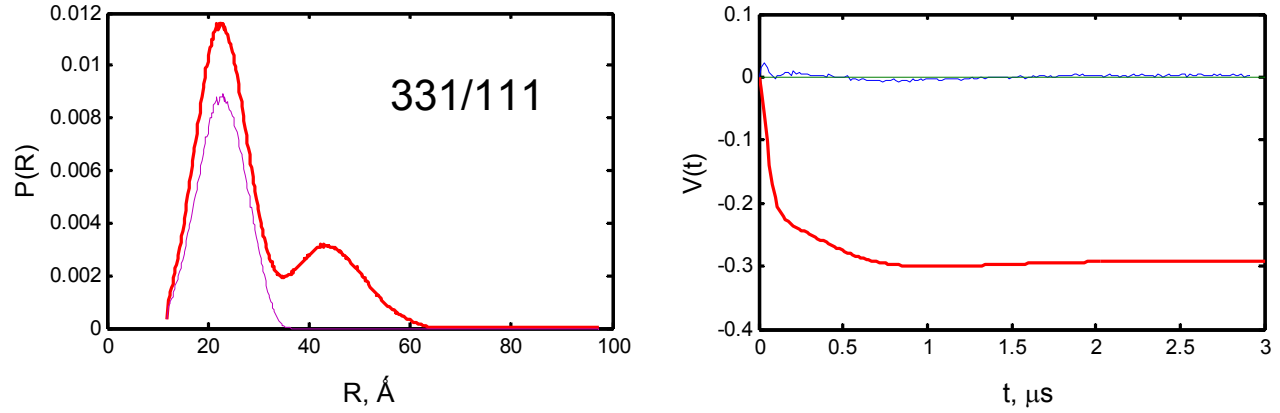


Fig. S4 Left panels show simulated distance distributions between the spin labels at site 331 on P3 and sites 100, 111 on receptor. The distance distributions for the case of unlabeled receptor are plotted in magenta. Model B produced substantially better representation of the experimental data after introducing an appropriate range of receptor positions that keep it away from P3.

Right panels – Time domain signals (in red) corresponding to distance distributions in the left panels. The residuals between the raw experimental and simulated data are plotted in blue.

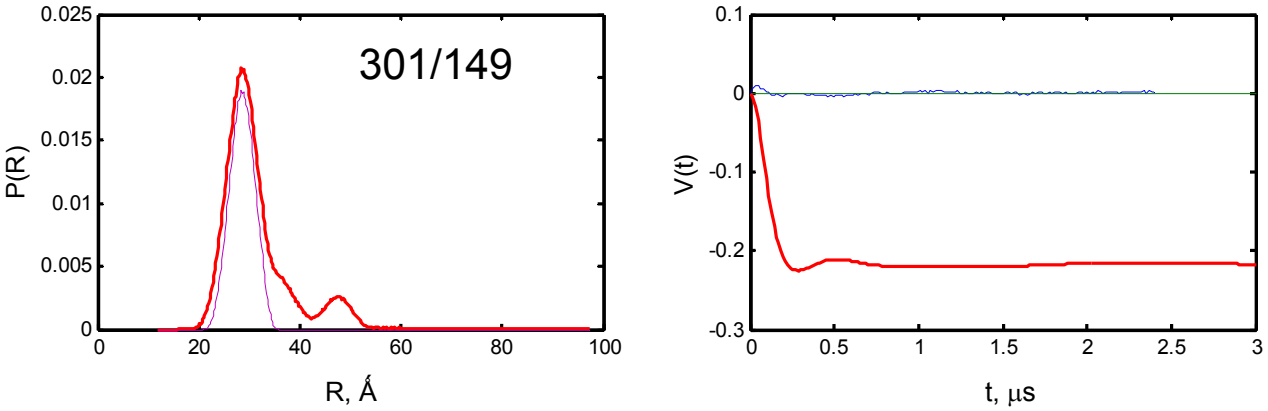


Fig. S5 Left panel shows simulated distance distribution between spin labels at site 301 in P3 and site 149 at the receptor tip. The distance distribution for the case of unlabeled receptor is plotted in magenta. Model B produced better representation of the experimental data, however the range of receptor positions required to obtain good fit can be rather narrow for this case, indicating that receptor and CheW, to which it is bound, occupy a relatively defined position with respect to site 301.

Right panel – time domain signal (in red) corresponding to distance distribution in the left panels. The residual between the raw experimental and simulated data is plotted in blue.

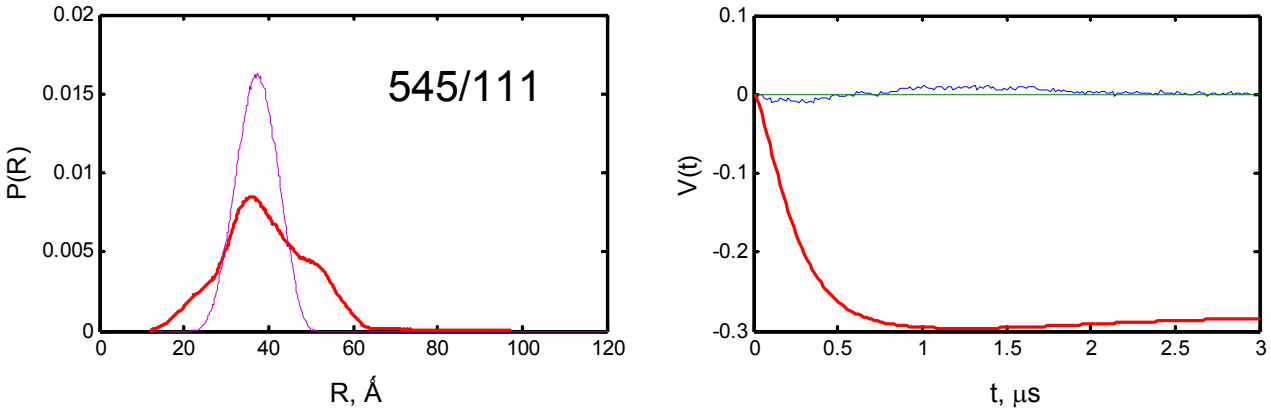


Fig. S6 Left panels shows simulated distance distribution between spin labels at site 545 in P5 and site 111 in the receptor. The distance distributions for the case of unlabeled receptor are plotted in magenta. Model B produced substantially better representation of the experimental data after a range of receptor positions was introduced. Simulations indicated that P5 and the receptor in the same domain are unlikely to be bound.

Right panel – Time domain signals (in red) corresponding to distance distribution in the left panel. The residual between the raw experimental and simulated data is plotted in blue.