THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

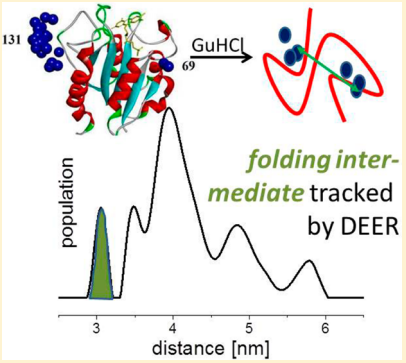
Double Electron−Electron Spin Resonance Tracks Flavodoxin Folding

Martin van Son,[†,#] Simon Lindhoud,[‡,§,#] Matthijs van der Wild,[†] Carlo P. M. van Mierlo,[‡] and Martina Huber*[,†]

[†]Department of Physics, Huygens-Kamerlingh Onnes Laboratory, Leiden University, PO Box 9504, 2300 RA Leiden, The Netherlands

[‡]Laboratory of Biochemistry, Wageningen University, 6700 ET Wageningen, The Netherlands

ABSTRACT: Protein folding is one of the important challenges in biochemistry. Understanding the folding process requires mapping of protein structure as it folds. Here we test the potential of distance determination between paramagnetic spin-labels by a pulsed electron paramagnetic resonance method. We use double electron−electron spin resonance (DEER) to study the denaturant-dependent equilibrium folding of flavodoxin. This flavoprotein is spin-labeled with MTSL ((1-oxy-,2,2,5,5-tetramethyl-D-pyrroline-3-methyl)-methanethiosulfonate) at positions 69 and 131. We find that nativelike spin-label separation dominates the distance distributions up to 0.8 M guanidine hydrochloride. At 2.3 M denaturant, the distance distributions show an additional component, which we attribute to a folding intermediate. Upon further increase of denaturant concentration, the protein expands and evidence for a larger number of conformations than in the native state is found. We thus demonstrate that DEER is a versatile technique to expand the arsenal of methods for investigating how proteins fold.



INTRODUCTION

Comprehending protein folding is one of the most fascinating research topics in current biochemistry research. Insight into this process requires structural information about the protein chain at different folding states. Novel methods are sought to obtain the corresponding experimental data. Advanced electron paramagnetic resonance (EPR) techniques are appropriate because they can track changes in distances between residues of a protein and can also probe protein dynamics during folding. Several EPR studies targeting protein folding, based on either steady-state[1,2] or flow-methods,[3−7] have been reported. These studies focus on folding-dependent local changes in mobility of an EPR-probe (see for example refs 1 and 3) or on detecting distance alterations between spin-labeled residues through EPR line broadening.[4] In a pioneering study, Dockter et al.[8] targeted the folding of a protein in the membrane by freezing the protein at different time points of folding.

Here we describe the use of double electron−electron spin resonance (DEER) to directly monitor *local* structure during protein folding by tracking distances between spin labels. We investigate the equilibrium folding of flavodoxin using guanidine hydrochloride (GuHCl) as denaturant. Flavodoxin folding has been the subject of intense study.[9−18] The native-state structure of flavodoxin is shown in Figure 1a. Site-directed mutagenesis was performed to replace the residue at position 131 by a cysteine (D131C mutation). The wild-type cysteine at position 69 serves as the second site for spin-labeling. In this study, we label both cysteines of the protein with the thiol-reactive nitroxide MTSL ((1-oxy-,2,2,5,5-tetramethyl-D-pyrroline-3-methyl)-methanethiosulfonate) to enable detection of DEER. In the following we refer to the spin-labeled protein as fdx-SL.

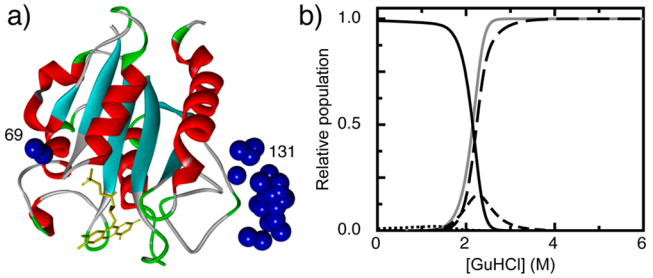


Figure 1. (a) Cartoon representation of flavodoxin (PDB entry 1YOB), showing in yellow the flavin mononucleotide cofactor. Blue spheres highlight representative locations of the spin-label nitroxides as derived from MMM simulation of monomer A in the crystallographic unit cell. (b) Normalized equilibrium population of native flavodoxin (thin solid line), native apoflavodoxin (dots), molten globule (short dashes), and unfolded protein (long dashes) as a function of denaturant concentration.[9] The gray, solid line represents the fraction of non-native molecules (i.e., the sum of MG and unfolded protein). Single-molecule FRET reveals that unfolded flavodoxin and MG are not distinct species because the MG forms continuously from the unfolded state (data to be published).

The in vitro folding of apoflavodoxin (i.e., flavodoxin without flavin mononucleotide (FMN)) occurs autonomously and involves significant formation of a folding intermediate. This intermediate is a molten globule (MG), a species that often

Special Issue: Wolfgang Lubitz Festschrift

Received: January 27, 2015
Revised: May 28, 2015
Published: June 23, 2015

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

forms during protein folding. The MG is an ensemble of interconverting conformers with a substantial amount of secondary structure but without the side-chain packing of native protein. Apoflavodoxin's MG is misfolded and kinetically off-pathway on the route to native protein.[17] Noncovalent binding of FMN to native apoflavodoxin is the last step in flavodoxin folding.[19] Figure 1b summarizes our current understanding of the equilibrium folding of flavodoxin and shows the denaturant-dependent population of folding species detected (i.e., unfolded, MG, native apoflavodoxin, and native flavodoxin).

This study demonstrates that DEER tracks flavodoxin folding through its ability to detect changes in local protein structure. Interpretation of the distance distributions as a function of denaturant concentration reveals the presence of a protein conformation that differs from the native state and has well-defined structure, indicative of a folding intermediate.

EXPERIMENTAL DETAILS

Purification and Spin-Labeling of Flavodoxin D131C. The flavodoxin variant D131C was generated and purified as described elsewhere.[20] This protein contains the wild-type cysteine residue at position 69 as well as a cysteine residue at position 131. Prior to labeling of the protein with the spin-label MTSL (Toronto Research Chemicals), the protein was unfolded in 5 M GuHCl and incubated with dithiothreitol DTT to reduce the thiol-groups of the cysteines. Subsequently, DTT was removed by gel filtration with a P6-DG column (Bio-Rad) that was equilibrated with 5 M GuHCl (Fluka) in 50 mM potassium phosphate (Sigma-Aldrich) buffer at pH 7.5. Labeling with MTSL was carried out during 16 h at 4 °C, using a 20-fold molar excess of spin-label over protein. The resulting double spin-labeled protein was purified from excess spin-label and GuHCl by gel filtration on a Superdex 75 10/30 HR column (Pharmacia) that was equilibrated in 100 mM potassium pyrophosphate (Sigma-Aldrich) at pH 6.0. Apoflavodoxin thus obtained was incubated with an excess of flavin mononucleotide (FMN) to reconstitute the holoprotein. Free FMN was separated from flavodoxin by gel filtration on a Superdex 75 10/30 HR column.

Sample Preparation. We made a stock solution of GuHCl in 100 mM potassium pyrophosphate (KPP_i), pH 6.0, with 20% glycerol. The buffer was 100 mM KPP_i, pH 6.0, with 20% glycerol. EPR samples of flavodoxin at various denaturant concentrations were prepared by appropriately mixing both solutions. The concentration of the denaturant was determined by measuring the refraction index and using the relation[21,22]

$$Z = 57.147\Delta N + 38.68(\Delta N)^2 - 91.60(\Delta N)^3 \quad (1)$$

where Z is the concentration of GuHCl in mol/L (M) and ΔN is the difference between the refraction indices of the buffer solution with and without GuHCl. Table 1 lists the concentrations GuHCl in the flavodoxin samples used in the DEER measurements.

In all measurements, protein concentration was ~0.1 mM. Protein solutions were placed in quartz tubes with inner and outer diameters (i.d. and o.d., respectively) of 2.3 and 3.0 mm, respectively. Samples that contained GuHCl were incubated at room temperature in the dark for 12 h. All samples were flash frozen in liquid nitrogen.

Spin-label concentration was determined by double integration of the 80 K cw-EPR spectrum and comparison to the spectrum of MTSL with known concentration. At least 83% of the cysteines of the flavodoxin sample were spin-labeled. In view of the error margin, which, because of uncertainties in the baseline correction of reference and sample spectra, in our case is in the order of 15%, this labeling degree is in agreement with the expected good accessibilities of the sites in the protein that are spin labeled in the unfolded state.

Table 1. Flavodoxin Folding from DEER Distance-Distribution Parameters[a]

concentration GuHCl (M)	coupled spins (%)[b]	mean distance $\langle r \rangle$ (nm)	standard deviation (nm)
0.3	109	4.08	0.45
0.8	117	4.01	0.51
2.3	94	4.25	0.58
3.5	40	5.30	1.55
4.5	48	5.38	1.62

[a]The regularization parameter α is 100. [b]Calculated from number of spins (see Figure 5a) as follows: 2 = 100%; 1 = 0%. The number of spins determined by DeerAnalysis has an overall error of ~5%, corresponding to an error of 10% in percentage of coupled spins.[32]

For modulation depth calibration, each of the DEER measurements was directly followed by a reference measurement. The reference sample is a solution of the biphenyl biradical PH2[23] in methyl tetrahydrofuran (MTHF) contained in a quartz tube with an i.d./o.d. of 2.3/3.0 mm. Oxygen was removed from solution by four repeated freeze−thaw cycles, followed by flame sealing to close the sample tube.

Continuous-Wave EPR Measurements. Continuous-wave EPR measurements were performed at 9.8 GHz using an ELEXSYS E 680 spectrometer (Bruker BioSpin GmbH, Rheinstetten, GE) equipped with a rectangular cavity and a cryostat. A flow of liquid helium was directed through the cavity to maintain a temperature of 80 K. Spectra were recorded at a microwave power of 0.16 mW with a field sweep of 20 mT and 2048 field points. Field modulation at a frequency of 100 kHz was employed with 0.2 mT amplitude. The time constant was 41 ms with a conversion time of 41 ms. Each sample was measured for 15−30 min, depending on the signal intensity.

DEER Measurements and Data Analysis. DEER measurements were performed at 9.3 GHz using an ELEXSYS E 680 spectrometer (Bruker BioSpin GmbH, Rheinstetten, GE) equipped with a 3 mm split-ring resonator (ER 4118XMS-3-W1) and a cryostat, Oxford model CF 935. All DEER experiments were performed at 40 K. The four-pulse sequence $(\pi/2)_1$-τ_1-$(\pi)_1$−$(\tau_1 + t)$-$(\pi)_2$−$(\tau_2$-$t)$-$(\pi)_1$−τ_2-[echo] was employed, where subscripts 1 and 2 indicate events occurring at the observer and pump frequency, respectively. Separation between frequencies was approximately 65 MHz. The observer pulses had lengths of 16 and 32 ns, and the pump pulse had a length of 16 ns. Delay times were τ_1 = 140 ns and τ_2 = 3.6 μs. Each DEER measurement lasted about 15 h.

The DeerAnalysis2011 and the DeerAnalysis2013 programs[24] were used to analyze the DEER data and obtain distance distributions. We assumed a homogeneous three-dimensional background. The validation option within the software was used to find a consistent background start, resulting in 600 ns for protein solutions with low denaturant concentrations (0, 0.3, 0.8, and 2.3 M GuHCl) and 1720 ns at high denaturant concentrations (3.5 and 4.5 M GuHCl). The optimum Tikhonov-regularization parameter, α, for distance distributions up to and including 2.3 M was 100 as determined by the L-curve criterion.[24] For modulation depth comparison at

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

GuHCl concentrations of 3.5 and 4.5 M, the same α value was chosen to ensure compatibility with the lower GuHCl concentration data (see Results and Discussion).

To model the spin-label linker conformation in the native state of flavodoxin, we used the multiscale modeling of macromolecular systems (MMM) software.[25]

RESULTS

We study folding of the double spin-labeled flavodoxin variant D131C (called fdx-SL), which contains a native cysteine at position 69 and an engineered cysteine at position 131. The protein sample is spin-labeled for 83% with MTSL (see Experimental Details) and contains no free spin-label, as determined by cw-EPR at room temperature (data not shown).

Figures 2 and 3 show results of frozen solution cw-EPR and DEER. The cw-EPR spectrum of a double spin-labeled species

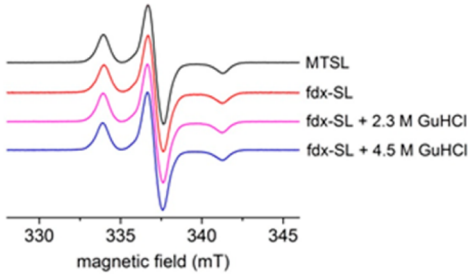


Figure 2. Continuous-wave EPR spectra of MTSL (black), native fdx-SL (red), and fdx-Sl at 2.3 (purple) and 4.5 M (blue) GuHCl. Spectra are recorded at 80 K and superimposed such that the central lines are aligned at the same field value.

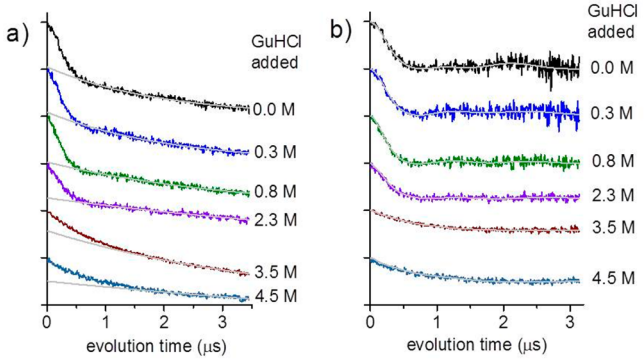


Figure 3. DEER time traces of fdx-SL at various concentrations of GuHCl. (a) Traces shown with their optimal background fit (in gray). The maximum of each individual trace is normalized to 1 and traces are vertically offset for better viewing. (b) Traces after correcting for background. In panel b, the fitted traces (in gray) correspond to distance distributions with regularization parameter α of 100.

is expected to show line broadening compared to the spectrum of a monoradical reference that is measured under identical conditions, if the spin-labels are separated by less than 2 nm.[26] Lineshapes of the cw-EPR spectra are identical to those of the spectrum of a monomeric reference (Figure 2), demonstrating the absence of line broadening. Figure 3 presents all DEER time traces collected. Figure 3a shows raw time traces with background, and Figure 3b provides traces corrected for background. First, we describe results obtained on native fdx-SL (i.e., at 0 M GuHCl). Then we interpret the data of the series of DEER experiments performed on fdx-SL at various GuHCl concentrations.

The DEER-time trace of native fdx-SL without GuHCl has an initial decay and some structure, but not much clearly visible modulation (Figure 3a). The modulation depth corresponds to what is expected for two coupled spins and shows that the entire protein population contributes to the distance distribution, which is shown in Figure 4 and discussed below. The

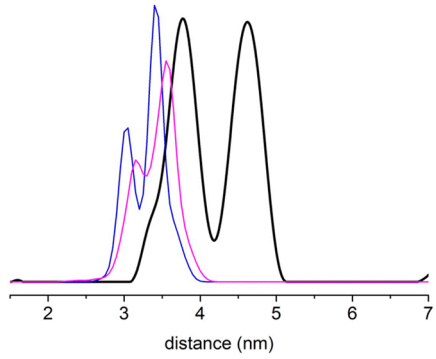


Figure 4. Experimentally derived distance distributions of native fdx-SL (black) and MMM-derived distance distributions based on the two flavodoxin monomers that are contained within the asymmetric unit of the crystal (PDB entry 1YOB; monomer A (blue), monomer B (purple)).

distribution of native fdx-SL has two peaks with maxima at 3.8 and 4.6 nm (i.e., a separation of 0.9 nm) and widths (full width at half-maximum (fwhm)) of 0.37 and 0.39 nm, respectively. The distribution centered at 3.8 nm has a shoulder at shorter distances, indicating an additional, third distance component.

To assay whether the obtained distance distributions are in agreement with the structure of native protein, possible spin-label linker conformations can be calculated by the rotamer-library-based method MMM.[25] This method takes into account that the measured distances are between the nitroxide groups of the two MTSL labels and that the nitroxide is separated from the protein backbone by a spin-label linker of about 0.5 nm. In addition to the rotamer library, the X-ray structure of flavodoxin (PDB entry 1YOB[27]) is also used as input. The protein crystallizes with two monomeric molecules in the asymmetric unit, called A and B. The MMM algorithm predicts spin-label positions for both monomers. The corresponding distance distributions are shown in Figure 4. Both monomers give rise to two peaks, which are separated by 0.38 nm (monomer A) or 0.40 nm (monomer B). These two peaks derive from two families of linker conformations because, in the X-ray structure, the protein backbone has a unique conformation. Possible locations of the nitroxides are shown as spheres in Figure 1. The spin-label at position 69 has few accessible conformations (two in monomer A (Figure 1a) and five in monomer B (not shown)), whereas the spin-label at position 131 shows an extensive cloud of accessible nitroxide positions (Figure 1a). This observation suggests that the two peaks seen in the distance distributions of flavodoxin in MMM (Figure 4) originate, at least in part, from different conformations of spin-label at position 69. In the measured distribution, the peaks are separated by 0.9 nm, whereas MMM show peak separations of 0.38 and 0.40 for monomers A and B, respectively. Also, the centers of the distance distributions of protein monomers A and B differ by 0.15 nm. Figure 4 further highlights that these centers are located at a shorter distance than the measured one (for discussion, see below).

The overall appearance of the DEER decays (Figure 3) and the modulation depth, given as the percentage of coupled spins

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

(Table 1), show that with respect to the GuHCl concentration the DEER data fall into two regimes. From 0.3 to 2.3 M GuHCl, the modulation depth accounts for >94% of the protein population. In that regime, the time traces (Figure 3) have a pronounced initial decay and some structure at later times of the trace, i.e., above 0.6 μs. From 2.3 to 3.5 M GuHCl, the protein population that contributes to the observed distance distribution halves, showing that at 3.5 M GuHCl a large fraction of the protein population has conformations where the interspin distances of the two nitroxides fall outside the measurement range of DEER (i.e., below 2 or above 6.5 nm). Also, the corresponding DEER curves show no structure and are less steep than those at lower GuHCl concentrations at early times. Table 1 also lists the calculated mean distances $\langle r \rangle$ of the distance distributions at various concentrations of denaturant and the corresponding overall widths of the distance distributions, given as standard deviations. The mean distance and the overall width remain almost constant between 0.3 and 2.3 M GuHCl, and upon further increase in denaturant concentration, the distance increases by about 1 nm with concomitant tripling of the standard deviation (Table 1), reflecting flavodoxin unfolding. Figure 5a shows that these parameters and the percentage of coupled spins have the same dependence on GuHCl concentration, suggesting them as overall indicators of folding.

Figure 5b,c shows distance distributions of flavodoxin at several GuHCl concentrations. Figure 5b shows the distributions obtained from Tikhonov regularization, regularization parameter[28] α of 100. Given the low modulation depth and smooth decay of the DEER time trace, the exact shapes of the distance distributions at 3.5 and 4.5 M GuHCl are less well-defined than those of the distributions obtained at lower denaturant concentrations; therefore, the distance distributions of fdx-SL at these high GuHCl concentrations will not be considered further.

Figure 5c shows Gaussian curves that are fitted directly to the time dependence of the DEER curves for GuHCl concentrations between 0.3 and 2.3 M. The minimum number of Gaussians was found to be two for 0.3 and 0.8 M GuHCl and three for 2.3 M GuHCl (for details, see Appendix). In Table 2, the parameters of these Gaussians are given. To monitor protein folding, we use the distance distribution of 0.3 M GuHCl as a reference. This distance distribution differs from that of fdx-SL without GuHCl (Figure 4), although in both samples the protein is fully folded.[29]

The distance distributions of flavodoxin at 0.3 and 0.8 M GuHCl are well described by two Gaussian functions (labeled N_1 and N_2 in Table 2), with slightly differing parameters (Table 2). For example, upon going from 0.3 to 0.8 M GuHCl, peak N_1 gains intensity from peak N_2 and both peaks broaden slightly (e.g., N_2 by 0.1 nm). Below, we show that peaks N_1 and N_2 arise from native protein.

For the GuHCl concentration of 2.3 M, the shape of the DEER decay for times above 0.6 μs is not flat; at 1.4 μs a broad maximum is seen, showing that also at that concentration a structured form of the protein prevails. A distribution corresponding to slightly modified parameters of the native peaks, or two peaks in general, are not satisfactory (see Appendix). Fitting the data by three Gaussians significantly improves the agreement with the experimental data, and the fit parameters are given in Table 2.

We expect the presence of a folding intermediate of flavodoxin at 2.3 M GuHCl (Figure 1b). Two of the three Gaussians (i.e., N_1 and N_2 in Table 2) are similar to the distributions observed

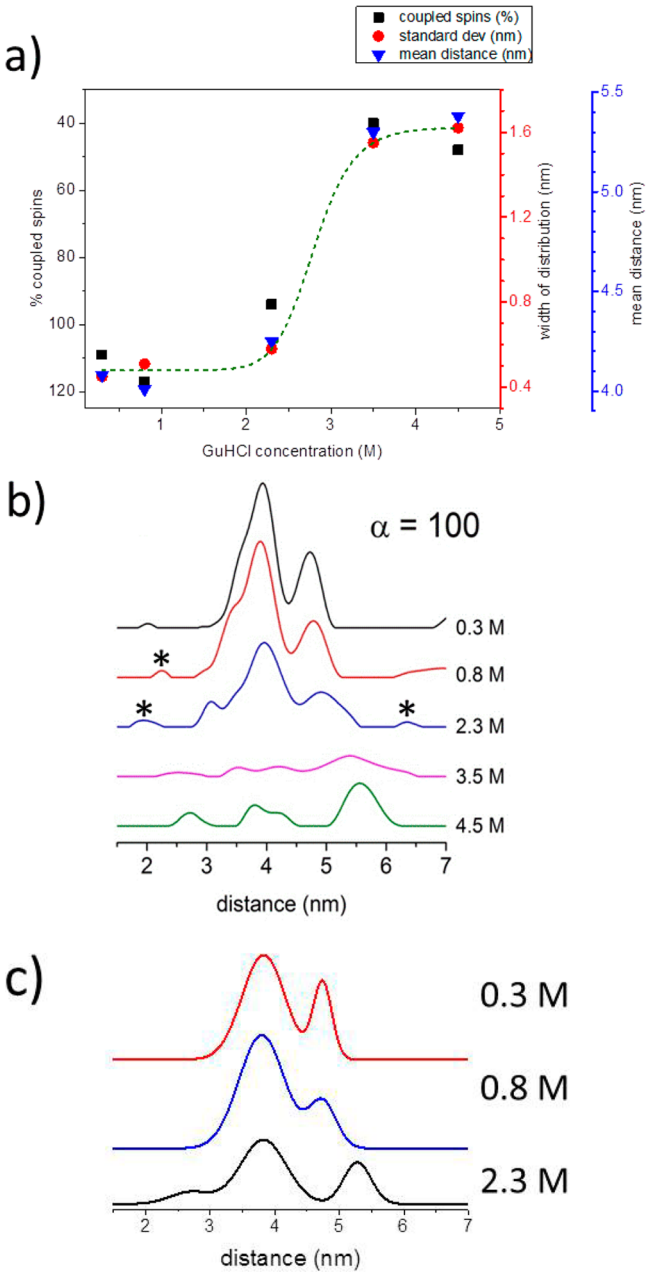


Figure 5. Effect of GuHCl concentration on DEER parameters of fdx-SL. (a) Global folding parameters: percent of coupled spins (black squares), the overall mean distance (blue triangles), and widths of the entire distance distribution (red dots). The green dotted line is a guide for the eye. (b) Distance distributions from Tikhonov analysis of the DEER time traces; regularization parameter α is 100. Peaks marked by an asterisk are not significant according to the suppression tool in DEER analysis. (c) Distance distributions from direct fitting of Gaussians to the DEER time traces.

at the lower GuHCl concentrations mentioned and are due to native protein. A new distribution arises at 2.7 nm (labeled MG in Table 2) which must be due to folding intermediate. The contribution of MG accounts for 11% of the total integral of the three distance distributions observed for flavodoxin at 2.3 M GuHCl. Below, we discuss the interpretation and significance of these distance distributions.

DISCUSSION

In the following we discuss DEER-derived distance distributions of native flavodoxin. Then we interpret denaturant-induced

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

Table 2. Parameters Obtained from Fitting of Gaussian Functions to DEER Time Traces Obtained for Denaturant-Dependent Flavodoxin Folding

		GuHCl concentration		
peaks		0.3 M	0.8 M	2.3 M
N_1	distance (nm)	3.8	3.8	3.8
	width (nm)	0.47	0.50	0.50
	contribution %	76	80	64
N_2	distance (nm)	4.7	4.7	5.3
	width (nm)	0.20	0.30	0.30
	contribution %	24	20	25
MG	distance (nm)	N.A.	N.A.	2.7
	width (nm)			0.45
	contribution %			11

changes in flavodoxin structure in relation to protein folding.

Native Protein. In its native state, flavodoxin is a globular protein with a well-defined structure (Figure 1). Spin labeling of apoflavodoxin does not impair FMN binding. Because binding of FMN to native apoflavodoxin occurs primarily through a very specific combination and geometry of hydrogen bonds and aromatic interactions, this observation shows that spin-labeled flavodoxin attains its native conformation. We observe two peaks in the DEER-derived distance distributions of native protein (Figure 4), the widths of which are typical for intrinsic flexibility of spin-label attached to a well-structured protein.[30]

Modeling the spin-label side-chain conformations with MMM also yields two-peaked distance distributions, arising from two families of spin-label linker conformations, albeit at shorter distances and with a smaller peak separation than those observed experimentally.

Several origins of the differences between the parameters derived from DEER as compared to those derived from the X-ray structure of flavodoxin are possible. First, in solution, a more extended conformation of the loop between residues 126 to 148 may play a role. Conformational dynamics in this loop would affect the distance between the spin labels at positions 69 and 131 because this loop contains one of the spin-labels. However, NMR spectroscopy shows that flavodoxins are rather rigid, although some effect of mutagenesis and subsequent spin labeling cannot be excluded. The second explanation for the overall longer distances measured by DEER is conformational variability of the cysteine at position 69. In native flavodoxin, the side-chain of this cysteine resides in the interior of the protein, and hence the low number of accessible conformations obtained by MMM for this residue. In spin-labeled flavodoxin this side-chain more likely switches to a conformation that projects the spin-label to the exterior of the protein, and as a result the distance between the nitroxides of both spin-labels increases. The separation between the two peaks in the experimental distance distribution is larger than what is commonly observed for different families of spin label conformations, suggesting that also backbone differences come into play. Such changes in side-chain and backbone conformations are not accounted for in MMM and can explain the observed difference between MMM and DEER-derived distance distributions of native flavodoxin. We reason that the distance distribution obtained by DEER is compatible with what is known about the native structure of flavodoxin.

At 0.3 M, GuHCl flavodoxin is still in the fully folded state.[10,31] However, the corresponding Tikhonov-derived distance distribution (Figure 5b) differs from that of flavodoxin without denaturant. We attribute this difference to a local influence of GuHCl on spin-label conformation.

DEER Tracks Flavodoxin Folding. For flavodoxin, we detect changes of DEER parameters upon altering denaturant concentration, which shows the capacity of DEER to follow protein folding. To exclude that the observed changes are caused by local influence of GuHCl on spin-label conformation, we use in the following analysis the distance distribution of native flavodoxin at 0.3 M GuHCl as signature of natively folded protein.

According to DEER data, the most notable change in flavodoxin structure happens between 2.3 and 3.5 M GuHCl (Figure 5 and Table 1). Up to 2.3 M denaturant, the distance distributions account for almost the entire protein population and the shape of this distribution is that of a structured protein. Upon increasing denaturant concentration, broadening of the distance distributions occurs, as well as a concomitant increase of the average distance, $\langle r \rangle$. Now, a larger fraction of fdx-SL molecules has distances that fall outside the DEER observation window (i.e., below 2 or above 6 nm) (Figure 5a). Existence of protein molecules with interspin distances below 2 nm can be excluded because these species would give rise to dipolar broadening in frozen solution cw-EPR spectra, a phenomenon which we do not observe (see Results). Hence, upon increasing GuHCl concentration above 2.3 M, a fraction of the flavodoxin molecules has interspin distances exceeding 6.5 nm. In conclusion, flavodoxin is relatively compact at denaturant concentrations below 2.3 M GuHCl. In contrast, a larger number of conformations, including extended ones typical for unfolded protein, exist at GuHCl concentrations of 3.5 M and above. As discussed above, for GuHCl concentrations above 2.3 M, the distance distributions are not meaningful because of the large width of the distribution and the low modulation amplitude; although they are shown in Figure 5, we do not consider them further.

The Tikhonov-derived distance distributions provide a good representation of the overall development of folding. However, the fine structure of the distance distributions can have artifacts, particularly for distance distributions in which broad and narrow components occur. Therefore, we compare different approaches to the analysis of the DEER traces obtained between 0.3 and 2.3 M GuHCl (see Appendix). Gaussian-distance distributions fitted to the time traces show that for 0.3 and 0.8 M GuHCl concentrations, a reasonable representation of the data can be obtained with two Gaussians (Table A1; for details see Appendix). The Tikhonov-derived distance distributions suggest that the peak N_1 could be constituted by two peaks, N_1 and N_2 (Table A2), which could explain the large width of the distance peak N_1 in Table 2.

For the GuHCl concentration of 2.3 M, the shape of the DEER decay above 0.6 μs has structure and requires a minimum of three Gaussians for fitting. Models with distance distributions derived from the native protein or distributions in which the native distance distribution, either from 0.3 M or from 0.8 M GuHCl concentration, is combined with a broad component do not satisfy the observation, excluding a superposition of the native and a completely random state. Acceptable solutions require three Gaussians of width below 0.5 nm (see Appendix). In the best solution, the peak N_2 shifts to a longer distance, by 0.5 nm. The widths of N_1 and N_2 and their relative ratios are similar to the ratio at 0.8 M GuHCl, showing considerable native state population. The new peak,

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

referred to as MG, is at a shorter distance than N_1 and N_2, and has a width which is in between N_1 and N_2, underlining a defined state of the protein. This observation shows that the protein state that gives rise to the MG peak must be rather compact and suggests it as a folding intermediate. We estimate the population corresponding to this state, taking into account the contribution of peaks N_1 and N_2, as well as the 6% of protein molecules with distances that fall outside the DEER range. Assuming that peaks N_1 and N_2 represent the native state, a total of 84% of flavodoxin molecules are native, 10% are folding intermediate, and 6% are unfolded. This is a lower limit for the population of the folding intermediate because if there was a folding intermediate that had an interspin distance similar to native protein, it could not be distinguished from native protein in the present analysis. The shift of N_2 to larger distances that occurs for 2.3 M GuHCl compared to the lower GuHCl concentrations may be an indication for additional folding−intermediate signature. The MG peak is at shorter distance than the peaks N_1 and N_2, suggesting a more compact arrangement of the protein in the region probed.

Shorter distances observed for the MG peak were also detected by Förster resonance energy transfer (FRET) measurements on apoflavodoxin. These experiments targeted the denaturant-dependent distance between the same two cysteines as the DEER study presented here.[10] FRET shows that below about 1.5 M GuHCl, the MG intermediate has an interdye distance shorter than that of native protein, which becomes apparent from a higher FRET efficiency. Note that the EPR measurements were done at 40 K, whereas the FRET experiments were performed at 298 K.

The presented case demonstrates that DEER tracks denaturant-dependent equilibrium folding and enables detection of distinct folding intermediates. As demonstrated by Dockter et al.,[8] time-dependent folding can also be addressed by DEER and even target protein assembly into a membrane, showing the versatility of the approach.

CONCLUSIONS

We show that DEER has the capacity to reveal important features of protein folding. Denaturant-dependent folding of flavodoxin could be tracked because DEER detects folding-related conformational changes. The methodology provides distance distributions between nitroxide labels that are covalently attached to the protein under investigation and shows in the case of flavodoxin folding involvement of MG intermediate. The width of a distance distribution informs about the degree of structuring of the folding state involved and allows discrimination between folding intermediate and unfolded protein. We conclude that DEER is a versatile technique to expand the current arsenal of methods for investigating how proteins fold.

APPENDIX

The challenge of protein folding by DEER is that broad distance distributions expected for fully unfolded protein can coexist with narrow distributions deriving from natively folded protein. For data analysis in the present study we tried many approaches, and here we present those that are relevant to describe the DEER data.

Gaussian Fits to DEER Traces

The DEER data of fdx-SL at GuHCl concentrations between 0.3 and 2.3 M were fitted with several combinations of Gaussians. All fits with a single Gaussian, i.e., those with broad, unstructured distance distributions, had significant deviation from the experimental data, showing that at these GuHCl concentrations, multiple distances with narrower distance distributions prevail. Beyond a certain width, broad distance distributions lead to flat DEER curves at times above 0.6 μs, neglecting structure that is present in the experimental curves.

Fits with two Gaussians yield satisfactory agreement with the data at 0.3 and 0.8 M GuHCl (Figure A1) using the parameters given in Table A1. Significantly increasing the width of the Gaussians or combinations of narrow and broad Gaussians did not fit the data. The Gaussian distributions for 0.8 M GuHCl differ from those of 0.3 M by a slight increase in width of both components and an increase in the contribution of the shorter distance component.

For fdx-SL at 2.3 M GuHCl, several scenarios were tried: (i) The first was fitting with two Gaussians with parameters that are close to those of the lower GuHCl concentrations (Table A1). As seen in Figure A1a, the fit deviates significantly from the data, showing that the data cannot be explained by the distance distribution of native fdx-SL. (ii) Attempts to optimize the parameters of the two Gaussians did not improve the agreement significantly, and in all cases corrected Akaike information criterion (AICc) scores[33] were higher than those with fits with three Gaussians described below. (iii) Fits with three Gaussians (Figure A1b,c) agree much better with the experimental data. [We thank one of the reviewers for providing the routine to perform three Gaussian fits in DEER analysis.] Two options are shown, traces b and c of Figure A1, to illustrate to which degree the fits are sensitive to the parameters of the Gaussians. The AICc scores of the latter fits, given in Table A1, are significantly lower than those of the fits with two Gaussians, underlining that the improvement for the model of three Gaussians is statistically significant. We therefore conclude that three Gaussians provide a better representation of the data than the two-Gaussian model or any combination with Gaussians of widths larger than 0.6 nm.

Gaussian Fits to the Tikhonov DEER Distance Distributions

Considering that the results of Tikhonov regularization fit the time traces better than the limited number of Gaussians directly

Table A1. Parameters of Gaussian Fits to the DEER Data of fdx-SL at GuHCl Concentrations from 0.3 to 2.3 M Including Different Models for the 2.3 M GuHCl Case[a]

	peak MG			peak N_1			peak N_2			
	r (nm)	Δw (nm)	weight (%)	r (nm)	Δw (nm)	weight (%)	r (nm)	Δw (nm)	weight (%)	AICc score
0.3 M	n.a.	n.a.	n.a.	3.8	0.5	76	4.7	0.2	24	−2529
0.8 M	n.a.	n.a.	n.a.	3.8	0.5[b]	80	4.7	0.3[b]	20	−2910
2.3 M	n.a.	n.a.	n.a.	3.8[b]	0.5[b]	72	4.7[b]	0.3[b]	28	−3084
2.3 M u.c.	2.7	0.4	11	3.8	0.5	64	5.28	0.3	25	−3221
2.3 M adj.	3.1[b]	0.2[b]	13	3.95	0.4[b]	57	5.1[b]	0.6[b]	30	−3199

[a]Distance of peak, r; width, Δw; contribution, weight; n.a., not applicable; u.c., unconstrained; adj., adjusted. [b]Parameter kept fixed in the fit.

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514

Table A2. Parameters Obtained from Fitting of Gaussian Functions to DEER Distance Distributions Obtained for Denaturant-Dependent Flavodoxin Folding[a]

		GuHCl concentration		
peaks		0.3 M	0.8 M	2.3 M
N_1	X_c	3.58	3.40	3.43
	W	0.34	0.37	0.27
	A	21%	19%	6%
N_2	X_c	3.96	3.91	3.97
	W	0.37	0.45	0.55
	A	52%	58%	60%
N_3	X_c	4.72	4.77	4.94
	W	0.36	0.41	0.58
	A	27%	23%	26%
MG	X_c	N.A.	N.A.	3.07
	W			0.25
	A			8%

[a] X_c, centre of Gaussian (nm); W width at half height (nm); A, area under corresponding Gaussian, expressed as percentage of the sum of the integrals of all Gaussians. Regularization parameter α equals 100.

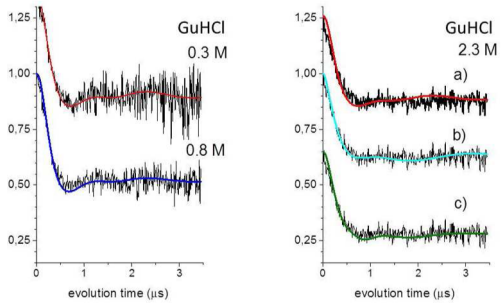


Figure A1. Direct fits to the DEER time traces of fdx-SL at 0.3, 0.8, and 2.3 M GuHCl. All raw time traces were treated as those described for Figure 3. For 2.3 M GuHCl, right-hand side of the figure, three possibilities are shown: (a) two Gaussian fit, (b) three Gaussian fit with parameters given in Table A1 "unconstrained" and (c) parameters "adjusted". See text for details.

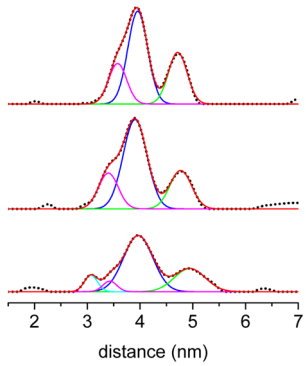


Figure A2. Gaussian fits to the distance distributions obtained by Tikhonov regularization of the DEER time traces.

fitted to the time traces (see above), we also considered which Gaussian functions fit the most intense peaks observed in the Tikhonov-derived distance distribution (Figure 5b). During fitting, several small peaks marked by an asterisk in Figure 5b need not be taken into account. These latter peaks are judged to be insignificant according to the suppression tool of the DEER analysis software because their contribution does not significantly alter the DEER time trace given the noise of the curve involved. The distributions obtained at 0.3 and 0.8 M denaturant are well described by the sum of three Gaussian functions (labelled N_1–N_3 in Table A2), with slightly differing parameters. For example, upon going from 0.3 to 0.8 M GuHCl, peak N_1 shifts by 0.2 nm to lower distances. Upon this change in denaturant concentration, the widths of peaks N_1–N_3 increase, with the largest increase (i.e. 0.08 nm) observed for peak N_2. The results of this treatment of the data agree in large measure with the direct fitting of Gaussians to the DEER decay data, discussed above, but it could be subject to artifacts, in particular the similar width of all four components for the data obtained at 2.3 M GuHCl could be an artifact of the Tikhonov regularization (Figure A2).

AUTHOR INFORMATION

Corresponding Author

*E-mail: huber@physics.leidenuniv.nl.

Present Address

[§]S.L.: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, The Netherlands.

Author Contributions

[#]M.v.S. and S.L.: These authors contributed equally. The manuscript was written through contributions of all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge Edgar Groenen for constant support and fruitful discussions. This work is part of the research program of the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)", which is financially supported by the "Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)", Grant (03BMP03). Financial support by NWO and by an NWO CW ECHO Grant (700.58.014) is also gratefully acknowledged.

ABBREVIATIONS

DTT, dithiothreitol; DEER, double electron–electron resonance; FMN, flavin mononucleotide; GuHCl, guanidine hydrochloride; MG, molten globule; MMM, modeling of macromolecular systems; MTSL, ((1-oxy-,2,2,5,5-tetramethyl-D-pyrroline-3-methyl)-methanethiosulfonate)

REFERENCES

(1) Kreimer, D. I.; Szosenfogel, R.; Goldfarb, D.; Silman, I.; Weiner, L. 2-State Transition Between Molten Globule and Unfolded States of Acetylcholinesterase As Monitored by Electron-Paramagnetic-Resonance Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91* (25), 12145–12149.

(2) Hubbell, W. L.; Mchaourab, H. S.; Altenbach, C.; Lietzow, M. A. Watching Proteins Move Using Site-Directed Spin Labeling. *Structure* **1996**, *4* (7), 779–783.

(3) DeWeerd, K.; Grigoryants, V.; Sun, Y. H.; Fetrow, J. S.; Scholes, C. P. EPR-detected Folding Kinetics of Externally Located Cysteine-directed Spin-labeled Mutants of iso-1-cytochrome c. *Biochemistry* **2001**, *40* (51), 15846–15855.

(4) Grigoryants, V. M.; DeWeerd, K. A.; Scholes, C. P. Method of Rapid Mix EPR Applied to the Folding of Bi-spin-labeled Protein as a Probe for the Dynamic Onset of Interaction Between Sequentially Distant Side Chains. *J. Phys. Chem. B* **2004**, *108* (27), 9463–9468.

(5) Grigoryants, V. M.; Veselov, A. V.; Scholes, C. P. Variable Velocity Liquid Flow EPR Applied to Submillisecond Protein Folding. *Biophys. J.* **2000**, *78* (5), 2702–2708.

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507–13514

(6) Qu, K. B.; Vaughn, J. L.; Sienkiewicz, A.; Scholes, C. P.; Fetrow, J. S. Kinetics and Motional Dynamics of Spin-labeled yeast iso-1-cytochrome *c*: 1. Stopped-flow Electron Paramagnetic Resonance as a Probe for Protein Folding/Unfolding of the C-terminal Helix Spin-Labeled at Cysteine 102. *Biochemistry* **1997**, *36* (10), 2884−2897.
(7) Sienkiewicz, A.; Ferreira, A. M. D.; Danner, B.; Scholes, C. P. Dielectric Resonator-based Flow and Stopped-flow EPR with Rapid Field Scanning: A Methodology for Increasing Kinetic Information. *J. Magn. Reson.* **1999**, *136* (2), 137−142.
(8) Dockter, C.; Volkov, A.; Bauer, C.; Polyhach, Y.; Joly-Lopez, Z.; Jeschke, G.; Paulsen, H. Refolding of the Integral Membrane Protein Light-harvesting complex II Monitored by Pulse EPR. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (44), 18485−18490.
(9) Bollen, Y. J. M.; Nabuurs, S. M.; van Berkel, W. J. H.; van Mierlo, C. P. M. Last In, First Out. *J. Biol. Chem.* **2005**, *280* (9), 7836−7844.
(10) Lindhoud, S.; Westphal, A. H.; Borst, J. W.; van Mierlo, C. P. M. Illuminating the Off-Pathway Nature of the Molten Globule Folding Intermediate of an α-β Parallel Protein. *PLoS One* **2012**, *7* (9).
(11) Nabuurs, S. M.; de Kort, B. J.; Westphal, A. H.; van Mierlo, C. P. M. Non-native Hydrophobic Interactions Detected in Unfolded apoflavodoxin by Paramagnetic Relaxation Enhancement. *Eur. Biophys. J. Biophys. Lett.* **2010**, *39* (4), 689−698.
(12) Nabuurs, S. M.; van Mierlo, C. P. M. Interrupted Hydrogen/Deuterium Exchange Reveals the Stable Core of the Remarkably Helical Molten Globule of α-β Parallel Protein Flavodoxin. *J. Biol. Chem.* **2010**, *285* (6), 4165−4172.
(13) Nabuurs, S. M.; Westphal, A. H.; aan den Toorn, M.; Lindhoud, S.; van Mierlo, C. P. M. Topological Switching Between an α-β Parallel Protein and a Remarkably Helical Molten Globule. *J. Am. Chem. Soc.* **2009**, *131* (23), 8290−8295.
(14) Nabuurs, S. M.; Westphal, A. H.; van Mierlo, C. P. M. Noncooperative Formation of the Off-Pathway Molten Globule during Folding of the α-β Parallel Protein Apoflavodoxin. *J. Am. Chem. Soc.* **2009**, *131* (7), 2739−2746.
(15) Nabuurs, S. M.; Westphal, A. H.; van Mierlo, C. P. M. Extensive Formation of Off-Pathway Species during Folding of an α-β Parallel Protein Is Due to Docking of (Non)native Structure Elements in Unfolded Molecules. *J. Am. Chem. Soc.* **2008**, *130* (50), 16914−16920.
(16) van Mierlo, C. P. M.; Van Dongen, W. M. A. M.; Vergeldt, F.; Van Berkel, W. J. H.; Steensma, E. The Equilibrium Unfolding of Azotobacter vinelandii apoflavodoxin II Occurs via a Relatively Stable Folding Intermediate. *Protein Sci.* **1998**, *7* (11), 2331−2344.
(17) Bollen, Y. J. M.; Sanchez, I. E.; van Mierlo, C. P. M. Formation of On- and off-pathway Intermediates in the Folding Kinetics of Azotobacter vinelandii apoflavodoxin. *Biochemistry* **2004**, *43* (32), 10475−10489.
(18) Alagaratnam, S.; Van Pouderoyen, G.; Pijning, T.; Dijkstra, B. W.; Cavazzini, D.; Rossi, G. L.; Van Dongen, W. M. A. M.; van Mierlo, C. P. M.; Van Berkel, W. J. H.; Canters, G. W. A Crystallographic Study of Cys69Ala flavodoxin II from Azotobacter vinelandii: Structural Determinants of Redox Potential. *Protein Sci.* **2005**, *14* (9), 2284−2295.
(19) Bollen, Y. J. M.; Nabuurs, S. M.; Van Berkel, W. J. H.; van Mierlo, C. P. M. Last In, First Out. *J. Biol. Chem.* **2005**, *280* (9), 7836−7844.
(20) Lindhoud, S.; Westphal, A. H.; Borst, J. W.; van Mierlo, C. P. M. Illuminating the Off-Pathway Nature of the Molten Globule Folding Intermediate of an α-β Parallel Protein. *PLoS One* **2012**, *7* (9).
(21) Shirley, B. A. Urea and Guanidine Hydrochloride Denaturation Curves. In *Protein Stability and Folding: Theory and Practice*; Humana Press: Totowa, NJ, 1995; pp 177−190.
(22) Nozaki, Y. *The Preparation of Guanidine Hydrochloride*; Academic Press: New York, 1972; Vol. *26*.
(23) Corvaja, C.; Giacometti, G.; Kopple, K. D. Ziauddin Electron Spin Resonance Studies of Nitroxide Radicals and Biradicals in Nematic Solvents. *J. Am. Chem. Soc.* **1970**, *92* (13), 3919−3924.
(24) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006 - A Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30* (3−4), 473−498.
(25) Polyhach, Y.; Bordignon, E.; Jeschke, G. Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Phys. Chem. Chem. Phys.* **2011**, *13* (6), 2356−2366.
(26) Jeschke, G. Determination of the Nanostructure of Polymer Materials by Electron Paramagnetic Resonance Spectroscopy. *Macromol. Rapid Commun.* **2002**, *23* (4), 227−246.
(27) Alagaratnam, S.; Van Pouderoyen, G.; Pijning, T.; Dijkstra, B. W.; Cavazzini, D.; Rossi, G. L.; Van Dongen, W. M. A. M.; van Mierlo, C. P. M.; van Berkel, W. J. H.; Canters, G. W. A Crystallographic Study of Cys69Ala flavodoxin II from Azotobacter vinelandii: Structural Determinants of Redox Potential. *Protein Sci.* **2005**, *14* (9), 2284−2295.
(28) Jeschke, G.; Koch, A.; Jonas, U.; Godt, A. Direct Conversion of EPR Dipolar Time Evolution Data to Distance Distributions. *J. Magn. Reson.* **2002**, *155* (1), 72−82.
(29) Lindhoud, S.; Westphal, A. H.; Visser, A. J. W. G.; Borst, J. W.; van Mierlo, C. P. M. Fluorescence of Alexa Fluor Dye Tracks Protein Folding. *PLoS One* **2012**, *7* (10), e46838 DOI: 10.1371/journal.-pone.0046838.
(30) Finiguerra, M. G.; Prudencio, M.; Ubbink, M.; Huber, M. Accurate Long-range Distance Measurements in a Doubly Spin-labeled Protein by a Four-pulse, Double Electron-electron Resonance Method. *Magn. Reson. Chem.* **2008**, *46* (12), 1096−1101.
(31) Steensma, E.; van Mierlo, C. P. M. Structural Characterisation of Apoflavodoxin Shows that the Location of the Stable Nucleus Differs Among Proteins with a Flavodoxin-like Topology. *J. Mol. Biol.* **1998**, *282* (3), 653−666.
(32) Bode, B. E.; Margraf, D.; Plackmeyer, J.; Dürner, G.; Prisner, T. F.; Schiemann, O. Counting the Monomers in Nanometer-Sized Oligomers by Pulsed Electron−Electron Double Resonance. *J. Am. Chem. Soc.* **2007**, *129* (21), 6736−6745.
(33) Dalmas, O.; Hyde, H.; Hulse, R. E.; Perozo, E. Symmetry-Constrained Analysis of Pulsed Double Electron-Electron Resonance (DEER) Spectroscopy Reveals the Dynamic Nature of the KcsA Activation Gate. *J. Am. Chem. Soc.* **2012**, *134* (39), 16360−16369.

DOI: 10.1021/acs.jpcb.5b00856
J. Phys. Chem. B 2015, 119, 13507−13514