4722–4730 *Nucleic Acids Research, 2006, Vol. 34, No. 17* *Published online 10 September 2006*
doi:10.1093/nar/gkl546

Site-directed spin labeling measurements of nanometer distances in nucleic acids using a sequence-independent nitroxide probe

Qi Cai[1], Ana Karin Kusnetzow[3], Wayne L. Hubbell[3], Ian S. Haworth[4], Gian Paola C. Gacho[1], Ned Van Eps[3], Kálmán Hideg[5], Eric J. Chambers[4] and Peter Z. Qin[1,2,*]

[1]Department of Chemistry and [2]Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089-0744, USA, [3]Jules Stein Eye Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA, [4]Department of Pharmaceutical Sciences, University of Southern California, Los Angeles, CA 90089, USA and [5]Institute of Organic and Medical Chemistry, University of Pécs, H-7643, Pécs, P.O. Box 99, Hungary

Received May 5, 2006; Revised June 30, 2006; Accepted July 13, 2006

ABSTRACT

In site-directed spin labeling (SDSL), local structural and dynamic information is obtained via electron paramagnetic resonance (EPR) spectroscopy of a stable nitroxide radical attached site-specifically to a macromolecule. Analysis of electron spin dipolar interactions between pairs of nitroxides yields the inter-nitroxide distance, which provides quantitative structural information. The development of pulse EPR methods has enabled such distance measurements up to 70 Å in bio-molecules, thus opening up the possibility of SDSL global structural mapping. This study evaluates SDSL distance measurement using a nitroxide (designated as R5) that can be attached, in an efficient and cost-effective manner, to a phosphorothioate backbone position at arbitrary DNA or RNA sequences. R5 pairs were attached to selected positions of a dodecamer DNA duplex with a known NMR structure, and eight distances, ranging from 20 to 40 Å, were measured using double electron-electron resonance (DEER). The measured distances correlated strongly (R^2 = 0.98) with the predicted values calculated based on a search of sterically allowable R5 conformations in the NMR structure, thus demonstrating accurate distance measurements using R5. Furthermore, distance measurement in a 42 kD DNA was demonstrated. The results establish R5 as a sequence-independent probe for global structural mapping of DNA and DNA–protein complexes.

INTRODUCTION

Structural and dynamic information on nucleic acids is key to a fundamental understanding of their functions. Methods that provide the capability of determining atomic-resolution structures of DNA and RNA, such as X-ray crystallography and NMR spectroscopy, have yielded deep insight into the structure-function relationship in nucleic acids. However, X-ray crystallography studies are constrained by the need of growing crystals and NMR studies of DNA or RNA molecules longer than 40-nucleotide (nt) remain challenging. Therefore, techniques that can provide structural and dynamic information on high molecular weight systems under physiological conditions are of great interest in structure-function relationship studies of DNA and RNA.

Studies reported here explored nanometer distance measurement in nucleic acids using the technique of site-directed spin labeling (SDSL). SDSL derives information on the local environment of a macromolecule via electron paramagnetic resonance (EPR) spectroscopy of a site-specifically attached stable nitroxide radical (1). In protein studies, SDSL has been established as a tool for investigating solution structure and dynamics, and has been particularly successful in those systems that are difficult to study with X-ray crystallography and NMR spectroscopy [see reviews (2–7)]. SDSL studies of nucleic acids, although having shown great potentials, lag behind the more mature protein SDSL applications. Previous SDSL studies of singly labeled nucleic acids have established that structural and dynamical information at the level of an individual nucleotide can be derived from EPR spectral analysis that reveals the dynamics of the nitroxide [see a recent review (8)]. Recently, solution conformation changes in RNA molecules with complex 3D structures have been studied via monitoring nitroxide spectral changes (9–12), and methods are being developed to systematically extract local

*To whom correspondence should be addressed. LJS-251, 840 Downey Way, Los Angeles, CA 90089 0744, USA. Tel: +1 213 821 2461; Fax: +1 213 740 0930; Email: pzq@usc.edu

structural and dynamic information from the observed nitroxide spectra (13).

In SDSL, distances between two nitroxides can be determined by measuring the electron spin dipolar coupling (2,4,7,14–17). Recent developments in pulse EPR techniques have enabled inter-nitroxide distance determinations up to 75 Å in synthetic polymers and polypeptides (18–22), proteins (23–29) and nucleic acids (30–33). The ability of SDSL in measuring nanometer distances renders it an attractive tool for global structural mapping of bio-molecules. Compared to fluorescence resonance energy transfer (FRET), the pulse EPR methods have the advantages that: (i) two identical spin labels can be employed, thus facilitating labeling, while in FRET two different fluorophores must be selected based on the expected distances; (ii) spin labels are smaller than a majority of fluorophores, presumably leading to less perturbation and a more precisely defined inter-probe distance and (iii) distances measured in pulse EPR methods are precise over a broader range (21). Inter-nitroxide distances in nucleic acids have been measured using continuous-wave EPR methods for distances ranging from 5 to 20 Å in RNA's (34) and in a RNA–peptide complex (35), and very recently using pulse EPR techniques for distances between 20 and 72 Å in DNA (31) and RNA duplexes (30,32,33).

Work reported here evaluated distance measurements in DNA using nitroxides attached to phosphorothioates that are chemically substituted at specific backbone positions in arbitrary nucleic acid sequences (Figure 1A). This nitroxide, designated as R5, has two features that are unique as compared to nitroxide probes that are attached to either the base or the sugar of a certain chemically-modified nucleotide (mostly modified uridines) (8). First, R5 can be efficiently attached to arbitrary nucleotides within DNA sequences, as well as in RNA sequences where the nucleotide to the 5′ side of the intended labeling site can be substituted by a deoxyribonucleotide (36). Second, R5 labeling is cost-effective, as the cost of introducing a phosphorothioate is less than 1/20 of that of a modified nucleotide. These two features enable one to measure a large number of distances between arbitrary backbone positions of a DNA or RNA, thus facilitating global structure investigation.

The dynamics of R5 has been used to monitor RNA–RNA interaction in solution (36). However, R5 has not been evaluated for use in distance measurements. A particular concern is the potentially large distributions in the inter-nitroxide distance between a pair of R5 labels due to the inherent flexibility of the label itself (Figure 1A). In addition, R5 might be attached to either one of the two phosphorothioate diastereomers (R_p and S_p, Figure 1A), introduced at a ~50/50 ratio upon substituting a sulfur to one of the non-bridging oxygens (37). This might further broaden inter-spin distance distributions. The present study assessed R5 for distance measurements utilizing a dodecameric DNA duplex with a known

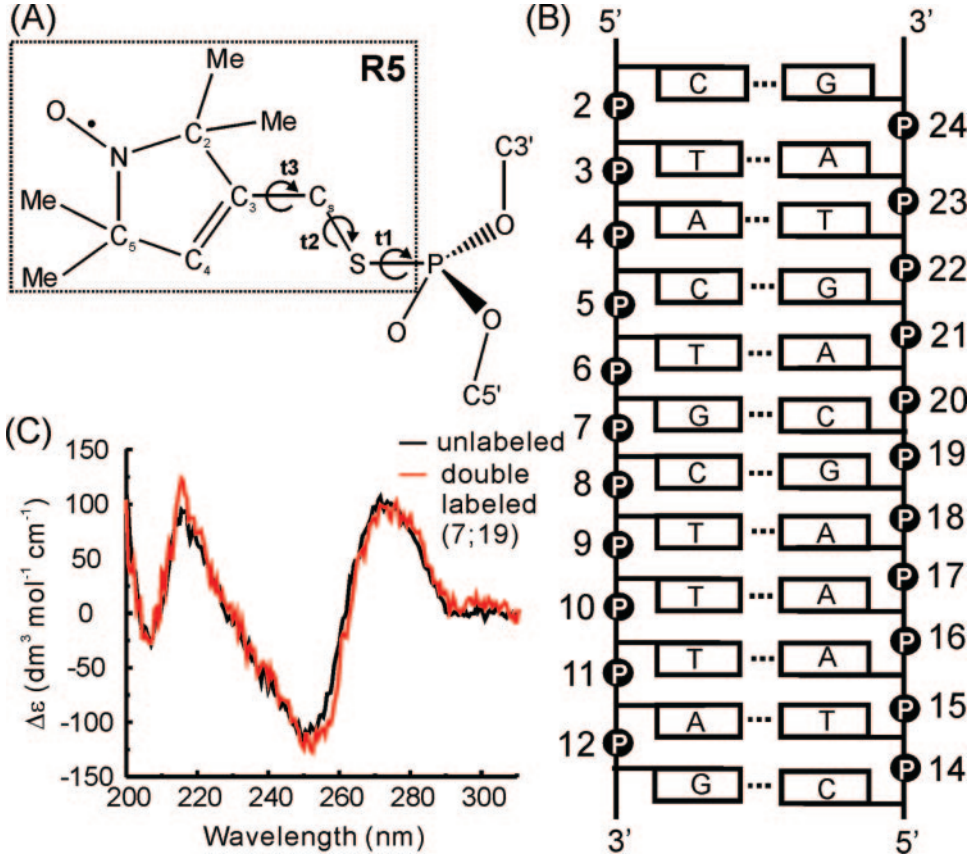

Figure 1. (**A**) The nitroxide attached to a phosphorothioate group in a nucleic acid. The R_p diastereomer is shown. (**B**) The model CS DNA. The phosphate backbone positions (represented by 'P') are numbered from 2 to 12 in the 5′ to 3′ direction in one strand and from 14 to 24 in the complementary strand. (**C**) A representative set of CD spectra of unlabeled and double-labeled CS DNA.

NMR solution structure. Using a 4-pulse double electron-electron resonance (DEER) method (18), distances ranging from 20 to 40 Å were measured between R5 pairs attached at selected DNA sites. The measured values were compared to inter-nitroxide distances predicted from modeling that accounts for the presence of both the R_p and the S_p diastereomers. The results demonstrate that R5 can be used to accurately measure distances in DNA. Furthermore, R5 was utilized to measure a distance in a 42 kD DNA, thus demonstrating its applicability in studying large nucleic acids. These studies firmly establish the utility of R5 for measuring distances in nucleic acids, thus expanding the capacity of SDSL in studying nucleic acids.

MATERIALS AND METHODS

DNA oligonucleotides

The sequence of the model DNA, designated as 'CS', is shown in Figure 1B. All DNA oligonucleotides, including those with a site-specific phosphorothioate modification, were synthesized using solid-phase DNA synthesis (Integrated DNA Technology, Coralville, IA). The individual DNA strands were purified using anion-exchange high-performance liquid chromatography (HPLC) following a reported procedure (36). HPLC fractions containing the DNA were desalted using a G-25 Sephadex column with water as the running medium. The purified DNA was lyophilized, then re-suspended in water and stored at −20°C. DNA concentrations were determined according to absorbance at 260 nm, using extinction coefficients of 108 200 $M^{-1}cm^{-1}$ and 125 800 $M^{-1}cm^{-1}$ for the respective DNA strands and their derivatives.

Spin labeling

A thiol-reactive nitroxide derivative, 3-iodomethyl-1-oxy-2,2,5,5-tetramethylpyrroline, was reacted with individual DNA strands that contain one phosphorothioate modification at a specific site. This reagent was prepared freshly from a precursor, 1-oxyl-2,2,5,5-tetramethyl-3-methane-sulfonyloxymethylpyrroline, following a published procedure (36,38,39) (Supplementary Data). In the DNA labeling reactions, up to 0.1 mM of phosphorothioate modified DNA was mixed with 60–100 mM of 3-iodomethyl-1-oxy-2,2,5,5-tetramethylpyrroline in a mixture containing 100 mM of MES [2-(N-Morpholino)-ethanesulfonic Acid (pH 5.8)] and 50% (v/v) formamide. After incubation in dark for 24 h at room temperature, excess nitroxide was removed using anion-exchange HPLC. The labeled DNA was desalted, lyophilized and stored as described above.

Characterizations of R5-labeled DNA

Thermal denaturation of DNA duplexes was carried out in a buffer containing 100 mM NaCl and 10 mM sodium phosphate (pH 6.8) following a previously reported procedure (13,40). The standard state free energy of duplex to single-strand transition ($\Delta G^0_{37°C}$) was obtained as described (13). Circular dichroism (CD) spectra were measured at room temperature between 310 and 200 nm on a JASCO spectropolarimeter. The duplexes were dissolved in the same buffer as those used for thermal denaturation measurements.

Four-pulse DEER EPR spectroscopy

To form a CS duplex for DEER measurement, appropriate singly labeled DNA strands were mixed in a 1:1 ratio, heated to 95°C for 3 min, then cooled at room temperature for 20 min. The heat-cool cycle was repeated, followed by the addition of appropriate amount of a buffer containing NaCl and Tris [2-Amino-2-(hydroxymethyl)- 1,3-propanediol (pH 7.5)]. The mixture was incubated at room temperature for >5 h and then lyophilized. The CS duplex was then re-suspended in 10 μl, with the final solution containing 100 mM NaCl, 50 mM Tris (pH 7.5), 20% (v/v) glycerol and 100–200 μM double-labeled DNA duplexes.

Four-pulse DEER measurements (18) were performed in a Bruker ELEXSYS 580 spectrometer fitted with a 2 mm split-ring resonator. Samples in 20% glycerol loaded into quartz capillaries were flash frozen and equilibrated at 50 K before measurements. The π/2 and π pulse widths were 8 and 16 ns, respectively. The observed frequency was set to the maximum of the low field region of the absorption spectrum and the pumping frequency was set to the maximum of the centerline. Each DEER data set was acquired with a repetition rate of either 500 or 1000 Hz, with the total acquisition time ranging from 5 to 14 h.

Echo decay data were analyzed using the DeerAnalysis2004 package (freely available at http://www.mpip-mainz.mpg.de/~jeschke/distance.html) (41). Data in the Bruker ELEXSYS format were directly loaded into the program and corrected for phase and zero time. Background echo decay was corrected using a homogeneous 3D spin distribution that gives an exponential decay signal. The distance distribution $P(r)$ was calculated based on the root mean square deviation (RMSD) between the experimental and the simulated data. $P(r)$ was first estimated using Hermite interpolation with different number of sampling points. The lowest number of Hermite interpolation sampling points from which there was no significant deviation of fitted from experimental data were selected. Final $P(r)$ was then obtained by Tikhonov regularization in distance domain with the constraint $P(r) > 0$ (41,42). The regularization parameter was adjusted to obtain the realistic resolution previously determined using Hermite interpolation. The mean distance and the standard deviation were calculated as:

$$\langle r_{\mathrm{DEER}} \rangle = \frac{\sum_{r_1}^{r_2} P(r) \cdot r}{\sum_{r_1}^{r_2} P(r)} \quad \mathbf{1}$$

$$\sigma = \sqrt{\frac{\sum_{r_1}^{r_2} P(r) \cdot (r - \langle r_{\mathrm{DEER}} \rangle)^2}{\sum_{r_1}^{r_2} P(r)}}, \quad \mathbf{2}$$

where r_1 and r_2 are the lower and upper boundary, respectively, of the selective population. The errors in the reported $\langle r_{\mathrm{DEER}} \rangle$ values were estimated to be 5% based on multiple measurements of the same sample.

Computational modeling

The input structure in the modeling program is the NMR-derived structures of the CS DNA duplex [(43); PDB ID 1CS2] or a generic B-DNA. The 1CS2 PDB data file contains three NMR-derived models: model 1 is the average of eight

structures obtained with JUMNA and the AMBER94 force field; model 2 is the average of eight structures obtained with JUMNA and FLEX and model 3 is the average of six structures obtained with X-PLOR (43). Only model 1 and model 2 in the PDB data file were used, since the coordinates for model 3 are inaccurate: the non-bridging oxygens in each nucleotide overlap with each other. The generic B-DNA duplex of the sequence d(CTACTGCTTTAG).d(CTAAAG-CAGTAG) was constructed using an in-house algorithm for nucleic acid construction, NASDAC (44).

Using NASDAC (ver.1.1), R5 was constructed at desired phosphorothioate sites on the DNA duplex. The geometry of R5 is as follows (Figure 1A): P-S, 1.99 Å (45); S-C_s, 1.83 Å, based on a phosphorylated cysteine in PDB ID 1H9C (46); Cs-C_3, 1.49 Å (47); C_3-C_2, 1.51 Å (48); C_2-Me, 1.52 Å (49); C_2-N, 1.49 Å (47); N-O, 1.27 Å (47); P-S-C_s, 101.1°, averaged over nine structures from PDB ID 1H9C (46); S-C_s-C_3, 109.4°; C_s-C_3-C_2, 124.9°; C_3-C_2-N, 99.4° (47); C_2-N-C_5, 114.4° (47); all other angles are of standard geometry. Torsion angle S-P-O-C5′ is set at the value in the original DNA duplex (with S substituted for O), torsion angles t1, t2 and t3 (Figure 1A) are variable, and the ring is planar. These geometry values give a ring-closing C_3-C_4 bond of 1.31 Å, in good agreement with experimental determination of this bond as 1.322 Å (50).

With the DNA coordinates fixed, the allowable conformational space for R5 attached to R_p and S_p phosphorothioate diastereomers was identified by variation of torsion angles t1, t2 and t3 in increments of 120° (torsions of 180°, 60° and −60°). Conformers in which any atom of the nitroxide and any atom of the DNA came within a contact distance were eliminated. The contact distance was defined as 75% of the sum of the van der Waals radii of any two interacting atoms. For a given DNA duplex with R5 modeled at two locations, distances were calculated between the nitroxide nitrogens. The corresponding mean and standard deviation of the distances were subsequently calculated based on the ensemble of distances. Calculations were carried out with inclusion of both R_p and S_p diastereomers at each site. For the NMR-derived structures, the ensemble of distances included all entries calculated for individual structures.

RESULTS AND DISCUSSION

A model DNA duplex for evaluating SDSL distance measurements

For evaluating SDSL distance measurements, an optimal model DNA duplex should contain non-self-complementary sequences with Watson–Crick base pairings, and have an available solution atomic-resolution structure (i.e. NMR structure) that is preferably a regular B-form helix. These criteria are satisfied by only a limited number of entries in the Protein Data Bank, and from these entries the model system is chosen to be a dodecamer DNA duplex designated as CS (following its PDB ID 1CS2, Figure 1B), the NMR structure of which revealed a slightly distorted B-form helix in solution (43).

Each CS DNA strand was labeled with one R5 at a desired site, and double-labeled duplexes were assembled from two singly labeled complementary strands. HPLC and denaturing gels showed that R5 labeling was efficient, and MALDI-TOF mass spectrometry analyses confirmed that each DNA strand has one, and only one, R5 attached (Supplementary Data). DNA duplex formation was confirmed on native gels (Supplementary Data). All studies reported here utilized phosphorothioate substituted DNA without separating the R_p and the S_p diastereomers.

Several lines of evidence indicated that R5 does not appreciably perturb the native CS DNA structure. First, CD spectroscopy yielded identical spectra between unlabeled and double-labeled CS duplexes (Figure 1C, Supplementary Data). All CD spectra showed a positive band centered at ~275 nm and a negative band around ~240 nm, which are characteristics of a B-form DNA (51). In addition, thermal melting measurements revealed that the differences in $\Delta G^0_{37°C}$ between unlabeled and double-labeled duplexes were <1.0 kcal/mol at all sites examined (Supplementary Data), again indicating a lack of significant perturbation due to R5. Finally, modeling studies showed that there is ample space to accommodate R5 without disrupting DNA base pairing and stacking (see Figure 3 for an example).

Distance measurements using DEER

The theory and implementation of the 4-pulse DEER experiment has been described in detail (18). The experiment measures the dipolar coupling between nitroxides as a modulation of the electron spin echo decay; the frequency of modulation is directly proportional to r^{-3}, where r is the inter-spin distance. For reliable distance determination, the echo decay must be observed for a time equal to at least one period of the modulation corresponding to the longest distance anticipated. In the present experiments, the echo decay was collected up to 2 μs, corresponding to a maximum measurable distance of 45 Å, which is sufficient to cover the distance range within the model dodecameric DNA.

Eight sets of double-labeled CS duplexes with predicted inter-nitroxide distances in the range of 20–40 Å were measured using DEER, and representative data are shown in Figure 2. Figure 2A shows the original data of the measured echo decay (black traces) and the corresponding exponentially decaying background due to spins with randomly distributed inter-spin distances (red traces). The differences between the background and the echo decay in each case clearly reveal oscillations with different periods (Figure 2B), which arise from dipolar interactions between spins at discrete distances. The distance distribution, $P(r)$, was calculated from the background corrected echo decay data using Tikhonov regularization (see Materials and Methods) (41,42). For each double-labeled CS duplex, the computed $P(r)$ reveals one major band of probable distances (indicated by shaded boxes, Figure 2C). A small number of minor bands, most of them at longer distances, are also observed (Figure 2C). The minor bands are artifacts in data fitting (52), and are excluded from further analysis.

The mean distance ($\langle r_{DEER} \rangle$) and the standard deviation (σ) were calculated for the major band of $P(r)$ (see Materials and Methods). The resulting values of $\langle r_{DEER} \rangle$ range from 21.2 to 38.8 Å (Table 1, Figure 2, Supplementary Data), and that of σ range from 1.7 to 5.0 Å (Table 2, Figure 2, Supplementary

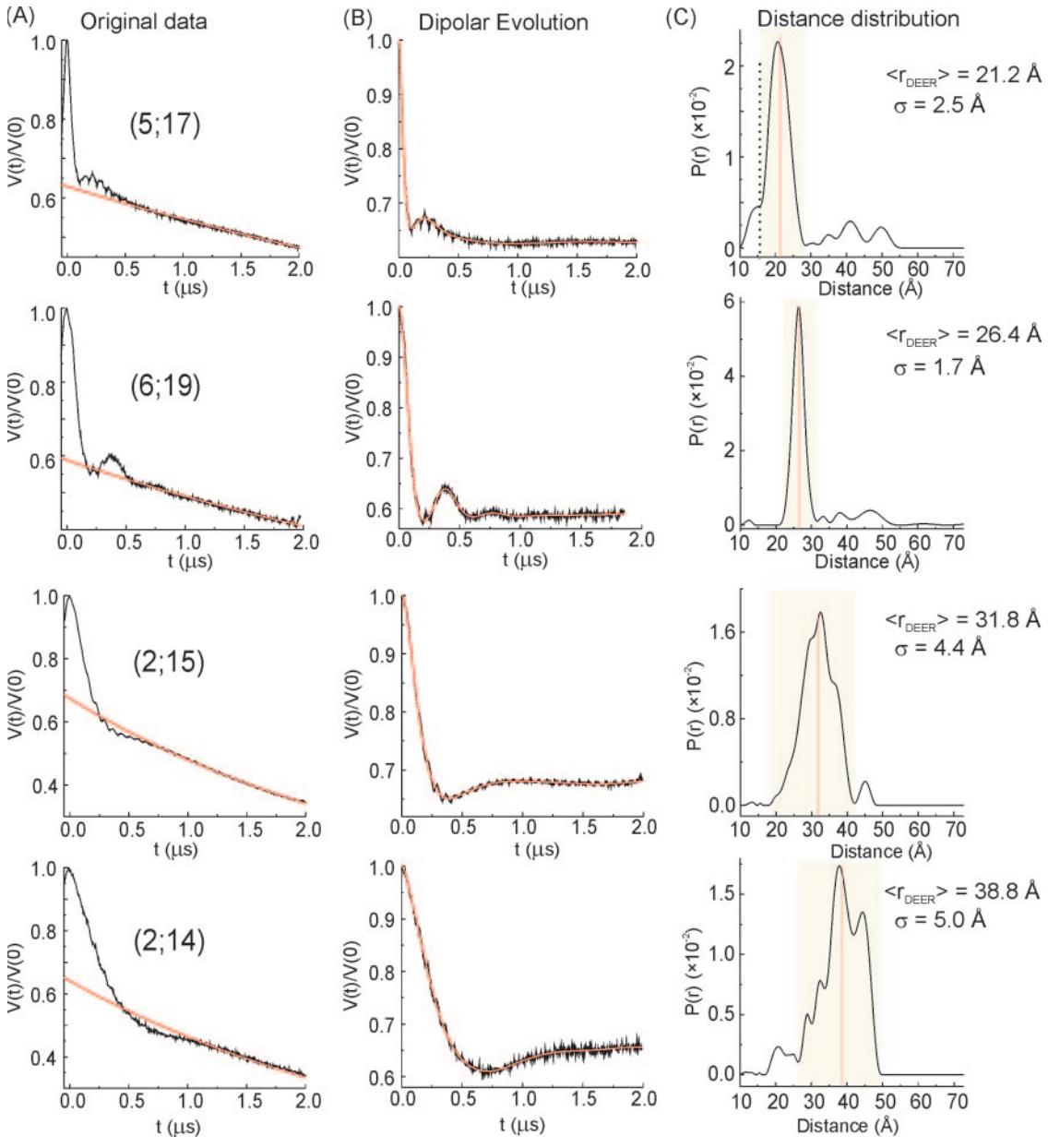

Figure 2. Representative DEER data for double-labeled CS duplexes, with the positions of R5 shown in the parenthesis. Additional DEER data are available in Supplementary Data. (**A**) Original echo decay data. The black traces are the measured echo amplitude that has been normalized to the amplitude at t = 0. The red traces are the background echo decay computed using a homogeneous 3D spin distribution. (**B**) Dipolar evolution functions. The black traces represent the differences between the measured echo decay and the background decay shown in (A). Electron Spin Echo Envelope Modulation (ESEEM) signals that manifest as high-frequency small amplitude oscillations for t <0.5 μs (54) are observed in the (5;17) and (2;15) data sets. The ESEEM signals do not affect data fitting. The red traces are the simulated echo decay computed according to the corresponding distance distributions shown in (C). (**C**) Distance distibutions $P(r)$ computed using Tikhonov regularization with a regularization parameter of 10 in the range from 1 to 8 nm. Changing the regulariztion parameter from 1 to 20 did not significantly alter $P(r)$. Shaded boxes indicate the the major bands in $P(r)$, and red lines mark the average distances calculated for each major band. The dotted line in the (5;17) dataset marks the lower limit of distances that are detectable by DEER with the reported experimental setup.

Data). The longest distance measured, $\langle r_{\mathrm{DEER}}(2;14)\rangle = 38.8$ Å, is between a pair of R5 at sites 2 and 14, which are located at the 5′ termini of the individual CS strands. The (2;14) distance distribution is also the broadest of all measured P(r) [$\sigma(2;14) = 5.0$ Å, Table 2]. This might reflect the higher flexibility at the terminus of the DNA duplex that broadens the inter-nitroxide distance distribution. It is also noted that most sets of $P(r)$ contain multiple peaks or shoulders in the major band, which might indicate the presence of multiple subsets of preferred R5 conformations.

Correlations between measured and expected distances

The measured $\langle r_{\mathrm{DEER}}\rangle$ values were compared to mean distances computed using an in-house modeling program. The program models R5 at a pair of selected nucleic acid sites

using experimentally determined bond lengths and bond angles (see Materials and Methods). With the nucleic acid coordinates fixed, allowable R5 conformations at each site are identified using a search based on exclusion of steric collisions, and an ensemble of inter-nitroxide distances is then calculated between all allowable conformations at a given site to that of the other site. An example of the CS DNA with R5 modeled at sites 2 and 14 is shown in Figure 3. This modeling program is efficient—each distance distribution can be computed in less than 10 s on a desktop PC. This study utilized $\langle r_{model} \rangle$, which is the predicted mean distance for the ensemble of R5 nitroxides modeled on both the R_p and the S_p phosphorothioate diastereomers, although results for R5 nitroxides attached to individual diastereomers (i.e. only R_p diastereomers) can be computed.

Table 1. Measured versus predicted average nitroxide–nitroxide distances

Nitroxide positions	Measured $\langle r_{DEER} \rangle$ (Å)	Modeled on NMR structure $\langle r_{model}^{NMR} \rangle$(Å)	Modeled on generic B-DNA $\langle r_{model}^{B\text{-}DNA} \rangle$(Å)
(2;14)	38.8	37.0	38.9
(2;15)	31.8	32.5	34.3
(3;15)	26.1	26.6	28.7
(4;15)	22.6	21.9	22.7
(5;17)	21.2	20.5	16.4
(6;18)	27.0	27.0	23.0
(6;19)	26.4	26.8	24.9
(7;19)	25.6	25.7	25.0
RMSD (Å)	—	0.8	2.6
R^2	—	0.98	0.88

Table 2. Measured versus predicted standard deviations in the nitroxide–nitroxide distance distributions

Nitroxide positions	Measured σ (Å)	Modeled on NMR structure σ_{model}^{NMR}(Å)	Modeled on generic B-DNA $\sigma_{model}^{B\text{-}DNA}$(Å)
(2;14)	5.0	3.9	3.1
(2;15)	4.4	3.4	2.7
(3;15)	3.6	3.2	2.5
(4;15)	3.9	3.4	2.3
(5;17)	2.5	5.4	5.4
(6;18)	2.1	3.9	4.6
(6;19)	1.7	2.8	3.3
(7;19)	1.9	1.5	1.9

Figure 4 shows the correlation between $\langle r_{DEER} \rangle$ and $\langle r_{model}^{NMR} \rangle$, which is the mean distance predicted based on the NMR structures of the CS DNA. The RMSD between $\langle r_{DEER} \rangle$ and $\langle r_{model}^{NMR} \rangle$ was calculated to be 0.8 Å, and the correlation coefficient (R^2) was 0.98 (Table 1, column 3). The data fit to $\langle r_{DEER} \rangle = 1.0 \times \langle r_{model}^{NMR} \rangle - 0.6$ Å, with a slope of unity and an offset that is small compared to the distances measured. This demonstrates that R5 can be used to accurately measure distances in DNA. Furthermore, the results indicate that distance distributions due to the presence of the R_p and the S_p diastereomers are adequately accounted for in the modeling, and it is not necessary to separate the two diastereomers in order to make meaningful distance measurements. Therefore, R5 can be used to measure distances in any DNA, even though in a majority of systems it is not practical to separate the R_p and the S_p phosphorothioate diastereomers.

The excellent correlation between $\langle r_{DEER} \rangle$ and $\langle r_{model}^{NMR} \rangle$ is gratifying given that the modeling program reported here uses a simple steric exclusion criterion for selecting allowable nitroxide conformations. This selection criterion is justified by continuous-wave EPR data. At all sites studied, continuous-wave EPR spectra of R5 showed characteristics of unrestricted, isotropic motion (Supplementary Data), suggesting the nitroxides have little interaction with the DNA.

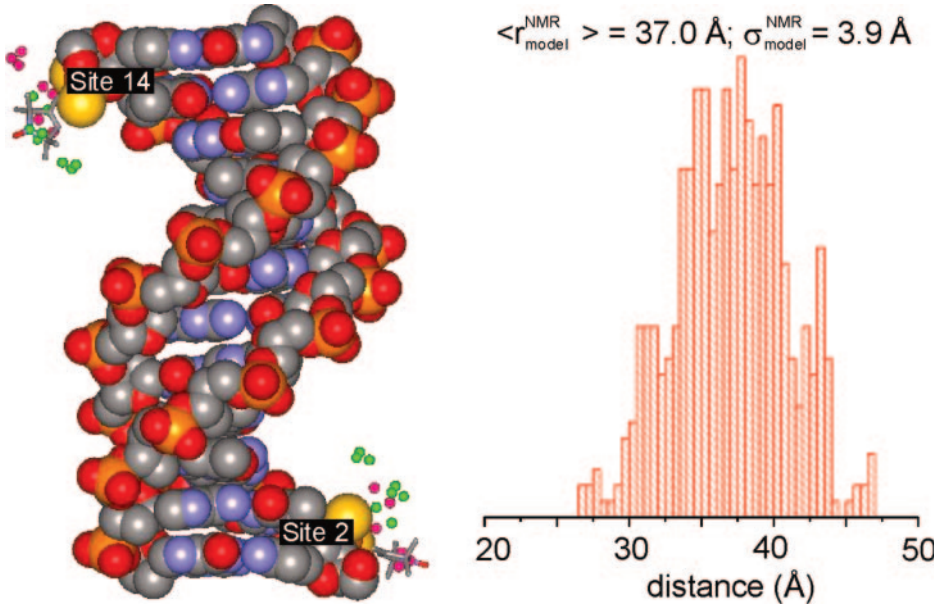

Figure 3. Ensembles of allowable R5 conformations modeled at sites 2 and 14 of the CS DNA. The NMR structure of the DNA (NMR model 2) is shown in CPK representation. One nitroxide in each ensemble is shown in stick representation. Green dots are the nitrogen atoms of nitroxides attached to the R_p phosphorothioate diastereomers, and pink dots are those attached to the S_p diastereomers. On the right is a histogram of the predicted distance distribution.

While the measured mean distances correlate very well with the predicted values, less satisfactory correlation is observed between the measured and the expected standard deviations, which report the widths of the distance distributions (Table 2). Among the eight sets of standard deviations, σ(3;15), σ(4;15) and σ(7;19) show matched measured and predicted values. The measured σ(5;17), σ(6;18) and σ(6;19) are significantly smaller than the predicted values. These larger predicted σ could be the result of an overestimation of allowable R5 conformations, as the reported modeling program gives the maximal allowable R5 conformation space. On the other hand, the measured σ(2;14) and σ(2;15) both exceed the corresponding predicted values by ~1.0 Å. The labeling sites in these two datasets are located at the 5′ termini, and the larger measured σ could reflect higher flexibility at the termini of the DNA duplex.

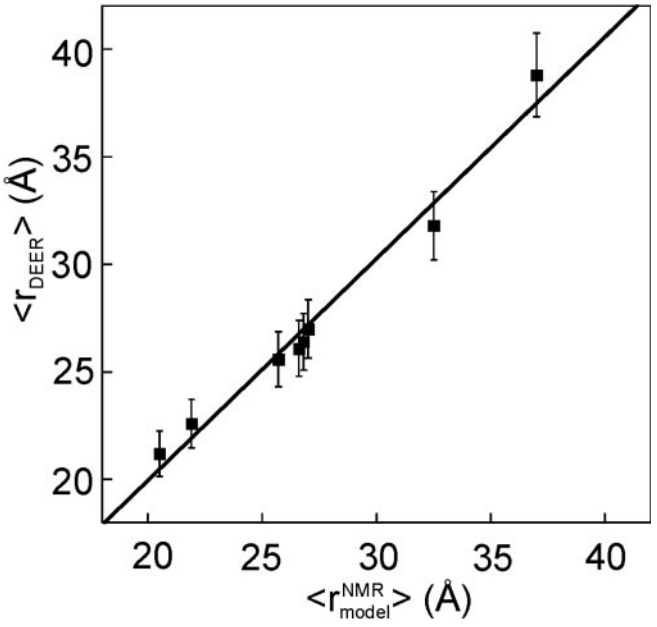

Figure 4. Correlation between $\langle r_{DEER}\rangle$ and $\langle r_{model}^{NMR}\rangle$. The error bars in $\langle r_{DEER}\rangle$ were set at 5% (see Materials and Methods). Error bars do not apply to $\langle r_{model}\rangle$ values, as they are calculated based on one set of DNA structure. The solid line represents a linear fit of $\langle r_{DEER}\rangle = 1.0 \times \langle r_{model}^{NMR}\rangle - 0.6$ Å.

Previous reports on calibrating SDSL distance measurements in nucleic acids correlated the measured distances to expected distances obtained from computer-generated structural models constructed using generic B-form (31) or A-form (34) parameters. In the present study, nitroxides were also modeled onto a computer-generated generic B-DNA with the CS DNA sequence (see Materials and Methods). This B-DNA is very similar to the average NMR structure, with an all atom root mean square distance of 2.3 Å. When the predicted $\langle r_{model}\rangle$ values based on the generic B-DNA ($\langle r_{model}^{B\text{-}DNA}\rangle$) were compared to the measured $\langle r_{DEER}\rangle$ values, an RMSD of 2.6 Å and an R^2 of 0.88 were obtained (Table 1, column 4). While the $\langle r_{DEER}\rangle/\langle r_{model}^{B\text{-}DNA}\rangle$ correlation is good, it is poorer than that between $\langle r_{DEER}\rangle$ and $\langle r_{model}^{NMR}\rangle$ (Table 1). Future studies will explore the feasibility of utilizing the $\langle r_{DEER}\rangle/\langle r_{model}\rangle$ correlations to discriminate DNA conformations. This may provide a diagnostic tool in structural analysis of DNA.

DEER measurement of a cross-strand distance in a 68 bp DNA

Previously reported SDSL distance measurements have been carried out on DNA or RNA duplexes ranging from 10 to 24 base-pair (bp) (6–15 kD) (30–34). With state-of-art nucleic acid chemical synthesis techniques, a phosphorothioate can be substituted at a specific site of a DNA that is over 100 nt long. This renders it feasible to carry out SDSL distance measurements in a large DNA. To illustrated this capability, studies were carried out in a 68 bp DNA duplex (41.9 kD, Figure 5A) that contains the origin of replication of the simian virus 40 (SV40), which serves as a system for studying eukaryotic DNA replications (53). Following similar protocols described above for CS DNA studies, the distance between a pair of R5 attached at sites directly across the origin DNA helix was measured (Figure 5). The average interspin distance was determined to be 25.2 Å, in very good agreement with the predicted value of 25.0 Å in a B-DNA

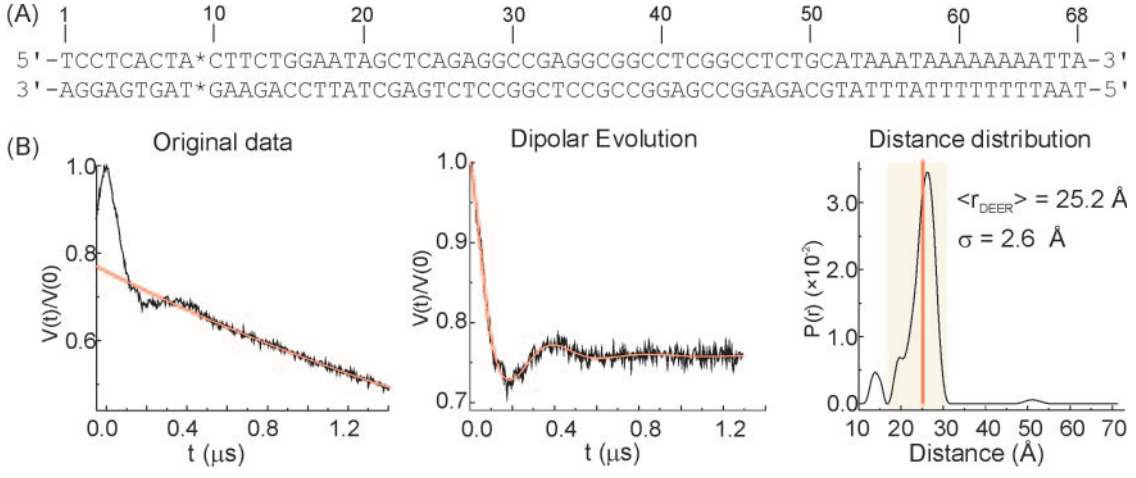

Figure 5. DEER measurement in a large DNA. (**A**) The sequence of a 68 bp duplex containing the entire DNA origin of replication in SV40, with '*' indicates the R5 labeling sites. (**B**) DEER data.

(i.e. $\langle r_{\text{model}}^{\text{BDNA}} \rangle$ between sites 7 and 19, Table 1). The data clearly demonstrate that R5 can be used to measure distances in a large DNA.

CONCLUSIONS AND PERSPECTIVE

The primary advantage of the R5 label is that it can be introduced, in an efficient and cost-effective manner, at any sequence position in a nucleic acid. A potential problem in using R5 in structure determination through nitroxide–nitroxide distance measurement is that a distribution of distances is generated due to the inherent flexibility of the label as well as the existence of two diastereomers in the phosphorothioate linkage, which complicate interpretation of nitroxide–nitroxide distances in terms of nucleic acid structure. However, the results of this study suggest that the distribution can be modeled sufficiently well to enable measured distances to be correlated to the structure of the parent molecule. Although the studies presented utilized duplex DNAs, the methods described are applicable to both duplex and non-duplex DNAs. The results thus lay a foundation for the application of SDSL for the unrestricted global structure mapping in nucleic acids and protein–nucleic acid complexes.

SUPPLEMENTARY DATA

Supplementary data are available at NAR online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr C. Altenbach for stimulating discussions on DEER measurements, Dr R. Langen and Dr B. Hegde for help with CD spectroscopy and H. He for help on DNA structural analyses. P.Z.Q. was supported by the Petroleum Research Fund # 39623-G4, the American Cancer Society IRG-58-007-45, the NSF CAREER award and a startup fund from the University of Southern California. W.L.H., A.K.K. and N.V.E. were supported by NIH Grant EY05216, the Jules Stein Professor Endowment, and a grant from the Ford Bundy and Anne Smith Bundy Foundation. K.H. was supported by a grant from Hungarian National Research Fund (OTKA T48334). The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Altenbach,C., Flitsch,S.L., Khorana,H.G. and Hubbell,W.L. (1989) Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry*, **28**, 7806–7812.
2. Hubbell,W.L. and Altenbach,C. (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling. *Curr. Opin. Struct. Biol.*, **4**, 566–573.
3. Feix,J.B. and Klug,C.S. (1998) Site-directed spin labeling of membrane proteins and peptide-membrane interactions. In Berliner,L.J. (ed.), *Biological Magnetic Resonance*. Plenum Press, NY, pp. 251–281.
4. Hubbell,W.L., Cafiso,D.S. and Altenbach,C. (2000) Identifying conformational chamges with site-directed spin labeling. *Nature Struct. Biol.*, **7**, 735–739.
5. Fajer,P.G. (2000) Electron spin resonance spectroscopy labeling in proteins and peptides analysis. In Meyers,R. (ed.), *Encyclopedia of Analytical Chemistry*. John Wiley & Sons, Chichester, pp. 5725–5761.
6. Columbus,L. and Hubbell,W.L. (2002) A new spin on protein dynamics. *Trends Biochem. Sci.*, **27**, 288–295.
7. Hubbell,W.L., Altenbach,C., Hubbell,C.M. and Khorana,H.G. (2003) Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. *Adv. Protein Chem.*, **63**, 243–290.
8. Qin,P.Z. and Dieckmann,T. (2004) Application of NMR and EPR methods to the study of RNA. *Curr. Opin. Struct. Biol.*, **14**, 350–359.
9. Edwards,T.E., Okonogi,T.M. and Sigurdsson,S.T. (2002) Investigation of RNA–protein and RNA–metal ion interactions by electron paramagnetic resonance spectroscopy: the HIV TAR-Tat motif. *Chem. Biol.*, **9**, 699–706.
10. Qin,P.Z., Feigon,J. and Hubbell,W.L. (2005) Site-directed spin labeling studies reveal solution conformational changes in a GAAA tetraloop receptor upon Mg2+-dependent docking of a GAAA tetraloop. *J. Mol. Biol.*, **351**, 1–8.
11. Kim,N.K., Murali,A. and DeRose,V.J. (2005) Separate metal requirements for loop interactions and catalysis in the extended hammerhead ribozyme. *J. Am. Chem. Soc.*, **127**, 14134–14135.
12. Edwards,T.E. and Sigurdsson,S.T. (2005) EPR spectroscopic analysis of U7 hammerhead ribozyme dynamics during metal ion induced folding. *Biochemistry*, **44**, 12870–12878.
13. Qin,P.Z., Jennifer,I. and Oki,A. (2006) A model system for investigating lineshape/structure correlations in RNA site-directed spin labeling. *Biochem. Biophys. Res. Commun.*, **343**, 117–124.
14. Borbat,P.P., Costa-Filho,A.J., Earle,K.A., Moscicki,J.K. and Freed,J.H. (2001) Electron spin resonance in studies of membranes and proteins. *Science*, **291**, 266–269.
15. Lakshmi,K.V. and Brudvig,G.W. (2001) Pulsed electron paramagnetic resonance methods for macromolecular structure determination. *Curr. Opin. Struct. Biol.*, **11**, 523–531.
16. Steinhoff,H.J. (2004) Inter- and intra-molecular distances determined by EPR spectroscopy and site-directed spin labeling reveal protein-protein and protein-oligonucleotide interaction. *Biol. Chem.*, **385**, 913–920.
17. Jeschke,G. (2005) EPR techniques for studying radical enzymes. *Biochim. Biophys. Acta.*, **1707**, 91–102.
18. Pannier,M., Veit,S., Godt,A., Jeschke,G. and Spiess,H.W. (2000) Dead-time free measurement of dipole–dipole interactions between electron spins. *J. Magn. Res.*, **142**, 331–340.
19. Milov,A.D., Tsvetkov,Y.D., Gorbunova,E.Y., Mustaeva,L.G., Ovchinnikova,T.V. and Raap,J. (2002) Self-aggregation properties of spin-labeled zervamicin IIA as studied by PELDOR spectroscopy. *Biopolymers*, **64**, 328–336.
20. Hinderberger,D., Schmelz,O., Rehahn,M. and Jeschke,G. (2004) Electrostatic site attachment of divalent counterions to rodlike ruthenium(II) coordination polymers characterized by EPR spectroscopy. *Angew. Chem. Int. Ed. Engl.*, **43**, 4616–4621.
21. Jeschke,G., Bender,A., Paulsen,H., Zimmermann,H. and Godt,A. (2004) Sensitivity enhancement in pulse EPR distance measurements. *J. Mag. Res.*, **169**, 1–12.
22. Pornsuwan,S., Bird,G., Schafmeister,C.E. and Saxena,S. (2006) Flexibility and lengths of bis-peptide nanostructures by electron spin resonance. *J. Am. Chem. Soc.*, **128**, 3876–3877.
23. Persson,M., Harbridge,J.R., Hammarstrom,P., Mitri,R., Martensson,L.G., Carlsson,U., Eaton,G.R. and Eaton,S.S. (2001) Comparison of electron paramagnetic resonance methods to determine distances between spin labels on human carbonic anhydrase II. *Biophys. J.*, **80**, 2886–2897.
24. Borbat,P.P., McHaourab,H.S. and Freed,J.H. (2002) Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme. *J. Am. Chem. Soc.*, **124**, 5304–5314.
25. Jeschke,G., Wegener,C., Nietschke,M., Jung,H. and Steinhoff,H.J. (2004) Interresidual distance determination by four-pulse double electron-electron resonance in an integral membrane protein: the Na+/proline transporter PutP of *Escherichia coli*. *Biophys., J.*, **86**, 2551–2557.
26. Sale,K., Song,L., Liu,Y.S., Perozo,E. and Fajer,P. (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.*, **127**, 9334–9335.
27. Hilger,D., Jung,H., Padan,E., Wegener,C., Vogel,K.P., Steinhoff,H.J. and Jeschke,G. (2005) Assessing oligomerization of membrane proteins

by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of *E. coli*. *Biophys. J*., **89**, 1328–1338.
28. Jeschke,G., Bender,A., Schweikardt,T., Panek,G., Decker,H. and Paulsen,H. (2005) Localization of the N-terminal domain in light-harvesting chlorophyll a/b protein by EPR measurements. *J. Biol. Chem*., **280**, 18623–18630.
29. Banham,J.E., Timmel,C.R., Abbott,R.J., Lea,S.M. and Jeschke,G. (2006) The characterization of weak protein–protein interactions: evidence from DEER for the trimerization of a von Willebrand Factor A domain in solution. *Angew. Chem. Int. Ed. Engl*., **45**, 1058–1061.
30. Schiemann,O., Weber,A., Edwards,T.E., Prisner,T.F. and Sigurdsson,S.T. (2003) Nanometer distance measurements on RNA using PELDOR. *J. Am. Chem. Soc*., **125**, 3334–3335.
31. Schiemann,O., Piton,N., Mu,Y., Stock,G., Engels,J.W. and Prisner,T.F. (2004) A PELDOR based nanometer distance ruler for oligonucleotides. *J. Am. Chem. Soc*., **126**, 5722–5729.
32. Borbat,P.P., Davis,J.H., Butcher,S.E. and Freed,J.H. (2004) Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes. *J. Am. Chem. Soc*., **126**, 7746–7747.
33. Piton,N., Schiemann,O., Mu,Y., Stock,G., Prisner,T. and Engels,J.W. (2005) Synthesis of spin-labeled RNAs for long range distance measurements by peldor. *Nucleosides Nucleotides Nucleic Acids*, **24**, 771–775.
34. Kim,N., Murali,A. and DeRose,V.J. (2004) A distance ruler for RNA using EPR and site-directed spin labeling. *Chem. Biol*., **11**, 939–948.
35. Macosko,J.C., Pio,M.S., Tinoco,I., Jr and Shin,Y.-K. (1999) A novel 5′ displacement spin-labeling technique for electron paramagnetic resonance spectroscopy of RNA. *RNA*, **5**, 1158–1166.
36. Qin,P.Z., Butcher,S.E., Feigon,J. and Hubbell,W.L. (2001) Quantitative analysis of the GAAA tetraloop/receptor interaction in solution: a site-directed spin labeling study. *Biochemistry*, **40**, 6929–6936.
37. Eckstein,F. (1985) Nucleoside phosphorothioates. *Ann. Rev. Biochem*., **54**, 367–402.
38. Hankovszky,H.O., Hideg,K. and Lex,L. (1980) Nitroxyls; VIII. Synthesis and Reactions of Highly Reactive 1-Oxyl-2,2,5,5-tetramethyl-2,5-dihydropyrrole-3-ylmethyl Sulfonates. *Synthesis*, 914–916.
39. Mchaourab,H.S., Kalai,T., Hideg,K. and Hubbell,W.L. (1999) Motion of spin-labeled side chains in T4 lysozyme: effect of side chain structure. *Biochemistry*, **38**, 2947–2955.
40. Qin,P.Z., Hideg,K., Feigon,J. and Hubbell,W.L. (2003) Monitoring RNA base structure and dynamics using site-directed spin labeling. *Biochemistry*, **42**, 6772–6783.
41. Jeschke,G., Koch,A., Jonas,U. and Godt,A. (2002) Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson*., **155**, 72–82.
42. Chiang,Y.W., Borbat,P.P. and Freed,J.H. (2005) The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson*., **172**, 279–295.
43. Leporc,S., Mauffret,O., Tevanian,G., Lescot,E., Monnot,M. and Fermandjian,S. (1999) An NMR and molecular modeling analysis of d(CTACTGCTTTAG). d(CTAAAGCAGTAG) reveals that the particular behavior of TpA steps is related to edge-to-edge contacts of their base-pairs in the major groove. *Nucleic Acids Res*., **27**, 4759–4767.
44. Chambers,E.J., Price,E.A., Bayramyan,M.Z. and Haworth,I.S. (2003) Computation of DNA backbone conformations. *J. Biomol. Struct. Dyn*., **306**, 177–185.
45. Goldstein,B.M. (1982) Disorder in the structure of trisodium phosphorothioate dodecahydrate. *Acta Cryst*., **B38**, 1116–1120.
46. Ab,E., Schuurman-Wolters,G.K., Nijlant,D., Dijkstra,K., Saier,M.H., Robillard,G.T. and Scheek,R.M. (2001) NMR structure of cysteinyl-phosphorylated enzyme IIb of the N,N′-diacetylchitobiose-specific phosphoenolpyruvate-dependent phosphotransferase system of *Escherichia coli*. *J. Mol. Biol*., **308**, 993–1009.
47. Langen,R., Oh,K.J., Cascio,D. and Hubbell,W.L. (2000) Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry*, **39**, 8396–8405.
48. Aarset,K., Page,E.M. and Rice,D.A. (2005) Molecular structure of 2,5-dihydropyrrole (C4NH7), obtained by gas-phase electron diffraction and theoretical calculations. *J. Phys. Chem*., **A109**, 4961–4965.
49. Cygler,M. (1981) Structure of 2,2,5,5-tetramethyl-3-phenylethynyl-3-pyrrolidinol. *Acta Cryst*., **B37**, 1765–1767.
50. Chen,S.Y., Nie,J.J., You,J.Z., Xu,D.J. and Chen,Y.Z. (2001) Synthesis and crystal structure of a pyrroline nitroxide radical, (N-2,2,5,5-tetramethyl-1-oxo-3-pyrroline-3-carbonyl) (N,N′-diisopropyl)urea *J. Chem. Crystallogr*. **31**, 339–343.
51. Rodger,A. and Norden,B. (1997) In *Circular Dichroism and Linear Dichroism*. Oxford University Press, Oxford, NY, pp. 23–31.
52. Jeschke,G., Panek,G., Godt,A., Bender,A. and Paulsen,H. (2004) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson*., **26**, 223–244.
53. Deb,S., DeLucia,A.L., Baur,C.P., Koff,A. and Tegtmeyer,P. (1986) Domain structure of the simian virus 40 core origin of replication. *Mol. Cell. Biol*., **6**, 1663–1670.
54. Schweiger,A. and Jeschke,G. (2001) Nuclear modulation effect I: basic experiments. In *Principles of Pulse Electron Paramagnetic Resonance*. Oxford University Press, Oxford, pp. 247–295.