

A Practical Guide for Solid-State NMR Distance Measurements in Proteins

FRANK A. KOVACS, DANIEL J. FOWLER, GREGORY J. GALLAGHER, LYNMARIE K. THOMPSON

Department of Chemistry, University of Massachusetts Amherst, Amherst, Massachusetts

ABSTRACT: Rotational resonance (R^2) and rotational echo double resonance (REDOR) are powerful solid-state NMR techniques that can be applied in a site-directed fashion for precise distance measurements in proteins. These tools are well suited for systems in which a few precise distance measurements are needed to understand a mechanism or map a binding site, particularly if this information is unavailable from x-ray crystallography or solution NMR, as is often the case for membrane proteins. Strategies and challenges in the design and implementation of such experiments are described and illustrated with experiments probing mechanisms of transmembrane signaling in bacterial chemotaxis receptors. © 2007 Wiley Periodicals, Inc. Concepts Magn Reson Part A 30A: 21–39, 2007

KEY WORDS: rotational resonance; R^2 ; REDOR; MAS; solid-state NMR; site-directed distance measurements

INTRODUCTION

Why Use Solid-State NMR to Study Proteins?

Solid-state NMR has become an important tool for structural studies of proteins. X-ray crystallography and solution NMR are in many ways more facile, and will continue to be the principal tools of struc-

tural biologists. The many structures obtained by crystallography and solution NMR have provided key insights into mechanisms of protein functions and the foundation for rational drug design. However, structure determination of membrane proteins has lagged behind that of soluble proteins because these techniques are not as useful for insoluble systems. X-ray crystallography requires high-quality crystals, which have proven difficult to obtain for many systems, including membrane proteins. Solution NMR techniques, on the other hand, require that the protein undergo rapid isotropic motion; because this is not the case for large proteins (>50 kDa) or for proteins associated with lipid bilayers, membrane proteins are often not amenable to complete structure determination by solution NMR. There has been a recent acceleration in the rate of solving crystal structures of membrane proteins (1), as well as advances in solution NMR methods applicable to structure determination of membrane proteins (2, 3).

Received 27 July 2006; accepted 28 August 2006

Correspondence to: Lynmarie Thompson; E-mail: thompson@chem.umass.edu

Current address for F.A.K.: Department of Chemistry, University of Nebraska at Kearney, Kearney, NE 68849-1150.

Current address for G.J.G.: Department of Chemistry, American International College, 1000 State Street, Springfield, MA 01075.

Concepts in Magnetic Resonance Part A, Vol. 30A(1) 21–39 (2007)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/cmr.a.20071

© 2007 Wiley Periodicals, Inc.

However, membrane protein structures are still less than 1% of the structures in the Protein Data Bank, which is much less than the predicted 20%–30% occurrence of membrane proteins in biological organisms.

Solid-state NMR provides excellent tools for such systems. Recent advances in solid-state NMR techniques are making it possible to determine complete structures of small proteins (4). Furthermore, the ability of solid-state NMR to measure selected distances and local structure in large proteins (5) can provide answers to specific, detailed questions about protein function in systems inaccessible to either crystallography or solution NMR. These targeted distance measurements within large proteins, to address questions of biological interest, are the focus of this article. Here, we describe how to identify cases where such an approach may be productive, how to target a measurement to the desired region, how to choose the appropriate spectroscopic measurement technique, and how to perform the measurements with sufficient accuracy and precision. Throughout the article, these points are illustrated with reference to our work on the *E. coli* serine chemotaxis receptor.

Isotopically labeled solid-state NMR samples must be prepared biosynthetically for proteins larger than ≈ 50 amino acids. Uniform and selective/extensive labeling yield spectra that are generally too complex for site-directed studies of large proteins (unless techniques are developed to focus on the site of interest). Thus a critical aspect of site-directed experiments is to develop a strategy for directing the isotopic labels to the site of interest. One approach has been the labeling of a small molecule that will bind to the protein, such as a ligand (6) or chromophore (7, 8). A more general approach discussed below is the use of site-directed mutagenesis for the incorporation of a single-labeled amino acid into the protein (9).

Solid-state NMR studies of proteins employ either oriented samples (not covered here, see (10) for a recent review) or magic angle spinning (MAS). In MAS experiments the sample is spun about an axis that is at the “magic angle” (54.7°) with respect to the magnetic field so that the chemical shift anisotropy and dipolar couplings that are less than the spinning speed will be averaged to zero. Because MAS averages dipolar couplings, distance measurements require “recoupling” techniques that reintroduce the dipolar coupling to be measured. There are two main types of MAS solid-state NMR distance measurement techniques: those used to measure homonuclear dipolar couplings between spin $\frac{1}{2}$ pairs such as ^{13}C , ^{13}C and those used to measure heteronuclear dipolar couplings between spin $\frac{1}{2}$ pairs such as

^{13}C , ^{15}N and ^{13}C , ^{19}F . For measurement of selected distances in large proteins, the NMR resonances are often extremely weak as a result of the small number of spins in the sample (e.g., for measurement of a distance between a single pair of carbons in a 60-kDa protein imbedded in a membrane vesicle). Thus the techniques employed must not sacrifice much of the signal and must be stable enough for the long acquisitions needed to increase the signal-to-noise ratio. This article focuses on two such techniques that have been widely used: rotational-echo double-resonance (REDOR) for heteronuclear distance measurements and rotational resonance (R^2) for homonuclear distance measurements (Fig. 1). Rotational-echo double-resonance (REDOR) uses rotor synchronized pi pulses to recover the dipolar interaction between the two nuclei (11, 12). R^2 recovers the dipolar coupling by choosing the MAS spinning frequency such that the isotropic chemical shift difference between the two nuclei of interest ($\Delta\sigma$) is equal to an integer multiple of the MAS frequency (ω_r): $\Delta\sigma = n\omega_r$ (13). These techniques have been applied to a number of proteins to measure selected distances for the development of structural models, characterization of ligand sites, and testing of proposed mechanisms (14–18). This article describes how to design and implement REDOR and R^2 distance measurements that address a biological question of interest in a large protein.

A Membrane Protein System for Investigating the Mechanism of Transmembrane Signaling with Solid-State NMR

Here, we use an experimental design example to demonstrate the practical use of the distance measurement techniques R^2 and REDOR to address an important biological question. We begin by briefly describing the system and the questions of interest to be studied with distance measurements. The serine receptor (Fig. 2) from *E. coli* is a member of a class of transmembrane receptors known as chemotaxis receptors. This receptor has a serine binding site that gives the bacterium the ability to sense the concentration of the amino acid serine in its environment and then bias its swimming in the direction of an increase in serine concentration. The 120-kDa dimeric receptor has a periplasmic domain that binds the ligand serine, a transmembrane domain that spans the inner bacterial membrane, and a cytoplasmic domain that controls the phosphorylation activity of an associated protein. For this well-studied chemotaxis receptor family, there are high-resolution crystal

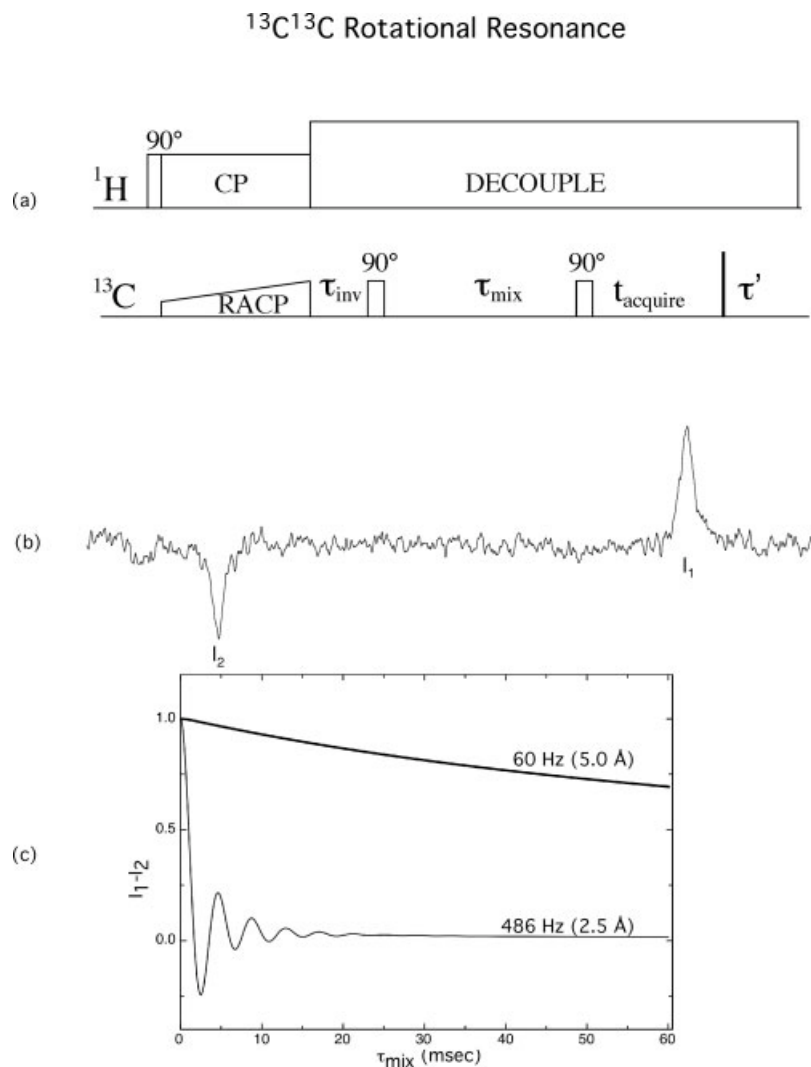


Figure 1 Overview of $^{13}\text{C}^{13}\text{C}$ R^2 (above) and $^{13}\text{C}\{^{19}\text{F}\}$ REDOR (next page) experiments. (a) Pulse sequence for constant-time version of rotational resonance first uses a delay inversion (τ_{inv}) to invert one of the spins, and then allows for the spins to exchange magnetization during a variable mixing time (τ_{mix}). The postacquisition decoupling time τ' is varied such that $\tau_{\text{mix}} + \tau' = \text{constant}$. (b) A resulting ^{13}C difference spectrum (0–60 ms mixing time) for the 5 Å $^{13}\text{C}^{13}\text{C}$ R^2 distance measurement in the Ser receptor shown in Fig. 2. The peak intensities of I_1 and I_2 are subtracted and the difference magnetization is plotted as a function of mixing time and compared to simulated magnetization exchange curves to determine the distances. (c) Exchange curves for both strong and weak $^{13}\text{C}^{13}\text{C}$ dipolar couplings (2.5 Å and 5.0 Å distances, respectively). (d) Pulse sequences for $^{13}\text{C}\{^{19}\text{F}\}$ REDOR (^{13}C -observed, ^{19}F -dephase). Single-refocusing pulse REDOR (option 1) is compared with the multiple refocusing pulse version (often called xy -8 because xy -8 phase cycling is used on both channels) of REDOR (option 2) in which the pi pulses on the rotor period (black) are all moved to the observe ^{13}C channel. Both versions are shown for four rotor cycles of dephasing (gray band on ^{13}C channel). The single refocusing pulse experiment (option 1) includes two central rotor periods without dephasing so that a multiple of eight dephasing pulses are used (for the optimum xy -8 phase cycling) (39). The multiple refocusing pulse version (option 2) has a total of four pi pulses on each channel that are xy -4 phase cycled. Both versions can be conducted to measure the full dipolar coupling (black pulses on the rotor periods plus white pulses on the half rotor period) or a scaled coupling (black pulses on the rotor periods plus gray pulses shifted away from the half rotor period as indicated by the arrows). (e) ^{13}C spectra showing a REDOR difference (ΔS) between the control spectrum (S_0) collected with no ^{19}F pi pulses and the dephased spectrum (S) collected with ^{19}F pi pulses on F-gramine. The $\Delta S/S_0$ data are plotted as a function of dephasing time ($= N_c T_r = \text{number of rotor periods} \times \text{rotor period}$) and simulated to obtain the distance. (f) REDOR dephasing curves for strong and weak $^{13}\text{C}^{19}\text{F}$ dipolar couplings (2.5 Å and 5.0 Å distances, respectively). The plot also shows that the scaled-coupled REDOR experiment makes a strong coupling look like a weak one, which is useful for optimization of the experimental parameters: the scaled-coupled curve for 2.5 Å with shift = 0.9 approaches the full coupling curve for 5.0 Å. Shift of 0.9 indicates pulses have been shifted from $0.5T_r$ to $0.9T_r$.

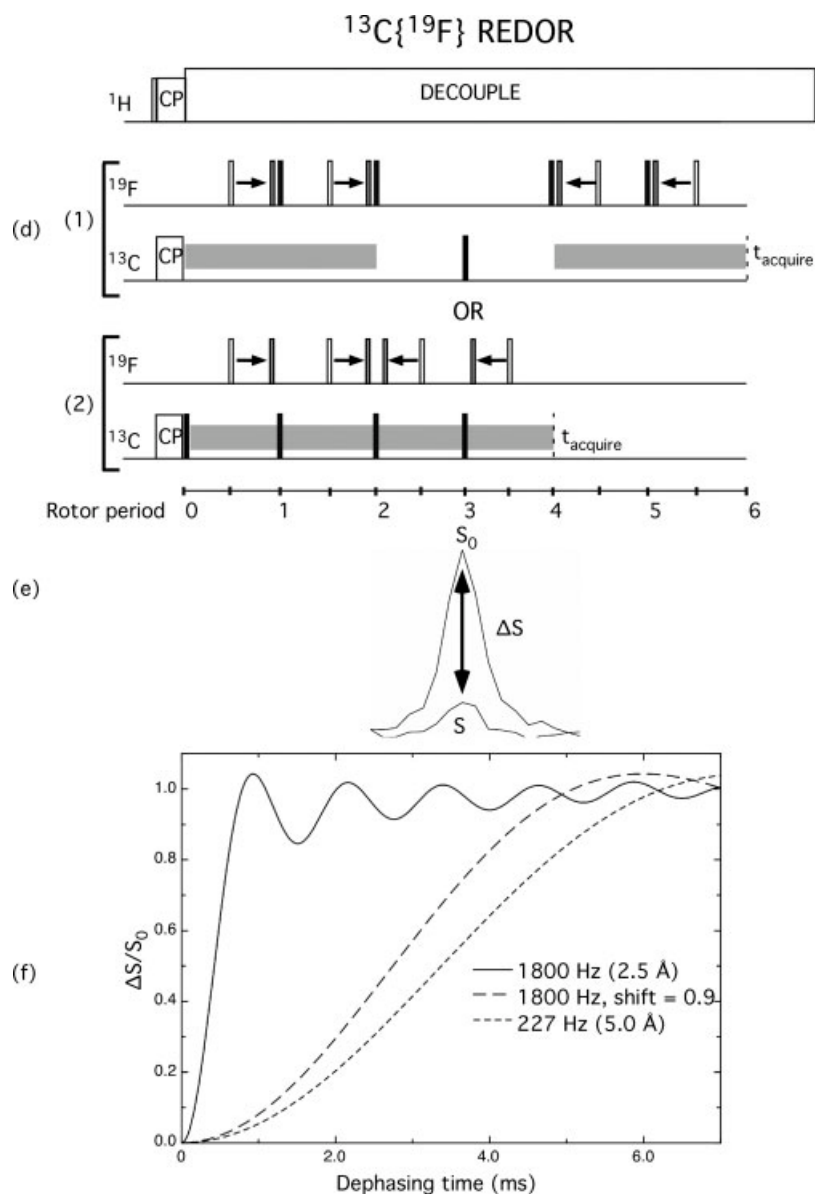


Figure 1 (Continued)

structures of protein fragments of the periplasmic and cytoplasmic domains but only a low-resolution model of the transmembrane domain obtained from cysteine cross-linking experiments. Though there are no structures determined for an entire receptor, a whole receptor structural model has been constructed using the various domain structures (19), and a generalized model is shown in Fig. 2.

A primary objective of the solid-state NMR structural studies of this receptor has been to determine what structural changes propagate information across the bacterial inner membrane. From crystal structures of the periplasmic domain with and without ligand bound, researchers proposed that ligand binding indu-

ces a scissors-like conformational change, where one of the periplasmic domains of the functional dimer rotates with respect to the other (20). However, when a different method of superimposing the crystal structures was used it was discovered that in the ligand-bound structure one of the periplasmic helices appears to have moved downward toward the membrane by approximately 1.6 Å in a piston-like motion (21). Our lab has endeavored to address this question using solid-state NMR distance measurements on the intact, membrane-bound receptor. We used the crystal structure of the periplasmic domain to select measurable distances between the α_4 helix proposed to move downward (α_4) and the other helices of the four-helix bundle

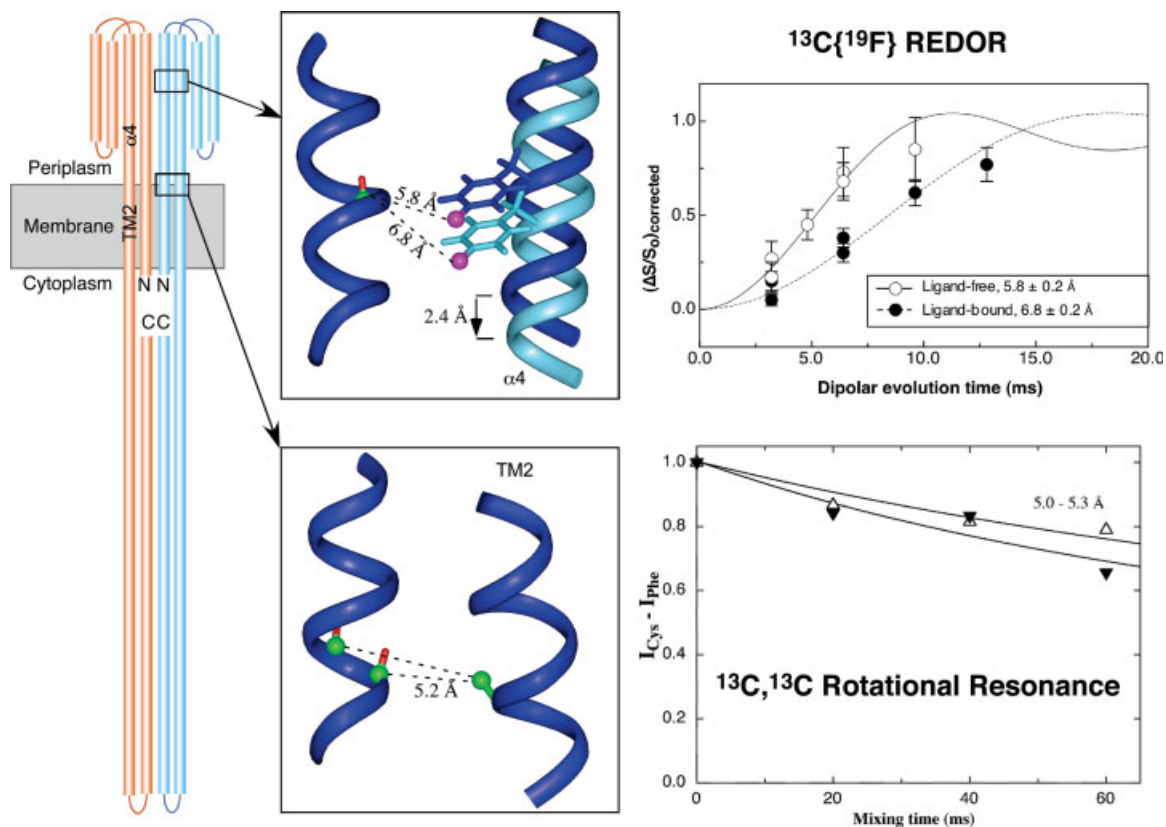


Figure 2 Site-directed distance measurements performed on the intact, membrane-bound Ser receptor. The piston model (21) proposes that ligand binding displaces helix $\alpha 4$ -TM2 toward the cytoplasm by about 1.6 Å. A ligand-induced change in interhelical distance is observed in the periplasmic domain (9) but is not detected in the transmembrane domain (15) (open symbols are ligand-free receptors; solid symbols are ligand-bound receptors). The experiments were directed to the sites of interest by using site-directed mutagenesis to introduce a unique Cys residue (the wild-type protein has no Cys residues). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

(α_1 , α_2 , and α_3). If the predicted piston-like movement of 1.6 Å occurs upon the binding of ligand, we should be able to measure this as a change in the interhelical distance and to perform additional measurements to follow the propagation of any conformational change into the transmembrane and cytoplasmic domains. Thus solid-state NMR distance measurements can provide insights into the mechanisms of transmembrane signaling by receptors.

EXPERIMENTAL DESIGN

Can Solid-State NMR Distance Measurements Address the Biological Question of Interest?

Distances can be measured by other more sensitive techniques. When is solid-state NMR the appropriate

choice? Solid-state NMR has the potential to provide fairly precise distance measurements, in part because distances can be measured between native sites and thus avoid the need to incorporate a bulky and dynamic spin label or fluorophore. When the biological question of interest requires precise short distance measurements, solid-state NMR may be the method of choice. For example, measurement of the subtle ligand-induced conformational changes proposed to transmit the signal in the serine receptor (e.g., 1.6 Å piston) requires such precise distance measurements. Furthermore, solid-state NMR distance measurements are typically impractical unless some structural information is already known. Because of the limited range of measurable distances (Table 1) and the time-consuming nature of both the sample preparation and the NMR experiment, a structural model is essential to guide the choice of distances to measure. Finally, solid-state NMR requires large sample quan-

Table 1 Distances Measurable by R² (Homonuclear Spin Pairs) and REDOR (Heteronuclear Spin Pairs), Shown in Angstrom

	Fluorine	Phosphorus	Carbon	Nitrogen
Fluorine	15.3	11.5	9.8	7.3
Phosphorus		8.7	7.4	5.5
Carbon			6.3	4.7
Nitrogen				3.5

Internuclear distances corresponding to a weak dipolar coupling of 30 Hz are listed. Such weak couplings are challenging to measure; the actual upper limit of measurable distance depends on various factors, including signal-to-noise ratio.

ties, on the order of 1–10 mg of isotopically labeled protein per sample, depending on the available magnetic field. This is a large amount for most biological samples, which typically requires efficient overexpression in *E. coli*. So far, unfortunately, efficient overexpression in *E. coli* is difficult for many eukaryotic membrane proteins of pharmaceutical interest. A further question is whether sufficient spins can be put into the very limited sample volume of the MAS rotor (70–300 μ l). Small volume rotors (\approx 70 μ l) are often required to achieve high spinning speeds (needed for rotational resonance) and high power decoupling (needed to get narrow resonances and the longest possible T₂ values, see below). Obtaining sufficient sensitivity is especially challenging for membrane proteins because a significant fraction of the sample will be lipid. Because a major advantage of solid-state NMR is the ability to investigate the protein in a functional state in its lipid bilayer environment, it is critical that the lipid/protein ratio and the hydration be high enough to maintain the functional state of the protein. Recent experiments on the serine receptor illustrate the sample quantities and signal averaging times needed; we packed approximately 10 mg of receptor (=0.167 μ moles of ¹³C spins at the target site) into a 98 μ l volume rotor and collected 40,000 scans (\approx 11 hours) per spectrum (S and S₀) for a REDOR distance measurement (48 and 64 rotor cycles of dephasing) on a 300 MHz spectrometer. For the natural abundance correction (see below), the same experiment must be performed on an ¹⁹F-only labeled sample. Thus to determine whether a distance changes upon ligand binding to the receptor requires a minimum of four spectra on each of three samples (ligand-free and ligand-bound ¹³C¹⁹F-labeled, and ¹⁹F-only labeled), totaling 12 spectra or about 6 days of spectrometer time.

The following are the key steps in choosing a spin pair for a solid-state NMR distance measurement to address a question of interest.

1. Structural model: Choose spin pairs within the distance and precision limits of the technique that address the question,
2. NMR requirements: Choose nuclei to satisfy the requirements of the technique, and
3. Sample labeling requirements: Choose a spin pair that includes a unique site (to target the experiment) and a low frequency site (to minimize natural abundance corrections).

Use a Structural Model to Choose Measurable Distances That Address the Question of Interest

Table 1 summarizes the distance ranges measurable by each technique and spin pair. A further consideration is whether sufficient precision is available to answer the question being posed. Typically, the stronger the dipolar coupling being measured, the higher the potential precision of the measurement. For ¹³C{¹⁹F} REDOR, 6 Å distances can be measured with high precision (\pm 0.3 Å, see Fig. 2) to detect subtle conformational changes between elements of secondary structure (helix-helix distances to detect possible helix piston motions). Longer ¹³C¹⁹F distance measurements of 9–10 Å are less precise (\pm 1 Å) but can provide distance restraints for distinguishing which of the typical side chain rotamers (22) is adopted by the p-F-Phe. Weaker dipolar coupling measurements, such as ¹³C¹⁵N REDOR or ¹³C¹³C R², can be used to measure distances of up to 5–6 Å, which can occur between elements of secondary structure and provide access to other types of sites within a protein. These shorter distance measurement tools may be best suited for measuring backbone-backbone distances to determine local secondary structure. Finally, the distance limits and precision depend on the quality of data one is able to obtain. Because REDOR and R² techniques rely on measuring the difference between spectra taken in the presence and absence of a dipolar interaction of interest, a low signal-to-noise ratio will reduce both the precision of the measurement and the ability to measure long distances (which correspond to small differences between the spectra).

Choose Nuclei That Satisfy the Requirements of the NMR Technique

For rotational resonance, two chemically different carbon sites must be chosen, as the distance measurement experiment requires spinning at the chemical shift difference between the two nuclei. In proteins,

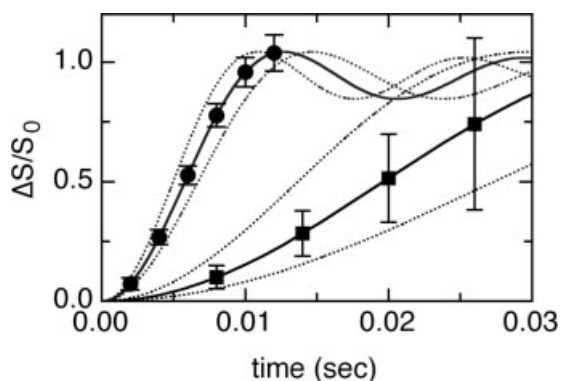


Figure 3 Effect of T_2 on precision of REDOR distance measurements. Circles are idealized $^{13}\text{C}\{^{19}\text{F}\}$ REDOR data for measurement of a distance of 6 Å (solid line) ± 0.3 Å (dotted lines). Squares are idealized data for measurement of a distance of 9 Å (solid line) ± 1.0 Å (dotted lines). Error bars are inversely proportional to the peak intensity as it decays with T_2 relaxation (error bar $\propto \exp[-\frac{t}{T_2}]$), assuming equal numbers of scans for all points.

this means one can use R^2 to measure the distance between a carbonyl carbon (C_1) and a C_β , but not between two carbonyls or two C_β s. For REDOR, the observe nucleus should have as long a T_2 as possible, for reasons described below. Carbonyl nuclei are a good choice, because they have relatively long T_2 values and ^{13}C -carbonyl-labeled amino acids are often less expensive than other labeled positions. Fluorine sites can allow longer distance measurements, but must be chosen based on what fluorinated amino acids are tolerated by the protein synthesis machinery of the cell and do not disrupt the activity of the protein of interest. Aromatic amino acids such as p-F-Phe and 5-F-Trp can be incorporated biosynthetically and have been shown not to disrupt function in some proteins (23), but the activity of the fluorinated protein should be checked. These considerations guide the analysis of the structural model for measurable distances that address the question of interest. One could focus for example on spin pairs in the region of interest that involve a carbonyl and a dissimilar carbon (for $^{13}\text{C}^{13}\text{C}$ R^2) or an aromatic amino acid (for $^{13}\text{C}^{19}\text{F}$ REDOR).

Signal losses due to T_2 are a significant concern in the REDOR experiment, because the observed spins are in the xy plane during the dephasing period. Weak dipolar couplings (long distances) require long dephasing times, at which the NMR signal intensities are smaller due to T_2 relaxation. Figure 3 illustrates the effect of T_2 on the precision of REDOR distance measurements; long distance measurements have lower precision due to the competition between T_2 relaxation and dephasing. As shown in Fig. 3, observing significant dephasing for 6 and 9 Å $^{13}\text{C}^{19}\text{F}$

distances requires dephasing times of 10 and 25 ms, respectively. A 9 ms T_2 (as we have observed for tyrosine carbonyl carbons in the serine receptor) predicts a signal decay to 33% at 10 ms and 6% at 25 ms dephasing times, making the 9 Å distance significantly more challenging to measure. The error bars in Fig. 3 demonstrate the increased error due to T_2 decay of the signal at longer dephasing times, which decreases the precision of the longer distance measurement: 6 ± 0.3 Å versus 9 ± 1.0 Å. Although the number of scans is typically increased for long dephasing times to partially offset the intensity loss due to T_2 decay, it is impractical to increase it much more than twofold (12 \rightarrow 24 hours). Thus it is important to choose sites and decoupling conditions to maximize T_2 . For example, the expected shorter T_2 for a CH_2 site (we have measured 5 ms for a cysteine CH_2 in the serine receptor) predicts a signal decay to 14% at 10 ms dephasing time. Therefore we typically choose to label Cys carbonyls for REDOR measurements. Finally, although incomplete decoupling has been shown not to interfere with the accuracy of the REDOR distance measurement (24), it can decrease the precision due to the shorter T_2 . In summary, choice of sites and decoupling conditions that maximize T_2 is important to increase the precision and distance limits of REDOR measurements.

For the R^2 experiment, the magnetization is stored on the z axis during magnetization exchange, so sites with shorter T_2 values can be used without suffering major losses in sensitivity. This makes it practical to measure CO to CH_2 distances as we have done in the serine receptor. The decoupling requirements in this experiment are different than they are in REDOR. High-intensity decoupling is not as important because, as mentioned, T_2 -dependent losses are much less significant. However, the stability of the decoupling intensity may become important. The dynamics of magnetization exchange in this experiment depend critically on the zero-quantum T_2 (T_2^{ZQ}) (25), and this may change if the decoupling level changes. Furthermore, exchange depends sharply on having the chemical shift difference between the two peaks be exactly a multiple of the spinning speed. The extent to which this condition is met depends on the linewidths, which can change with the decoupling efficiency. Given these dependencies it is possible that in some cases fluctuations in the decoupling power throughout the acquisition of several points could lead to distortions of the exchange curve. In practice, however, it appears that fairly severe power fluctuations can be tolerated. Figure 4 shows that similar RR exchange occurs in $2,4\text{-}^{13}\text{C}$, ^{15}N tyrosine even for experiments with grossly different decoupling powers in the high

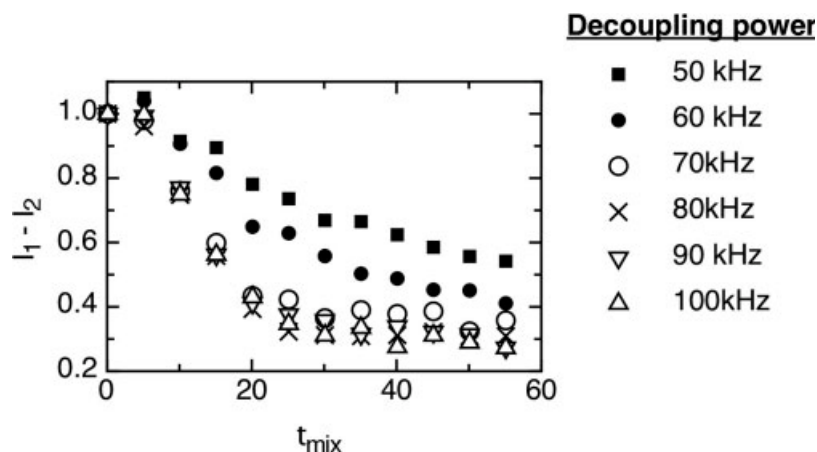


Figure 4 Effect of decoupling power on R^2 magnetization exchange. Magnetization exchange (normalized $I_1 - I_2$) for $n = 1$ R^2 experiment on $2,4\text{-}^{13}\text{C}$, ^{15}N tyrosine is similar over a range of high decoupling powers (80–100 kHz) but is significantly slower at low decoupling powers (≤ 60 kHz).

power range (80–100 kHz), but not at lower powers (≤ 60 kHz). Although this power dependence might differ in a sample with larger inhomogeneous line-widths (chemical shift dispersion), Fig. 4 suggests that decoupling power stability is not a critical concern in the high power range.

Design a Labeling Strategy to Create a Spin Pair Consisting of One Unique and One Low-Frequency Label

The experiment is directed to the site of interest by creating a uniquely labeled site. For some protein applications, one can use a labeled ligand or chromophore or label an amino acid that occurs uniquely in the protein sequence. A more general approach is to introduce a unique amino acid into a protein at the site of interest by site-directed mutagenesis. For example, the serine receptor sequence lacks Cys residues, so a unique Cys can be introduced. The mutant protein must be tested for functionality with activity assays. The widespread use of Cys mutagenesis for a variety of techniques may make it possible to select a site that is already known to tolerate Cys substitution.

The other half of the spin-pair, which typically is not unique, should occur at low frequency in the protein to minimize the natural abundance corrections. This nonunique site will contribute additional dephasing of or magnetization exchange with the natural abundance ^{13}C throughout the protein. This contribution is measured on a single-labeled sample and then subtracted to obtain the distance at the site of interest (see below). Choice of a rare amino acid in the protein sequence is important to minimize the magnitude of this correction. For example, for the REDOR

and R^2 experiments on the serine receptor described in Fig. 2, the nonunique site is a Phe, which occurs 13 times in the serine receptor. This leads to a smaller correction for natural abundance exchange or dephasing (measured on the Phe-labeled sample) than would be obtained for an abundant amino acid such as Gly, which occurs 36 times in the receptor.

Finally, important practical considerations in choosing amino acids are scrambling and cost. One must avoid labeling amino acids such as Glu that are so central in metabolism that a label in Glu would be scrambled into many amino acids. Ideally, one chooses amino acids that are commercially available at a reasonable price per liter of defined media. The amino acid cost alone does not determine which amino acids are practical for biosynthetic labeling because the required amounts in the growth medium vary considerably. For example, although $1\text{-}^{13}\text{C}$ -Cys costs about 10 times as much as $1\text{-}^{13}\text{C}$ -Gly, labeling costs are about the same for these two amino acids because the media contains about a tenth as much Cys as Gly.

To summarize, design of an experiment to measure a distance of interest requires choosing a pair of nuclei whose internuclear distance probes a question of interest, is measurable with sufficient precision, includes a carbonyl site, and can be labeled with one site (the carbonyl for REDOR) being unique (e.g., the site can be made into a unique, functional Cys mutant) and the other site occurring at low frequency in the protein. In designing the serine receptor distance measurements we wanted to measure distances to test the proposed ligand-induced piston motion of $\alpha 4$ -TM2. We scanned the structure for measurable interhelical distances between this helix and the others. For R^2 in the transmembrane region we

scanned the structural model for TM1 to TM2 distances ≤ 6 Å between the low-frequency amino acids such as Phe (carbonyl label, because others are too costly) and sites for potential Cys mutants (C_β label, because C_α is too costly). For REDOR we first found a Phe on $\alpha 4$ that is isolated (at least 15 Å away from all other Phe) and then scanned for carbonyl sites on the neighboring helices at distances ≤ 9 Å. Such site-directed distance measurements are worthwhile if a few precise distances can address a key question that cannot be more easily addressed by other approaches. Due to the subtle nature of the conformational changes proposed to transmit the signal in the serine receptor, solid-state NMR is an ideal probe of mechanism in this protein.

METHODOLOGY FOR ACCURATE DISTANCE MEASUREMENTS IN LARGE PROTEINS

Preparation of a Native, Homogeneous Protein Sample Labeled at the Sites of Interest

To express NMR quantities of isotopically labeled proteins, it is essential to have an efficient overexpression plasmid, and it is simplest to do the expression in *E. coli*. Site-directed mutagenesis to create unique sites in the protein (eg Cys in the serine receptor) is routine. Easy-to-use kits are available (Stratagene and Invitrogen) that give guidance on everything from how to design mutagenesis primers to troubleshooting with control reactions. DNA sequencing is essential to confirm the presence of the mutation and the absence of any new errors in the protein sequence. In the case of highly overexpressed cysteine-less proteins like the serine receptor, a quick screen for introduction of a Cys can first be performed via SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) on the cell cultures to determine whether Cys-crosslinked dimers of the protein are present. The cell cultures are treated with an oxidizing agent and then prepared for SDS-PAGE under nonreducing conditions (15). The activity of the mutant is checked by an assay appropriate for the protein under study.

Protein expression for NMR sample preparation is performed in a bacterial strain and media that allows specific incorporation of the desired labeled amino acids. When possible, the strain of bacteria used for expression of the isotopically labeled protein for NMR should be auxotrophic for (unable to synthesize) the desired labeled amino acids (see [\[cgsc.biology.yale.edu\]\(http://cgsc.biology.yale.edu\) or <http://www.atcc.org> for information on available *E. coli* genetic stocks\). This eliminates synthesis by the bacteria of the unlabeled amino acid. Then the level of label \(which must be known for the distance calculation\) is equal to that of the supplied amino acid. A Phe auxotroph is essential for p-F-Phe labeling, or there will be very low p-F-Phe incorporation levels, due to the toxicity of this amino acid analog. Auxotrophy of stocks should be checked periodically by inoculating equal numbers of cells into parallel cultures of defined media with and without the amino acid in question; the desired outcome is no growth in the absence of the amino acid. The bacteria are grown in defined growth media \(26\) containing all of the amino acids necessary for protein synthesis, including the desired labeled amino acids. When preparing labeled proteins, single-labeled and unlabeled proteins should be prepared in parallel as needed for the measurement of natural abundance contributions \(see below\).](http://</p>
</div>
<div data-bbox=)

Because p-¹⁹F-Phe is toxic to *E. coli*, it is necessary to use a media shift for p-F-Phe labeling. Cells are first grown to near saturation ($OD_{600} \sim 1.0$) in defined media containing only unlabeled amino acids, pelleted via centrifugation, and resuspended in fresh defined media containing the ¹³C labeled amino acid and p-¹⁹F-Phe. These manipulations are done using sterile technique to avoid contamination of the media with other bacteria. Protein expression is then activated by addition of the appropriate inducer. Typically, *E. coli* will express protein but will not continue to divide in the p-F-Phe-containing media. The label incorporation level for p-F-Phe must be measured; two approaches are amino acid analysis (but most facilities do not routinely measure p-F-Phe) or a simple assay involving proteolytic release of soluble fragments from the membrane for analysis by mass spectrometry (27).

The details will vary for each protein, but the goal of sample preparation is a functional, structurally homogeneous sample. The functionality of the protein is determined with activity assays. Structural homogeneity is judged by linewidths at half height; 1.5–2.5 ppm is typical for noncrystalline solids and narrower resonances have been observed for microcrystalline preparations (28). It may be necessary to perform NMR experiments on frozen samples for maximum sensitivity; frozen serine receptor samples give threefold stronger cross-polarization signals than the unfrozen sample. It is important to choose appropriate buffer conditions for freezing. For example, one should avoid sodium phosphate buffers, because they can drop 3 pH units during freezing (29). Rapid freezing with cryoprotectants minimizes freeze-thaw

damage that can denature proteins. For the serine receptor, 0.1% PEG-8000 (polyethylene glycol) is effective at reducing freezing-induced line-broadening. NMR experiments on unfrozen samples must use low duty cycles so that sample heating does not damage the protein (see below). Additionally, for NMR experiments on frozen samples, sample heating can damage the protein if it undergoes multiple freeze-thaw cycles.

Lyophilized samples can be used to increase the amount of protein in the NMR rotor (by removing water), but one should be concerned about whether the protein remains native. This can be judged in part by evaluating linewidths. We performed initial $^{13}\text{C}\{^{19}\text{F}\}$ REDOR studies of the serine receptor on lyophilized samples, with one data point on a frozen sample, and reasoned that the consistency of the dephasing data on the frozen and lyophilized samples indicated that the sample was not perturbed by lyophilization.

NMR Hardware: Cooling and Stability

REDOR and R^2 experiments are fairly robust, but care must be given to the experimental setup to obtain accurate distance measurements in proteins. One important issue is adequate cooling to prevent compromising of both the sample and the NMR experiment. As discussed previously, efficient proton decoupling (typically 70–100 kHz) is important for obtaining the longest possible T_2 and T_2^{ZQ} , which are key to being able to measure long distances in REDOR and R^2 , respectively. These high decoupling powers have been shown to cause significant sample heating (30). Furthermore, the high spinning speeds needed for R^2 also cause frictional heating of the sample. Typically, experiments are conducted on frozen samples, but if heating during the decoupling pulse causes sample thawing, two problems result: (1) probe detuning, which decreases the NMR signal and can be mistaken for magnetization exchange (see constant time R^2 experiment discussed below); and (2) sample degradation due to multiple freeze-thaw cycles and heating. The solution is to monitor the reflected proton power during the experiment for signs of probe detuning, to ensure that the chosen nominal temperature is low enough that the spinning and decoupling do not thaw the sample. Furthermore, the actual sample temperature is not known unless temperature calibration experiments are performed under comparable spinning, decoupling, and ionic strength conditions. Low-temperature cooling devices (FTS Systems, Inc.) greatly simplify maintaining low sample temperatures over the course of long experiments.

The challenge of observing a single spin in a large protein requires long signal-averaging times (≥ 6 days for three samples, see above). Instabilities in the room temperature and other factors can lead to changes in power levels over these periods, which can alter the signal strength due to changes in cross-polarization efficiency. This can be addressed by referencing each NMR spectrum to a control spectrum collected during the same time period. This is best achieved by cycling between S and S_0 spectra in a REDOR experiment or cycling between 0 and 1 or more nonzero mixing times in an R^2 experiment. Any instabilities will reduce intensities in both spectra equally so that the ratio S/S_0 or $t_{\text{mix}}/t_{\text{mix}} = 0$ will correct for instabilities in the CP signal intensity. Because this does not correct for all problems (see below), active feedback control of the amplifier outputs is probably the best solution for the long experiments required for NMR of biological solids (31).

Note the different effects of hardware instability on the R^2 and REDOR experiments. Both require that the spinning speed remain stable over the course of the experiment or the distance measurement will not be accurate. The safest course of action is to keep a log of the speed throughout the experiment and check it for deviations. Both experiments also will suffer reduced sensitivity if there are pulse power fluctuations, because these will reduce the cross-polarization efficiency. Power instability will compromise the accuracy of the distance measurement in different ways in a REDOR and R^2 experiment. For REDOR, if power levels change so that the dephasing pi pulses are no longer properly calibrated, incomplete dephasing will occur and the distance will appear longer than it is. For R^2 , if the proton decoupling power level changes enough to change the T_2^{ZQ} or linewidths, this can alter the magnetization exchange rate and lead to an inaccurate distance measurement (see Fig. 4). An internal control (discussed below) should be used to check whether such instabilities have compromised the accuracy of the experiment.

REDOR: Optimizing Pulses and Testing Distance Accuracy

There are several important variations of the REDOR pulse sequence that are appropriate for different applications (see Fig. 1). REDOR sequences that alternate the pi pulses between the dephasing and observe channels (see Fig. 1(d), option 2) have been observed to give more complete dephasing (32) but are *not* recommended if homonuclear couplings are present for the observe nucleus. For the latter case, the REDOR sequence with a single observe pi pulse (see Fig. 1(d), option 1) is considered to be best because it does not recouple the homonuclear interaction (that can inter-

ferre with the REDOR dephasing). For guidance as to the best REDOR sequence for a particular application, SIMPSON (33) can be used to simulate dephasing for the anticipated coordinates of the spin system for both versions of REDOR to determine which will perform best. The REDOR pulse sequence should also be checked carefully to be sure that all the pulse lengths are included in the calculated delays so that the pi pulses are really centered in the rotor periods.

Another variable in the REDOR pulse sequence is the position of the dephasing pulse. Maximum dephasing (and thus maximum sensitivity for the ΔS signal) is obtained when the pi pulses are centered in the rotor periods, and these pulse sequences should be used for the actual measurement of an unknown distance. A reduced-coupling version of REDOR, in which the pi pulses are shifted away from the center of the rotor period, is useful for optimization and testing of the REDOR experiment (see Fig. 1(d) with gray pulses (32)). The shift of the pi pulses decreases the dephasing so that the REDOR experiment can be conducted with a larger number of rotor cycles. This is important because distance accuracy should be checked in a REDOR experiment that uses as many pi pulses as will be needed for the distance measurement of interest. Imperfect pi pulses lead to incomplete dephasing, and this deviation grows larger with increasing numbers of pi pulses due to accumulation of the errors. This can cause the REDOR dephasing data to fall below the predicted REDOR curve at high numbers of rotor cycles. Thus a pulse calibration problem that may not be apparent in an eight-rotor cycle experiment could accumulate into a significant error in a 64-rotor cycle experiment. Because the shifted REDOR pulse sequences allow for experiments with larger numbers of rotor cycles, they are ideal for optimization but not for the actual measurement of an unknown distance.

Optimization of the REDOR experimental parameters is best done on a small molecule with strong NMR resonances. Calibration of the dephasing 180° pulse length (the ^{19}F pulse in a $^{13}\text{C}\{^{19}\text{F}\}$ REDOR experiment) can either be done directly, in an ^{19}F -observe experiment, or indirectly, in a ^{13}C -observe, ^{19}F -dephase REDOR experiment. Calibration via the REDOR experiment works well and is essential in spectrometers in which the path of the ^{19}F rf power differs for ^{19}F -observe versus ^{13}C -observe experiments. The approach is to choose a safe ^{19}F power level and perform the $^{13}\text{C}\{^{19}\text{F}\}$ REDOR experiment using a range of pulse lengths. The pulse length that gives the minimum remaining signal, which corresponds to maximum dephasing, is chosen for the 180° pulse.

Testing the accuracy of the REDOR distance measurement must be done on a sample with the spin pair of interest at a known distance and isolated from other spins. The latter is typically achieved by isotopic dilution; the labeled compound is mixed before crystallization with a known amount of the corresponding unlabeled compound (typically, labeled/unlabeled ratios of 1:10), to dilute out intermolecular spin pairs and leave only the isolated intramolecular spin pair. For REDOR involving ^{19}F , isotopic dilution can be difficult to achieve, as the fluorinated and unfluorinated compounds may phase separate. We have used a fluoropolycarbonate sample kindly provided by Jake Schaefer, in which the fluorine is self-diluted because it occurs on every fourth phenyl ring. The natural abundance ^{13}C spectrum of fluoropolycarbonate contains one resolved signal arising from the fluorinated rings, which can be used to test the accuracy of a REDOR measurement of a 2.36 \AA ^{13}C to ^{19}F distance (32). For this very strong dipolar coupling (rigid lattice predicts 2,177 Hz; lattice motions reduce this by $\sim 5\%$ to 2068 Hz (32)), REDOR dephasing would be complete in less than eight rotor cycles using the normal "centered" REDOR experiment, so the "shifted" REDOR experiment with attenuated dephasing is used (Fig. 1(d) with gray pulses). REDOR experiments on this molecule (similar to those described by Weldeghiorghis and Schaefer (32) except with a spinning speed of 7.5 kHz and a shift of 0.1) demonstrate that measurement of this strong dipolar coupling is accurate; the REDOR data fit a dipolar coupling of $1976 \pm 31 \text{ Hz}$ ($2.43 \pm 0.01 \text{ \AA}$), with no corrections needed. This indicates that the experimental parameters are optimized sufficiently to give accurate REDOR dephasing for up to 32 rotor cycles. However, distances of interest are likely to require more rotor cycles of dephasing (e.g., the serine receptor experiments shown in Fig. 2 measured dephasing to 64 rotor cycles). We have recently demonstrated accurate REDOR dephasing at larger numbers of rotor cycles with an easily prepared complex containing an isolated $^{13}\text{C}^{19}\text{F}$ spin pair separated by 5.6 \AA (D. Fowler, unpublished results).

Because the self-dilution of the $^{13}\text{C}^{19}\text{F}$ pairs makes the signal of interest weak, experiments on fluoropolycarbonate are time-consuming. After verifying that the experimental parameters were optimized for accurate short ^{13}C to ^{19}F distance measurements using fluoropolycarbonate, we tested to see whether a commercially available compound, 6-fluorogranine, could be used for more rapid optimization of REDOR experimental parameters. Figure 5(a) shows the spectrum and the resolved signal of a ring carbon (black circle) adjacent to the fluorinated site; this resonance was used for the pi

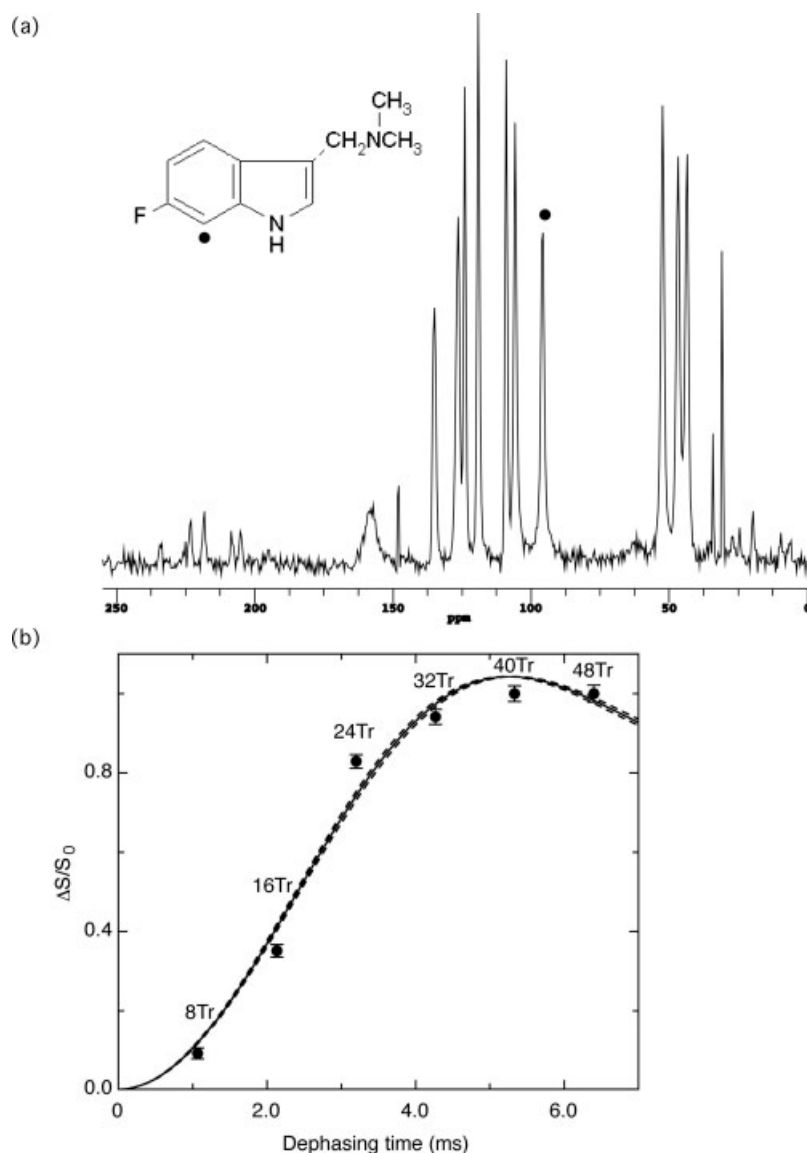


Figure 5 REDOR on 6-fluorogranine provides an efficient means to demonstrate accurate short-distance measurements. (a) 6-Fluorogranine and its natural abundance ^{13}C spectrum. The black circle indicates the 96.1 ppm peak of carbon 7 at a two-bond distance of 2.4 Å from the fluorine. (b) The REDOR data fit a distance of 2,047 (solid line) \pm 25 Hz (dashed lines), corresponding to 2.40 \pm 0.01 Å. REDOR was performed at a spinning speed of 7.5 kHz with a shift of 0.1.

pulse calibration described above. The REDOR data in Fig. 5(b) demonstrate that dilution of 6-fluorogranine is not needed to obtain REDOR data that fits the expected 2.4 Å distance; apparently the intermolecular $^{13}\text{C}^{19}\text{F}$ dipolar couplings are weak enough that they do not interfere with this measurement of a strong $^{13}\text{C}^{19}\text{F}$ dipolar coupling. Thus, 6-fluorogranine is a readily available compound for optimization of experimental parameters and testing dephasing accuracy, but only to 32 rotor cycles.

Because the B_1 field is not perfectly homogeneous in any probe, the appropriate sample size must be

chosen for each application. The B_1 homogeneity can be improved by restricting the sample to the center of the rotor with symmetrically placed spacers, but the smaller sample volume will reduce sensitivity. The effects of B_1 inhomogeneity on the experiment in each probe should be investigated by comparing results on various sample sizes to determine what is the largest sample (strongest signal) that gives accurate REDOR distance measurements using the full number of rotor cycles needed for the distance measurement of interest. The above experiments yielded accurate distances with a full rotor of 6-fluoro-

gramine, indicating that B_1 homogeneity is adequate; a full rotor sample can be used in this probe for REDOR experiments to 32 rotor cycles.

R²: Optimizing Pulses and Testing Distance Accuracy

A pulse sequence that minimizes artifacts for rotational resonance distance measurements in proteins (see Fig. 1a) employs a delay to achieve selective inversion and a constant decoupling time (34). The selective inversion is not achieved via a long weak pulse because this method was found to introduce artifacts (35) that are absent when a delay inversion is used. The use of a constant decoupling time is important so that the zero mixing time spectrum is a more comparable control experiment (36). As mentioned previously, the combination of high spinning speeds and long mixing times with high-power decoupling can significantly detune the probe and even thaw a frozen sample. Furthermore, as the mixing time is increased the detuning will also increase, causing a decrease in signal intensity that is indistinguishable from magnetization exchange; this can result in a systematic error in the distance (distance appears shorter due to detuning contribution). Use of a constant mixing time ensures that all time points undergo the same extent of detuning. This can be achieved with an additional delay in the pulse sequence (34), but it is simpler to add a postacquisition decoupling period (τ') as shown in Fig. 1a. Another approach is to do a sum polarization experiment (without selective inversion) at each mixing time as a control to measure any intensity losses resulting from processes other than magnetization exchange (25). However, because this approach requires collecting twice as many spectra (two spectra per mixing time), it is less practical than using a constant decoupling time for distance measurements in proteins.

A critical element of the pulse program is that it must include a phase cycle that removes any magnetization left in the xy plane during the mixing time. Such xy magnetization, which results from imperfect pulses, must be removed because its rapid T_2 decay is indistinguishable from magnetization exchange. As long as the pulse sequence includes phase cycling to remove this effect, imperfect pulses will degrade the signal but not compromise the accuracy of the distance measurement. This is illustrated in Fig. 6, which shows R² magnetization exchange data for 2,4'-¹³C, ¹⁵N tyrosine (carbon carbon distances of 5.12 Å) that were collected with optimum ¹³C 90° pulses (filled triangles, 5 μ s pulse length) and purposely mis-set pulses (open triangles, 2.5 μ s pulse length). The experiment performed with mis-set pulses suffers a loss of sensitivity (error bar

doubles, becoming approximately equal to the symbol size), but no loss in distance accuracy, even for this long distance. Thus calibration of the pulses is less critical for R² than for REDOR (except that sensitivity is critical for distance measurements in proteins), and there is less need to check R² experimental parameters on a molecule with a known long distance spin pair.

To maximize the sensitivity of detection of the magnetization exchange, one should optimize the CP conditions (typically ramped CP) on a molecule with similar carbons to obtain the maximum total signal for the two carbons of interest (Gly carboxyl and methylene carbons were used to calibrate CP for an experiment measuring distances between a backbone carbonyl and a Cys methylene). Because the R² experiment will maintain a constant sum polarization, both signals approach a final intensity equal to half of their initial sum, for a total intensity change equal to the total initial signal. For example, initial signals with a 1:–1 or a 1.5:–0.5 ratio (after selective inversion) would both change by a total of 2 to reach full exchange, at 0:0 and 0.5:0.5, respectively). Thus to maximize the sensitivity of the distance measurement, CP conditions should be chosen to maximize the sum of the signals at the two sites rather than to equalize their intensities.

The most challenging aspect of obtaining accurate distances with R² is the need to have accurate values for the T_2^{ZQ} . T_2^{ZQ} is estimated by performing Hahn echo experiments to measure the single quantum T_2 values for each site, which are combined as follows:

$$\frac{1}{T_2^{ZQ}} \approx \frac{1}{T_2^{(1)}} + \frac{1}{T_2^{(2)}}.$$

Furthermore, magnetization exchange rates are also affected by chemical shift dispersion, which puts the spins slightly off the R² condition. The linewidths of each site and T_2^{ZQ} are included in the R² simulation and are critical to obtaining an accurate distance. A plot of the deviation between predicted and experiment for all values of dipolar coupling and T_2^{ZQ} shows no unique minimum because a range of values for this pair can fit the data (see Fig. 9 (d) in (25)). R²W (R² width) is an elegant variation of the R² experiment that eliminates this problem, making it easier to obtain accurate distances (25). R²W measures magnetization exchange as a function of the spinning speed (the offset from the R² condition) rather than the mixing time. Because the T_2^{ZQ} and dipolar coupling affect the exchange differently, the resulting data uniquely determine the dipolar coupling with no knowledge of the T_2^{ZQ} . A plot of the deviation between predicted and experiment for all values of dipolar coupling and T_2^{ZQ} shows a unique minimum (see Fig. 9 (b) in (25)), so the dipolar coupling is determined with no knowledge of the value of T_2^{ZQ} . Although we are unaware of

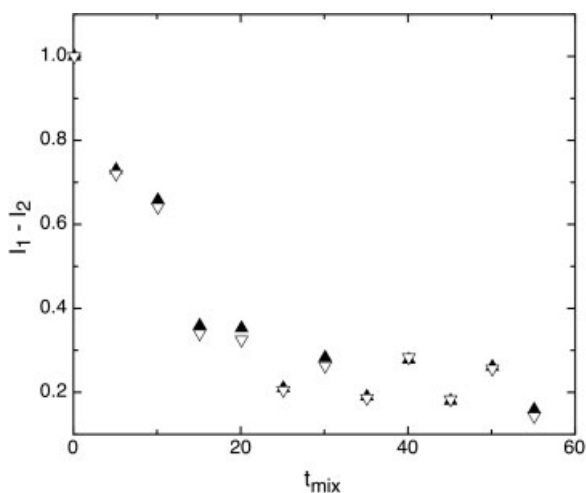


Figure 6 R^2 magnetization exchange is not affected by errors in the 90° pulse length. R^2 experiment ($n = 1$) performed on $2,4\text{-}^{13}\text{C}$, ^{15}N tyrosine (diluted 1:9 in natural abundance tyrosine) with correct ^{13}C 90° pulse length of $5\ \mu\text{s}$ (solid triangles) and incorrect ^{13}C 90° pulse length of $2.5\ \mu\text{s}$ (open triangles) shows equivalent magnetization exchange. Error bars are approximately twofold larger for the $2.5\ \mu\text{s}$ experiment, becoming approximately equal to the open triangle symbol size.

any applications of R^2 W to distance measurements in a large protein, the experiment appears practical. Sensitivity and exchange comparable to R^2 are expected; if a constant decoupling time version is used, then only two spectra are needed at each spinning speed (zero and the single chosen mixing time) and the total number of spectra needed appears comparable to an R^2 experiment.

Natural Abundance Correction: Determination of an Accurate Distance

For analysis of REDOR and R^2 data it is critical to correct for the contribution of natural abundance to the experiment and thus determine the amount of dephasing or magnetization exchange due to the sites of interest, which can then be analyzed to determine an accurate internuclear distance. Applications to the serine receptor illustrate the fact that this correction can be significant, particularly for measurement of weak dipolar couplings. For a $^{13}\text{C}^{13}\text{C}$ R^2 measurement in the serine receptor, the natural abundance correction changed the distance from 4 to $5\ \text{\AA}$ (15). Similarly, for a $^{13}\text{C}\{^{19}\text{F}\}$ REDOR measurement of a weak dipolar coupling, the natural abundance correction changed the distance from 8.9 to $10.1\ \text{\AA}$. In contrast, the correction was much less significant for the strong $^{13}\text{C}^{19}\text{F}$ dipolar coupling measurements shown in Fig.

2(a); the correction changed a $6\ \text{\AA}$ distance by $0.2\ \text{\AA}$ (9). The natural abundance correction will typically be larger for experiments on proteins than for synthetic peptides because of the large number of unlabeled residues and the multiplicity of the nonunique site.

Two types of corrections are needed. The contribution of natural abundance signals to both the S_0 signal and the ΔS (REDOR dephasing or R^2 magnetization exchange) must be removed in order to obtain an accurate distance for the site of interest. The correction to the S_0 signal is obtained by comparison with spectra of samples that are unlabeled at the observe nucleus (e.g., ^{19}F -labeled samples are used for $^{13}\text{C}\{^{19}\text{F}\}$ REDOR, and unlabeled samples are used for $^{13}\text{C}^{13}\text{C}$ R^2). We have found that preparation of these samples in parallel gives the best match between spectra of the different samples, minimizing difference artifacts. Unless equal masses of protein can be packed in each rotor (difficult for hydrated membrane-bound proteins), one spectrum must be scaled to match the other. The scaling factor is chosen using a region with a strong signal that should be equivalent in both samples (e.g., the C_α signal). It is difficult to avoid difference artifacts when matching spectra taken on two different samples, which makes this one of the weakest aspects of the data analysis. This correction to S_0 is for the observe nucleus only in a REDOR experiment, and for both the I_1 and I_2 sites in an R^2 experiment.

The correction to the ΔS signal is illustrated in Fig. 7(a) for a REDOR experiment measuring the distance between a unique site and a nonunique site (e.g., a single $1\text{-}^{13}\text{C}$ -Cys and $13\ \text{p-F-Phe}$ in the serine receptor, Fig. 2). To determine the dephasing between the sites of interest (pathway a, if only one Phe is near the unique Cys), the contributions of the other pathways must be removed. Pathway d, natural abundance dephasing of natural abundance, is often negligible. For $^{13}\text{C}\{^{19}\text{F}\}$ REDOR, pathway b is also negligible (as there is no natural occurrence of fluorine in proteins). Pathway c, the $13\ \text{p-}^{19}\text{F-Phe}$ dephasing the natural abundance ^{13}C , is significant. The contribution of pathway c is measured for each dephasing time with a REDOR experiment on a ^{19}F -labeled sample (SL, single labeled).

Similarly, Fig. 7(b) illustrates the natural abundance contribution to magnetization exchange for rotational resonance between a unique site and a nonunique site (e.g., a single $^{13}\text{C}_\beta$ -Cys and $13\ 1\text{-}^{13}\text{C-Phe}$ in the serine receptor, Fig. 2). Pathway d will again be negligible. Pathway b is also negligible because there is only a single Cys; only atoms having the correct chemical shift and located within a few angstroms of the single Cys $^{13}\text{C}_\beta$ can contribute significant exchange, and at natural abundance these few sites are occupied at only 1%. The contribution of pathway c is

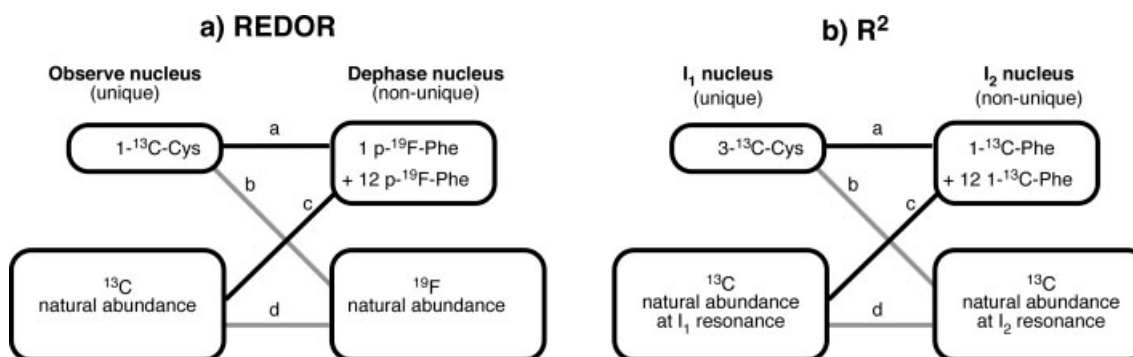


Figure 7 REDOR dephasing and R^2 exchange pathways illustrating the potential natural abundance contributions. Gray lines indicate dephasing/exchange pathways that are not significant for this application, due to the identity of the nuclei and the number of sites. Dephasing and exchange involving the nonunique site is the most significant (pathway c, black), and is measured experimentally in a SL (single-labeled) sample (labeled with the nonunique site label, which is the dephase nucleus for REDOR or I_2 for R^2). The dephasing/exchange measured on the SL sample is subtracted from that measured on the DL (double-labeled) sample to obtain the dephasing/exchange due to the site of interest.

measured by performing the R^2 experiment on a $1\text{-}^{13}\text{C}$ Phe sample (SL, single labeled with the nonunique site label) to determine the magnetization exchange between the natural abundance signal at the $^{13}\text{C}_\beta\text{-Cys}$ resonance position and the 13 $1\text{-}^{13}\text{C}\text{-Phe}$. For each exchange time point, exchange is measured on both double-label (DL, $^{13}\text{C}_\beta\text{-Cys}$ and $1\text{-}^{13}\text{C}\text{-Phe}$) and single-label (SL, $1\text{-}^{13}\text{C}\text{-Phe}$) samples, and the difference gives the exchange due to the site of interest. A test of this analysis on a known distance in the serine receptor (15) demonstrated this method to be reliable and important to the distance accuracy; the correction increased the measured distance from 4 to 5 Å.

The following briefly illustrates how the data are processed to correct for natural abundance contributions to S_0 and ΔS in a $^{13}\text{C}\{^{19}\text{F}\}$ REDOR experiment on a unique ^{13}C -labeled site. REDOR experiments measure both the S signal (with dephasing pulses) and the S_0 signal (without dephasing pulses), where $\Delta S = S_0 - S$. Because $\frac{\Delta S}{S_0} = \frac{S_0 - S}{S_0} = 1 - \frac{S}{S_0}$, REDOR dephasing data is plotted either as $\Delta S/S_0$ or S/S_0 versus dephasing time. We compute S/S_0 to minimize arithmetic operations that can increase propagated errors. As discussed previously, the REDOR experiment must be performed on both a $^{13}\text{C}, ^{19}\text{F}$ -labeled sample (DL) and a ^{19}F -labeled sample (SL) for each dephasing time point, t , to yield a series of data $S_0^{DL}(t)$, $S^{DL}(t)$, $S_0^{SL}(t)$, $S^{SL}(t)$. Let m = scaling factor needed to correct for differences in sample mass and cross-polarization efficiencies between the DL and SL spectra; m is deduced by matching the spectra at regions other than that of the labeled carbonyl resonance (e.g., match the C_α resonance), so that $S_0^{DL}(t) = m \times S_0^{SL}(t)$ outside the carbonyl region for

all dephasing time points. This value of m is used in the following equations to correct for natural abundance contributions to the carbonyl region. Natural abundance contributions to S_0 are removed by computing $S_0^{site}(t) = S_0^{DL}(t) - m \times S_0^{SL}(t)$. Corrections for natural abundance contributions to S spectra are performed similarly; the dephasing of the carbonyl at the site of interest (pathway a) is deduced by subtracting natural abundance dephasing (pathway c, SL sample) from the total dephasing (pathways a and c, DL sample):

$$\frac{S^{site}}{S_0^{site}}(t) = \frac{S^{DL}(t) - m \times S^{SL}(t)}{S_0^{DL}(t) - m \times S_0^{SL}(t)}$$

Finally, one further correction is needed to account for any incomplete labeling of the dephasing nucleus. Dephasing (ΔS) occurs only at ^{13}C sites that have a neighboring ^{19}F , but the S_0 signal due to all the ^{13}C sites must be multiplied by the fractional labeling. The fraction of p-F-Phe label in the sample ($f = [\text{p-F-Phe}]/[\text{total Phe}]$) must be measured independently (see above) and used to correct $\Delta S/S_0$ as follows:

$$\begin{aligned} \left[\frac{\Delta S^{site}}{S_0^{site}}(t) \right]_{corrected} &= \left[\frac{1}{f} \right] \frac{\Delta S^{site}}{S_0^{site}}(t) \\ &= \left[\frac{1}{f} \right] \left\{ 1 - \frac{S^{site}}{S_0^{site}}(t) \right\} \\ &= \left[\frac{1}{f} \right] \left\{ 1 - \frac{S^{DL}(t) - m \times S^{SL}(t)}{S_0^{DL}(t) - m \times S_0^{SL}(t)} \right\} \end{aligned}$$

Error Bars: Determination of the Precision of the Distance Measurement

The precision of the measurement is critical for addressing many biological questions. To address whether ligand binding induces changes in interhelical distances in the serine receptor, for example, requires knowing the precision of each distance measurement so that it is clear when a distance change is significant. It is not clear whether the 5.8 to 6.8 Å ligand-induced change measured in the serine receptor is significant unless one includes the precision of the measurement: 5.8 ± 0.2 to 6.8 ± 0.2 Å is significant.

The baseline error can be estimated as follows. The standard deviation of the noise is determined in a signal-free region of the spectrum. The chosen noise region must not be close to the edges of the spectrum, so that the noise is not artificially lowered by the filters. For a peak height measurement, this standard deviation provides the error bar (\pm SD). For an integral measurement, the standard deviation of the noise must first be multiplied by the square root of the number of points integrated to obtain the appropriate error bar. These errors are then propagated by standard methods through all the arithmetic operations in the REDOR and R^2 data analysis to estimate the final error in the distance. This method was used to estimate the precision of the distance measurements on the serine receptor.

In addition to the baseline error, the signal error may contribute to the precision of the experiment. For example, variations in cross-polarization efficiency due to fluctuations in amplifier output would produce a signal error not included in the baseline error estimate above. The total error, which is the sum of the baseline error and signal error, can be estimated by measuring the block-to-block variations of the signal. Collecting blocks of spectra over the course of a long experiment is generally good practice so that if any problems occur (such as probe arcing or deviations in temperature, tuning, or spinning speed), the affected blocks can be discarded. The unaffected blocks can be combined into several large blocks of data to determine the standard deviation of the peak intensity among these, which provides a conservative estimate of the precision that includes the signal error. This was not done for the serine receptor distance measurements shown in Fig. 2. In recent experiments (Varian 300 MHz InfinityPlus, room temperature range of $\leq 3^\circ$, no active amplifier output stabilization), we observed that for three consecutive blocks of ≈ 18 K scans each, the total error (standard deviation of the peak intensity among the three blocks) is 1.24 times the baseline error within each block. This comparison provides a useful means

to determine if signal variations are a significant issue for the experiment, and provides a conservative approach to estimate the total error; the total 54K scans can be treated as three replicate 18K scan experiments to provide an average and standard deviation. A less conservative approach is to combine all the $3 \times 18\text{K} = 54\text{K}$ scans (to increase the signal to noise by $\sqrt{3}$) and estimate that the total error will continue to be 1.24 times the baseline error in the final block of 54K scans. Either of these approaches provides a better estimate of the precision than the final baseline error.

Improving the Accuracy of the Distance Measurement

We have described several means for improving the accuracy of REDOR and R^2 distance measurements, which are limited by the pi pulse calibration and the need to know the T_2^{ZQ} , respectively. Control experiments are critical to test the REDOR pi pulse calibration; the R^2 W experiment avoids the critical dependence of R^2 on T_2^{ZQ} .

Even with good initial pulse calibrations, drifts in the power levels over the course of the experiment can still compromise the accuracy of REDOR and R^2 distance measurements. If the data are collected in blocks, these can be analyzed to determine if such power drifts have compromised the distance accuracy. Block-to-block variation of the S_0 signal intensity (no REDOR dephasing pulses or zero R^2 magnetization exchange time) are due to the baseline error plus signal error. Block-to-block variation of the S signal intensity (dephased or exchanged) will contain these errors plus any additional error due to variations in the pi pulses (altering dephasing) or in the decoupling (altering exchange). Therefore, if the variation in S_0 and S data is comparable, it is likely that power drifts that compromise the distance accuracy have not occurred over the time course of the experiment. If the standard deviation is not comparable, a plot of the S_0 and S data for all blocks might reveal blocks with substantial deviations that could be discarded. This provides an internal control to determine whether amplifier power fluctuations (e.g., due to room temperature fluctuations) are a problem limiting the accuracy of a distance measurement for the particular spectrometer and experiment.

Finally, the remaining weak link in the determination of an accurate distance is the need to match different samples to make natural abundance corrections. The cleanest means of avoiding this is to incorporate an additional label near the site of interest and use NMR filtering to eliminate natural abundance

contributions. For example, Schaefer and coworkers have developed TEDOR-REDOR (37) and double-REDOR (38) experiments to remove natural abundance contributions in various applications. The double time needed for an NMR-filtered experiment is no longer than that needed to collect spectra on two different samples, and the result is significantly more reliable because it is performed on a single sample. Unfortunately, such REDOR experiments are not accessible in many labs due to the need for four-channel NMR probes.

CONCLUSIONS

Site-directed REDOR and R^2 distance measurements in large proteins are powerful but challenging experiments. We have described strategies for designing and implementing these experiments, illustrated by the application to bacterial chemotaxis receptors to probe mechanisms of transmembrane signaling. Such strategies can be applied to investigate local structure and mechanisms in other important protein complexes. The large sample volumes and long signal averaging times required to measure a single distance may look unattractive when compared with sensitive techniques, such as FRET, or largely parallel ones, such as simultaneously derived NMR distance constraints. Furthermore, the “short ruler” they provide can sometimes be frustrating. However, the precision with which they can determine these distances, and the fact that they can do so in disordered solids, makes them exceptionally powerful in certain cases. If informed by previous structural data, and by competing or testable hypotheses, these techniques can discriminate the subtle structural features that lead to detailed understanding of complex biochemical problems. Solid-state NMR techniques for wide-scale whole-structure determination continue to develop rapidly; but rather than supplanting the high-precision/low-throughput methods described here, it is likely that they will increase the number of cases where REDOR or R^2 experiments can provide the small but crucial datum for a deeper understanding of the structure and mechanism.

ACKNOWLEDGMENTS

This research was supported by U.S. Public Health Service Grant GM47601. F.A.K. was supported by National Research Service Award F32GM20711, and D.J.F. was supported in part by National Research Service Award T32 GM08515 from the National Institutes of Health. We thank Clifford J. Unkefer

and the National Stable Isotope Resource at Los Alamos National Laboratory for synthesis of the $2,4\text{'-}^{13}\text{C}_2,^{15}\text{N}$ tyrosine, Edward G. Voigtman for helpful discussions of error analysis, and L. Charles Dickinson for NMR instrumentation support. This article is dedicated to the memory of Sarah Elizabeth Kovacs (1991–2005).

REFERENCES

1. Wiener AE. 2004. A pedestrian guide to membrane protein crystallization. *Methods* 34:364–372.
2. Hwang PM, Kay LE. 2005. Solution structure and dynamics of integral membrane proteins by NMR: a case study involving the enzyme PagP. *Methods Enzymol* 394: 335–350.
3. Sanders CR, Sonnichsen F. 2006. Solution NMR of membrane proteins: practice and challenges. *Magn Reson Chem* 44(S1):24–40.
4. McDermott AE. 2004. Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. *Curr Opin Struct Biol* 14:554–561.
5. Thompson LK. 2002. Solid-state NMR studies of the structure and mechanisms of proteins. *Curr Opin Struct Biol* 12:661–669.
6. McDowell LM, Poliks B, Studelska DR, O'Connor RD, Beusen DD, Schaefer J. 2004. Rotational-echo double-resonance NMR-restrained model of the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase. *J Biomol NMR* 28:11–29.
7. Petkova AT, Hatanaka M, Jaroniec CP, Hu JG, Belenky M, Verhoeven M, et al. 2002. Tryptophan interactions in bacteriorhodopsin: a heteronuclear solid-state NMR study. *Biochemistry* 41:2429–2437.
8. Crocker E, Eilers M, Ahuja S, Hornak V, Hirshfeld A, Sheves M, et al. 2006. Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. *J Mol Biol* 357:163–172.
9. Murphy OJ III, Kovacs FA, Sicard EL, Thompson LK. 2001. Site-directed solid-state NMR measurement of a ligand-induced conformational change in the serine bacterial chemoreceptor. *Biochemistry* 40:1358–1366.
10. Opella SJ, Marassi FM. 2004. Structure determination of membrane proteins by NMR spectroscopy. *Chem Rev* 104:3587–3606.
11. Gullion T. 1998. Introduction to rotational-echo, double-resonance NMR. *Concepts Magn Reson* 10:277–289.
12. Gullion T, Schaefer J. 1989. Rotational-echo double-resonance NMR. *J Magn Reson* 81:196–200.
13. Raleigh DP, Levitt MH, Griffin RG. 1988. Rotational resonance in solid-state NMR. *Chem Phys Lett* 146:71–76.
14. Goobes G, Raghunathan V, Louie EA, Gibson JM, Olsen GL, Drobny GP. 2006. A REDOR study of diammonium hydrogen phosphate: a model for distance measurements from adsorbed molecules to surfaces. *Solid State Nucl Magn Reson* 29:242–250.

15. Isaac B, Gallagher GJ, Balazs YS, Thompson LK. 2002. Site-directed rotational resonance solid-state NMR distance measurements probe structure and mechanism in the transmembrane domain of the serine bacterial chemoreceptor. *Biochemistry* 41:3025–3036.
16. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proc Natl Acad Sci U S A* 101:10048–10053.
17. Sharpe S, Kessler N, Anglister JA, Yau WM, Tycko R. 2004. Solid-state NMR yields structural constraints on the V3 loop from HIV-1 Gp120 bound to the 447-52D antibody Fv fragment. *J Am Chem Soc* 126:4979–4990.
18. Toke O, Maloy WL, Kim SJ, Blazyk J, Schaefer J. 2004. Secondary structure and lipid contact of a peptide antibiotic in phospholipid bilayers by REDOR. *Biophys J* 87:662–674.
19. Kim KK, Yokota H, Kim SH. 1999. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* 400(6746):787–792.
20. Milburn MV, Prive GG, Milligan DL, Scott WG, Yeh J, Jancarik J, et al. 1991. Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* 254(5036):1342–1347.
21. Chervitz SA, Falke JJ. 1996. Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc Natl Acad Sci U S A* 93:2545–2550.
22. Lovell SC, Word JM, Richardson JS, Richardson DC. 2000. The penultimate rotamer library. *Proteins* 40:389–408.
23. Danielson MA, Falke JJ. 1996. Use of 19F NMR to probe protein structure and conformational changes. *Annu Rev Biophys Biomol Struct* 25:163–195.
24. Mehta AK, Hirsh DJ, Oyler N, Drobny GP, Schaefer J. 2000. Carbon-proton dipolar decoupling in REDOR. *J Magn Reson* 145:156–158.
25. Costa PR, Sun B, Griffin RG. 2003. Rotational resonance NMR: separation of dipolar coupling and zero quantum relaxation. *J Magn Reson* 164:92–103.
26. Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. 1989. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol* 177:44–73.
27. Gallagher GJ. 2006. Solid-state NMR measurements of depths and distances in membrane peptides and proteins [Ph.D. thesis]. University of Massachusetts, Amherst, MA.
28. Tycko R. 2006. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q Rev Biophys* 39:1–55.
29. Pikal-Cleland KA, Rodriguez-Hornedo N, Amidon GL, Carpenter JF. 2000. Protein denaturation during freezing and thawing in phosphate buffer systems: monomeric and tetrameric beta-galactosidase. *Arch Biochem Biophys* 384:398–406.
30. Stringer JA, Bronnimann CE, Mullen CG, Zhou DH, Stellfox SA, Li Y, et al. 2005. Reduction of RF-induced sample heating with a scroll coil resonator structure for solid-state NMR probes. *J Magn Reson* 173:40–48.
31. Stueber D, Mehta AK, Chen Z, Wooley KL, Schaefer J. 2006. Local order in polycarbonate glasses by 13C{19F} rotational-echo double-resonance NMR. *J Polymer Sci [B]* 44:2760–2775.
32. Weldeghiorghis TK, Schaefer J. 2003. Compensating for pulse imperfections in REDOR. *J Magn Reson* 165:230–236.
33. Bak M, Rasmussen JT, Nielsen NC. 2000. SIMPSON: a general simulation program for solid-state NMR spectroscopy. *J Magn Reson* 147:296–330.
34. Balazs YS, Thompson LK. 1999. Practical methods for solid-state NMR distance measurements on large biomolecules: constant-time rotational resonance. *J Magn Reson* 139:371–376.
35. Smith SO, Jonas R, Braiman M, Bormann BJ. 1994. Structure and orientation of the transmembrane domain of glycoporphin A in lipid bilayers. *Biochemistry* 33:6334–6341.
36. Tomita Y, O'Connor EJ, McDermott A. 1994. A method for dihedral angle measurement in solids—rotational resonance NMR of a transition-state inhibitor of triose phosphate isomerase. *J Am Chem Soc* 116:8766–8771.
37. McDowell LM, Schmidt A, Cohen ER, Studelska DR, Schaefer J. 1996. Structural constraints on the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase from rotational-echo double-resonance NMR. *J Mol Biol* 256:160–171.
38. Beusen DD, McDowell LM, Slomczynska U, Schaefer J. 1995. Solid-state nuclear magnetic resonance analysis of the conformation of an inhibitor bound to thermolysin. *J Med Chem* 38:2742–2747.
39. Gullion T, Baker DB, Conradi MS. 1990. New, compensated Carr-Purcell sequences. *J Magn Reson* 89:479–484.

BIOGRAPHIES



Frank A. Kovacs received a B.S. in biology from the University of West Florida. He obtained a Ph.D. in molecular biophysics from Florida State University in 1999, where he worked with Timothy Cross using solid-state NMR to study the structure of oriented proteins. He then worked as a postdoctoral researcher in the chemistry department at University of Massachusetts Amherst with Lynmarie

Thompson using REDOR to study structure/function relationships in the serine receptor of *E. coli*. He currently teaches biochemistry and does research on protein structure and function at University of Nebraska at Kearney.



Daniel J. Fowler graduated from Colgate University in 2002 with a B.A. in chemistry and creative writing. He is currently pursuing graduate studies at the University of Massachusetts Amherst with Lynmarie Thompson. His scientific interests are broad, but have focused on the structural biology of chemotaxis and the design and application of NMR experiments.



Gregory J. Gallagher received a B.S. in chemistry and biochemistry from Worcester Polytechnic Institute. He obtained a Ph.D. in chemistry from the University of Massachusetts Amherst in 2006. He did his doctoral work with Lynmarie Thompson investigating signal transduction in a bacterial chemotaxis receptor using solid-state NMR. Currently, he is an assistant professor of chemistry at American International College in Springfield, Massachusetts, where he teaches organic chemistry and biochemistry.



Lynmarie K. Thompson earned her Ph.D. from Yale University in 1989, working with Gary Brudvig on photosynthetic proteins. Her postdoctoral work with Robert Griffin at MIT applied solid-state NMR methods to the study of bacteriorhodopsin. In 1991 she moved to University of Massachusetts Amherst. Her research focuses on mechanistic studies of membrane proteins, using a combination of solid-state NMR and other biophysical and biochemical methods to investigate fundamental processes such as transmembrane signaling. She is a faculty member in the Department of Chemistry, the Department of Biochemistry and Molecular Biology, and the Graduate Program in Molecular and Cellular Biology. She also serves as director of the Chemistry-Biology Interface Training Program.