Inter-Tryptophan Distances in Rat Cellular Retinol Binding Protein II by Solid-State NMR†

Lynda M. McDowell,‡ Susan M. Holl,‡ Shi-jun Qian,§ Ellen Li,§ and Jacob Schaefer*,‡

Department of Chemistry, Washington University, St. Louis, Missouri 63130, and Department of Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

Received October 13, 1992; Revised Manuscript Received February 11, 1993

ABSTRACT: Structural constraints for the tryptophans in rat cellular retinol binding protein II (CRBP II) have been obtained by rotational-echo double-resonance (REDOR) solid-state NMR. CRBP II was labeled with L-[6-^{19}F]tryptophan and L-[2-^{13}C]tryptophan. The ^{13}C–^{19}F dipolar coupling was determined for various possible tryptophan geometries. The allowed distance between the closest two of the four tryptophans in CRBP II was obtained for each geometry. The minimum possible distance between these two tryptophans in CRBP II is 7 Å, and the maximum possible distance is 11 Å.

Rotational-echo double-resonance (REDOR)[1] is a new, magic-angle spinning, solid-state NMR technique (Gullion & Schaefer, 1989a,b) that is suitable for the analysis of the structure of large, noncrystallizeable proteins. REDOR involves measuring the heteronuclear dipolar coupling between two stable-isotope labels. Magnetization on one rare-spin radiofrequency channel is dephased by rotor-synchronized pulses on the other. Interactions with protons are suppressed by dipolar decoupling. The extent of dephasing has a simple relation to the strength of the dipolar coupling and hence the internuclear distance.

All applications of REDOR to protein structure determinations so far have used specific rare-spin labels and specific dipolar couplings. For example, REDOR has been used to measure the interaction between a single ^{13}C label of an inhibitor in the binding site of an enzyme and a ^{31}P label in the same inhibitor (Christensen & Schaefer, 1993). The C–P coupling leads to a determination of the internuclear distance and so becomes a measure of the conformational change induced in the inhibitor on binding. REDOR has also been used to measure the coupling between a single ^{31}P of a negatively charged phosphonate group of a bound inhibitor and the ϵ-^{15}N of a nearby positively charged lysyl residue (Schmidt et al., 1993). Although the ^{15}N was introduced by nonspecific labeling and the ϵ-nitrogens of all lysines of the protein are labeled, the N–P interaction is specific if there is only a single lysyl residue within, say, 5 Å of the phosphonate ^{31}P. The coupling to more distant lysines is weak and can be neglected.

Specific dipolar couplings arise in a natural way for protein complexes because the binding site is unique. The ability of REDOR to determine precise distances in proteins and peptides with great accuracy from specific dipolar couplings is well documented (Marshall et al., 1990; Holl et al., 1992; Christensen & Schaefer, 1993). A more difficult application of REDOR is to the analysis of protein structure when there is no unique complex and there are no isolated specific couplings. In this paper, we examine such a situation for ^{19}F and ^{13}C labels in all four of the tryptophan residues of rat cellular retinol binding protein II (CRBP II).

Rat cellular retinol-binding protein II (CRBP II) is an abundant intestinal protein that binds *all-trans*-retinol and *all-trans*-retinal. The observation that CRBP II is found in high concentrations exclusively in the small intestinal villus enterocytes (MacDonald & Ong, 1987) suggested that CRBP II is uniquely adapted for the intestinal uptake and metabolism of retinol. Using an efficient prokaryotic expression vector, we have overexpressed rat CRBP II in *Escherichia coli*, in order to generate large amounts of the apoprotein for structural analysis (Li et al., 1987). The *E. coli*-derived CRBP II has ligand-binding properties that are indistinguishable from that of CRBP II isolated from rat intestine (Li et al., 1987).

The structure of CRBP II has been probed by incorporating fluorine at the 6 position of the four Trps at positions 9, 89, 107, and 110 (Li et al., 1989, 1990, 1991). Fluorine NMR studies of 6-fluorotryptophan (L-[6-^{19}F]Trp) labeled CRBP II in solution demonstrated ligand-induced changes in the chemical shifts of two of the four Trp residues (9 and 107). Definitive resonance assignments were made by analysis of four Trp substitution mutants (Li et al., 1990). These experiments were performed with the expectation that the signal corresponding to the substituted Trp residue would be missing and that the resonances corresponding to the remaining three Trp residues would undergo minimal changes in chemical shifts. The perturbations appeared to be relatively small for the three of the four mutant proteins. However, replacement of Trp_{110} by phenylalanine resulted in a large downfield change in the chemical shift of the Trp_9 signal. This perturbation in the Trp_9 signal raised the possibility that the Trp_{110} and Trp_9 residues were in proximity. Aternatively, the perturbation could reflect a local change in conformation resulting from this mutation.

Molecular modeling of the three-dimensional structure of CRBP II was performed (Li et al., 1989) using the coordinates of a homologous protein, intestinal fatty-acid-binding protein, to predict the positions of the four Trp residues of CRBP II. [*E. coli*-derived CRBP II has been crystallized (Sacchettini et al., 1988), but the X-ray structure analysis is still in progress.] The model predicted that CRBP II would consist of two nearly orthogonal β-sheets formed by 10 antiparallel β-strands. The four Trp residues in this model were predicted

† This work was supported by the National Institutes of Health (GM-40634, J.S.; DK-40172 and Research Career Development Award DK-02072, E.L.). S.M.H. thanks IBM for support through a predoctoral fellowship.

‡ Department of Chemistry.

§ Department of Medicine.

[1] Abbreviations: CRBP II, rat cellular retinol binding protein II; REDOR rotational-echo double resonance; CP, cross-polarization; MeA, methylalanine; Bzl, benzyl; N_c, number of rotor cycles.

0006-2960/93/0432-4560$04.00/0

to be roughly equidistant and approximately 10 Å apart. However, inspection of this model indicated that Trp_{110} would be exposed to solvent. (The model is *in vacuo*.) Twisting of the H β-strand so that Trp_{110} was less exposed to solvent appeared to place Trp_{110} in closer proximity to Trp_9 (5 Å).

To investigate the inter-tryptophan distances, we labeled CRBP II derived from *E. coli* both with L-[6-^{19}F]Trp and L-[2-^{13}C]Trp. We show that even though specific C–F distances cannot be determined, REDOR provides enough information to conclude that (i) a mutant CRBP II protein has a conformation different from that of the wild-type protein and (ii) a CRBP II structure which has two tryptophan residues within 5 Å of one another is disallowed. REDOR constraints such as these may prove to be generally useful in guiding modeling or X-ray studies of protein tertiary structure.

MATERIALS AND METHODS

L-[6-^{19}F]Trp was purchased from Sigma. L-[2-^{13}C]Trp (99%) was purchased from Cambridge Isotopes. The labeled CRBP II was obtained by growing a CRBP II expressing *E. coli* tryptophan auxotroph in the presence of 0.05 mM L-[2-^{13}C]Trp and 0.05 mM L-[6-^{19}F]Trp (Li et al., 1989). CRBP II (approximately 1 mM) was dialyzed against distilled water and lyophilized for solid-state NMR. The sample was packed into a 7-mm high-performance zirconia rotor and fitted with plastic (Kel-F) endcaps and spacers. The sample weight was about 80 mg.

Experiments were run on a home-built spectrometer operating with a 4.7-T wide-bore Oxford magnet (Oxford, England). The four-radiofrequency (RF), quadruple-tuned probe makes use of a coaxial transmission line connecting tuning components to a single, seven-turn, 9-mm solenoidal coil. Only the coil is in the magnet (Holl et al., 1990). The spectrometer uses a Chemagnetics (Fort Collins, CO) spinning system, ENI (Rochester, NY) rf transmitters for frequencies below 100 MHz, and Kalmus (Woodinville, OR) transmitters for the higher radiofrequencies. The pulse generator and acquisition system are from Tecmag (Houston, TX). Specially designed circuits control the rf pulse amplitude and magic-angle-spinning speed.

REDOR utilizes magic-angle sample spinning (MAS) to produce high-resolution spectra and measures the heteronuclear dipolar coupling (D_{IS}) between isolated pairs of labeled nuclei, I and S. The basic experiment consists of preparation of transverse S-spin magnetization by cross-polarization (CP) of the S spins from the abundant proton reservoir, followed by a period of I–S dipolar evolution that reintroduces the weak heteronuclear dipolar coupling removed by MAS, and S-spin signal detection. The dipolar evolution period contains two sets of rotor-synchronized, interleaved pulse trains; the first set consists of I-spin π pulses in the middle of each rotor cycle, and the second set consists of an S-spin π pulse at the end of each rotor cycle. (Placing the dephasing pulses at half-rotor period intervals maximizes the dephasing during the dipolar evolution time.) REDOR actually requires two spectra to be collected, one with the pulses on the I channel to produce the signal S and one without to produce the signal S_0. In a powder, the ratio of the difference between the two spectra ($\Delta S = S_0 - S$) and S_0 can be directly related to D_{IS} (Pan et al., 1990). Given D_{IS}, the internuclear distance (r_{IS}) can be easily calculated:

$$r_{IS} = [(\gamma_I \gamma_S h)/(2\pi D_{IS})]^{1/3}$$

where γ_I and γ_S are the gyromagnetic ratios of the I- and S-spins, respectively, and h is Planck's constant.

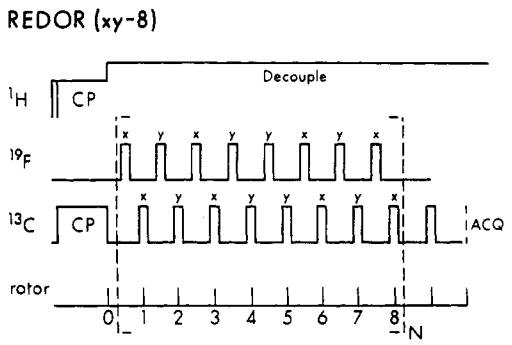

FIGURE 1: Pulse sequence for the solid-state, magic-angle-spinning experiment rotational-echo double-resonance (REDOR). The phases of the pulses follow the *xy*-8 pulse alternation scheme. The REDOR experiment consists of two separate acquisitions: one using the sequence as drawn to produce the signal S, and the other using the sequence without the ^{19}F pulses to produce the full echo spectrum S_0. The difference between these two signals is ΔS. The ratio of ΔS and S_0 allows the calculation of a ^{19}F–^{13}C internuclear distance for an isolated pair of spins.

The REDOR experiment was performed with *xy*-8 phase cycling (Gullion et al., 1990) of the π pulses on both channels. Pulses on the rotor period were on the carbon channel, while those on the half-rotor period were on the fluorine channel (Figure 1). The spectrometer operates at 199.951 MHz for ^{1}H, 188.122 MHz for ^{19}F, 50.283 MHz for ^{13}C, and 20.268 MHz for ^{15}N. The ^{15}N channel was not used in the experiments described here. RF field amplitudes of 38 kHz (^{1}H CP, ^{13}C CP and refocusing pulses) and 50 kHz (^{19}F dephasing pulses) were used, with ^{1}H decoupling at 80 kHz. Magic-angle spinning was at 5 kHz. The CP time was 2 ms, which guarantees representative intensities for all carbons (Jacob et al., 1985), and the experiment was repeated after a 1-s delay. Scans alternated between dephased (S) and full echo (S_0) spectra. Measurements were made with both 40 and 80 rotor cycles of dephasing.

Solids are not truly rigid at room temperature. Because molecular motion partially averages the dipolar coupling, measured couplings are somewhat less than the rigid-lattice values. To correct for the motion in CRBP II, we compared our measured D_{CF} with the X-ray determined ^{13}C-^{15}N internuclear distance in [2-^{13}C, ^{15}N]alanine of 1.49 Å (r_{CN}). The experimentally measured dipolar coupling (D_{CN}) is 900 Hz (Pan et al., 1990), somewhat less than the calculated rigid-lattice value of 921 Hz. Using a D_{CN} corrected for molecular motion, the ^{13}C–^{19}F internuclear distance (r_{CF}) is given by

$$r_{CF} = (\gamma_F D_{CN}/\gamma_N D_{CF})^{1/3} r_{CN}$$

Cumulative pulse imperfections and rotor instability which lead to less efficient dephasing and refocusing than is expected can cause errors in distance measurements. These errors should be insignificant when the number of rotor cycles is small and a phase compensated pulse sequence (Figure 1) is used, but they can accumulate nonlinearly with increasing dephasing. We calibrated REDOR dephasing after 40 and 80 rotor cycles to dephasing after 20 rotor cycles using a triple-labeled, nine-residue fragment of the peptide antibiotic emerimicin: ^{19}FCH$_2$CO-Phe-MeA-MeA-[1-^{13}C]MeA-[^{15}N]-Val-Gly-Leu-MeA-MeA-OBzl. The ^{13}C–^{19}F distance determined by solid-state NMR after 20 and 40 rotor cycles is in good agreement with the X-ray crystal structure result (Holl et al., 1992). After 80 rotor cycles, the C-F REDOR-determined distance is too long by 12%, corresponding to a 40% underestimate of $\Delta S/S_0$.

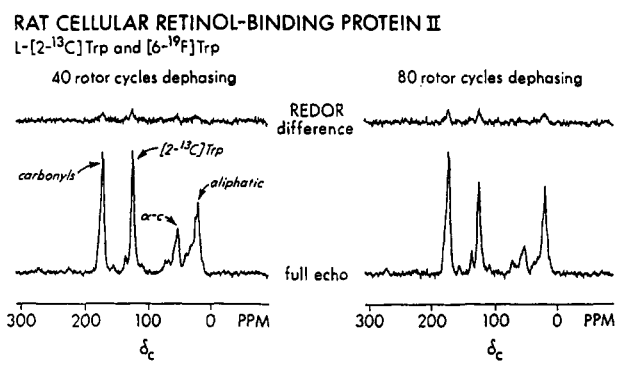

FIGURE 2: Rotational-echo, double-resonance (REDOR) ^{13}C NMR spectra of rat cellular retinol-binding protein II. The spectra correspond to 40 rotor cycles of dephasing (left) and 80 rotor cycles of dephasing (right). The REDOR difference spectra are plotted above their full-echo spectra. The full echo spectra are the average of 124 040 scans (40 rotor cycles) and 285 300 scans (80 rotor cycles).

RESULTS

Ignoring the minor effect arising from the difference in CP rates (Jacob et al., 1985), the relative intensities of the carbonyl-carbon peak (175 ppm) and the L-[2-^{13}C]Trp peak (124 ppm) after a 2-ms cross-polarization transfer in a two rotor-cycle spin-echo spectrum can be used to estimate the fraction of L-[2-^{13}C]Trp incorporation. The ratio of intensities of the two peaks is given by

$$\frac{[2\text{-}^{13}C]\text{Trp intensity}}{\text{carbonyl intensity}} = \frac{(\text{number of Trp} \times \text{average Trp label fraction})}{(\text{number of carbonyls} \times \text{carbonyl label fraction})}$$

There are 175 backbone and side-chain carbonyl carbons in CRBP II (Li et al., 1989) at 1.1% natural abundance ^{13}C (label fraction 0.011). There are four Trps, the L-[2-^{13}C]-Trps incorporated with 99% label and the L-[6-^{19}F]Trps with 1% natural-abundance ^{13}C. The experimental L-[2-^{13}C]Trp to carbonyl intensity ratio is 1.2 (data not shown). The calculated average label fraction for L-[2-^{13}C]Trp is therefore 0.57 (±20% estimated error), which implies that the ^{13}C and ^{19}F labels have been distributed close to uniformly. Efficient incorporation of L-[6-^{19}F]Trp has been previously demonstrated in this cell line (Li et al., 1989). Therefore, in the following discussion we assume that L-[2-^{13}C]Trp and L-[6-^{19}F]Trp are equally represented in this CRBP II sample and that any nonlabeled Trp contribution is negligible.

The observed $\Delta S/S_0$ at 124 ppm after 40 rotor cycles of dephasing is 0.10, and after 80 rotor cycles it is 0.16 (Figure 2). There is no significant background signal at 124 ppm, so the observed full-echo signal is S_0 (Gullion & Schaefer, 1989). No correction to $\Delta S/S_0$ was needed for the 40 rotor cycle experiment, but a multiplicative correction factor of 1.4 was used for the 80 rotor cycle experiment.

No evidence of large-amplitude motions which could affect the determination of internuclear distances was found. The Trps in particular are not likely to be part of a dynamically active structural region because the amplitudes of their rotational spin echoes were not greatly reduced with respect to those of the more rigid carbonyls on the backbone. In fact, after 8 ms, 50% of the original CP amplitude of the L-[2-^{13}C]Trp peak remained.

DISCUSSION

The presence of four labeled Trps, which are not resolved in the carbon NMR spectrum, and the combination of ^{13}C and ^{19}F labels in the CRBP II sample make the translation of $\Delta S/S_0$ into a unique distance between a specific pair of Trps impossible. We can, however, determine the allowed range of the distances between the 2-C and 6-F of the closest pair of Trps. To do this, we describe Trp geometries and distances which are allowed or disallowed on the basis of the NMR data.

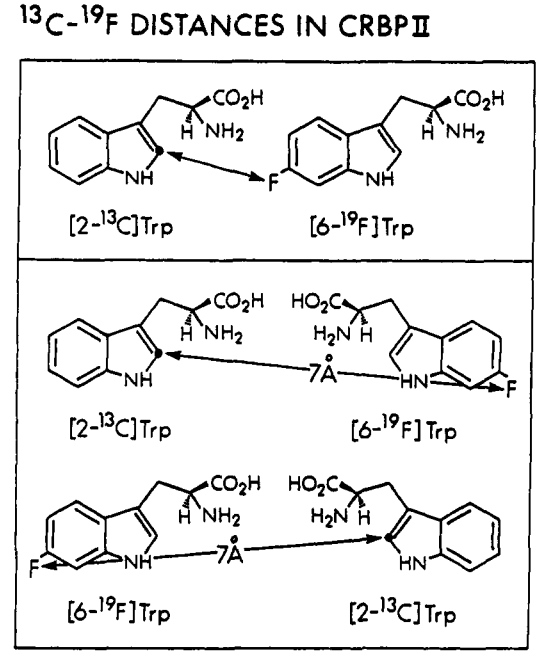

FIGURE 3: Models of labeled tryptophan geometry used to interpret the REDOR data. Although the tryptophans are coplanar in this figure, the analysis is valid for any spatial arrangement. Both panels represent the hypothesis that only two of the four Trp in CRBP II contribute to the REDOR difference signal (ΔS). In the upper panel only one label combination contributes to ΔS. This model is not consistent with the REDOR data. In the lower panel, two label combinatons contribute to ΔS with an average ^{13}C–^{19}F internuclear distance of 7 Å.

The shortest distance we could obtain from our data occurs if only two of the four Trps are close enough to contribute to ΔS and their relative orientation is as shown in Figure 3 (top). Only the indicated combination of L-[2-^{13}C]Trp and L-[6-^{19}F]Trp will make significant contributions to ΔS. (The distance across each indole is about 5.5 Å, and the maximum internuclear distance observable by ^{13}C–^{19}F REDOR is about 12 Å.) However, the full echo spectrum (S_0) would have contributions from all four Trps and all but one of the 16 possible label combinations. In this case, the statistics indicate that 8 times as many carbons contribute to S_0 as to ΔS. Applying this hypothesis to the 80 rotor cycle REDOR data yields $\Delta S/S_0 = 1.79$ with the imperfect-pulse-rotor correction factor, or 1.28 without it. Neither value is possible because the maximum theoretical value of $\Delta S/S_0$ is 1.028 (Figure 4).

The next case we consider still has only two of the four Trps close enough to contribute to ΔS, but the geometry is such that we measure an average dipolar coupling which comes from two combinations of labels. One example of such a geometry is shown in Figure 3 (bottom). Now the statistics say that 4 times as many carbons contribute to S_0 as to ΔS. This hypothesis is plausible because reasonable values of $\Delta S/S_0$ (solid circles in Figure 4) are obtained for both the 40 and 80 rotor cycle experiments. The corresponding dipolar couplings are 84 Hz ($N_c = 40$) and 78 Hz ($N_c = 80$), and the theoretical curve (Figure 4) corresponds to the average D of 81 Hz. The distances calculated under this hypothesis are

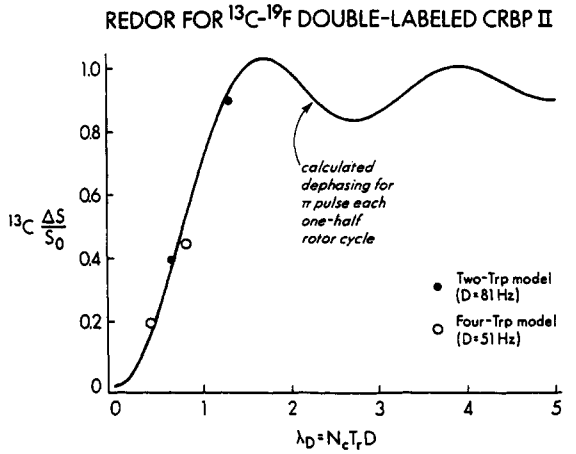

FIGURE 4: Theoretical relationship between $\Delta S/S_0$ and λ_D, the product of the number of rotor cycles, N_c, the rotor period, T_r (1/spinning speed, in seconds), and the heteronuclear dipolar coupling constant, D (in Hz). The solid circles are the experimental results of Figure 2 assuming a two-Trp model (Figure 4, bottom), and the open circles are the results assuming a four-Trp model (see the text).

6.9 Å ($N_c = 40$) and 7.1 Å ($N_c = 80$). These values represent the minimum average distance between the 2-^{13}C of one Trp and the 6-^{19}F of another Trp.

A third model has two pairs of Trp close enough to contribute to ΔS, but the individual pairs are far apart and have only weak dipolar couplings between their members. Now the label statistics indicate that twice as many carbons contribute to S_0 as to ΔS (open circles in Figure 4). If we assume that the two pairs have the same average dipolar coupling between the 2-^{13}C and the 6-^{19}F, the distances determined are 7.9 Å ($N_c = 40$) and 8.5 Å ($N_c = 80$). We tried various combinations of one short (7–8.5 Å) distance and one longer distance (9–11 Å), but none of them fit both the 40 and 80 rotor cycle data.

Finally, we determine the longest distance between the two closest Trps that is consistent with the REDOR data. This occurs when all four Trps are equidistant (an equilateral trigonal pyramid of Trps). Now only two of the 16 label combinations (all L-[2-^{13}C]Trp and all L-[6-^{19}F]Trp) fail to contribute to ΔS and the statistical multiplier is $^8/_7$. The experimental data calculated with the curve in Figure 4 under this model yield distances of 8.7 Å ($N_c = 40$) and 9.5 Å ($N_c = 80$). However, Figure 4 is calculated for an isolated spin pair and that assumption is violated in this model. Most of the label combinations have two or three fluorines in the vicinity of a single labeled carbon. Multiple fluorines can produce extra dephasing which translates into a stronger apparent dipolar coupling and thus an underestimate of the true distance. Assuming that these fluorines interact only with the carbon and not with each other, and that the fluorine dephasings of the carbon are independent of one another (Holl et al., 1990), we estimate the maximum distance between Trps to be 11 Å.

On the basis of the REDOR results, the shortest possible average distance between the 2-C and the 6-F of a pair of Trps in CRBP II is 7.0 Å, and the remaining two Trps are then constrained to be more than 12 Å from that pair and from each other. It is therefore unlikely that the perturbation in chemical shift of the Trp_9 fluorine resonance observed by solution NMR when Trp_{110} is replaced by a phenylalanine residue is simply due to the proximity of Trp_9 and Trp_{110}. It is more likely that the shift is due to a local change in protein conformation. The solid-state NMR data show that the longest possible average distance between Trps occurs if all four are at comparable distances, and then the average 2-C to 6-F distance must be between 9 and 11 Å. This result is consistent with the original CRBP II molecular model. An intermediate possibility supported by the NMR data is that two pairs of Trps have average 2-C to 6-F internuclear distances of 8 (±1) Å, and these pairs are more than 12 Å apart.

ACKNOWLEDGMENT

We gratefully acknowledge Mr. Robert McKay for assistance with the spectrometer and Dr. Robert Sacchettini for sharing insights into CRBP II modeling.

REFERENCES

Christensen, A. M., & Schaefer, J. (1993) *Biochemistry* (in press).
Crow, J. A., & Ong, D. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4707–4711.
Gullion, T., & Schaefer, J. (1989a) *Adv. Magn. Reson. 13*, 57–83.
Gullion, T., & Schaefer, J. (1989b) *J. Magn. Reson. 81*, 196–200.
Gullion, T., Baker, D. B., & Conradi, M. S. (1990) *J. Magn. Reson. 89*, 479–484.
Holl, S. M., Marshall, G. R., Beusen, D. D., Kociolek, K., Redlinski, A. S., Leplawy, M. T., McKay, R. A., & Schaefer, J. (1992) *J. Am. Chem. Soc. 114*, 4830–4833.
Holl, S. M., McKay, R. A., Gullion, T., & Schaefer, J. (1990) *J. Magn. Reson. 89*, 620–626.
Jacob, J. S., Schaefer, J., & Wilson, G. E. (1985) *J. Biol. Chem. 260*, 2777–2781.
Jones, T. A., Bergfors, T., Sedzik, J., & Unge, T. (1988) *EMBO J. 7*, 1597–1604.
Levin, M. S., Locke, B., Yang, N. C., Li, E., & Gordon, J. I. (1988) *J. Biol. Chem. 263*, 17715–17723.
Li, E., Locke, B., Yang, N. C., Ong., D. E., & Gordon, J. I. (1987) *J. Biol. Chem. 262*, 13773–13779.
Li, E., Qian, S., Nader, L., Yang, N. C., d'Avignon, A., Sacchettini, J. C., & Gordon, J. I. (1989) *J. Biol. Chem. 264*, 17041–17048.
Li, E., Qian, S., Yang, N. C., d'Avignon, A., & Gordon, J. I. (1990) *J. Biol. Chem. 265*, 11549–11554.
Locke, B. C., MacInnis, J. M., Qian, S., Gordon, J. I., Li, E., Fleming, G. R., & Yang, N. C. (1992) *Biochemistry 31*, 2376–2383.
MacDonald, P. N., & Ong, D. E. (1987) *J. Biol. Chem. 262*, 10550–10556.
Marshall, G. R., Beusen, D. D., Kociolek, K., Redlinski, A. S., Leplawy, M. T., Pan, Y., & Schaefer, J. (1990) *J. Am. Chem. Soc. 112*, 963–966.
Pan, Y., Gullion, T., & Schaefer, J. (1990) *J. Magn. Reson. 90*, 330–340.
Sacchettini, J. C., Gordon, J. I., & Banaszak, L. J. (1988) *J. Biol. Chem. 263*, 5815–5819.
Schmidt, A., Cohen, E. C., McDowell, L. M., Beusen, D. D., & Schaefer, J. (1993) 34th Experimental Nuclear Magnetic Resonance Conference, St. Louis, MO.