Journal of Biomolecular NMR, 5 (1995) 415–419
ESCOM

J-Bio NMR 293

Probing internal water molecules in proteins using two-dimensional ^{19}F-^{1}H NMR

David P. Cistola* and Kathleen B. Hall*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8231, St. Louis, MO 63110-1093, U.S.A.

Received 16 March 1995
Accepted 8 May 1995

Keywords: Bound water; Proteins; Fatty acid-binding protein; HOESY NMR

Summary

A simple approach for detecting internal water molecules in proteins in solution is described. This approach combines ^{19}F-detected heteronuclear Overhauser and exchange spectroscopy (HOESY) with site-specific ^{19}F substitution. The model system employed was intestinal fatty acid-binding protein complexed with [2-mono-^{19}F]-palmitate. An intense cross peak was observed between the fluorine and a buried water molecule, as defined in the 1.98 Å crystal structure of the complex. From HOESY spectra, the fluorine–water distance was estimated to be 2.1 Å, in agreement with the crystal structure. This approach may be applicable to macromolecules that are too large for ^{1}H-detected NMR methods.

Individual water molecules buried within cavities in proteins are thought to serve important roles in proton transport, enzymatic catalysis, ligand binding and protein stability (Edsall and McKenzie, 1983; Meyer, 1992; Berndt et al., 1993). These internal water molecules have been identified most often in crystal structures, using high-resolution X-ray and neutron diffraction methods (Cheng and Schoenborn, 1991; Williams et al., 1994). For proteins in solution, NMR methods have recently been developed to characterize the locations and exchange rates of internal waters (Otting and Wüthrich, 1989; Clore et al., 1990,1994; Forman-Kay et al., 1991; Otting et al., 1991a,b; Grzesiek and Bax, 1993; Xu et al., 1993). These ^{1}H-detected NMR methods require special solvent suppression schemes to avoid obliteration of the resonances of interest and somewhat complicated 3D pulse sequences to achieve adequate spectral resolution for all but the smallest proteins. In principle, these difficulties may be circumvented if the bound waters could be probed indirectly by detecting rare nuclei other than ^{1}H. In this report, we present a simple approach for probing bound water molecules using heteronuclear Overhauser and exchange spectroscopy (HOESY) with ^{19}F detection.

As a model system to demonstrate this method, we have used rat intestinal fatty acid-binding protein (I-FABP) complexed with palmitate. This soluble 15.4 kDa complex is well characterized, both structurally and energetically (Sacchettini et al., 1989; Jakoby et al., 1993), and sequence-specific ^{1}H, ^{13}C and ^{15}N assignments are now available (Hodsdon et al., 1995). Its backbone fold exhibits a β-clam topology comprising two five-stranded antiparallel β-sheets and two short α-helices. The single molecule of bound fatty acid is buried in an interior cavity, sequestered from bulk solvent. This cavity, with a volume of 850 Å^3, also contains at least seven ordered water molecules, as observed by X-ray crystallography (Sacchettini et al., 1989). The most nonpolar portion of the cavity is adjacent to the C-2 to C-4 positions of the fatty acid, and one of the bound water molecules is directly adjacent to the C-2 position. We employed [2-mono-^{19}F]-palmitate in a ^{19}F-detected HOESY experiment in order to probe the bound water in this buried region of the complex.

Two-dimensional HOESY spectra for [2-mono-^{19}F]-palmitate (Matreya, Inc., Pleasant Gap, PA), bound to I-FABP in a 1:1 stoichiometry (Cistola et al., 1989), are shown in Fig. 1. The single ^{19}F atom in the complex gives rise to two well-resolved ^{19}F resonances at −103.7 and −105.6 ppm. These peaks have an intensity ratio of 1:1 and represent the *R*- and *S*-enantiomers of [2-^{19}F]-palmi-

*To whom correspondence should be addressed.

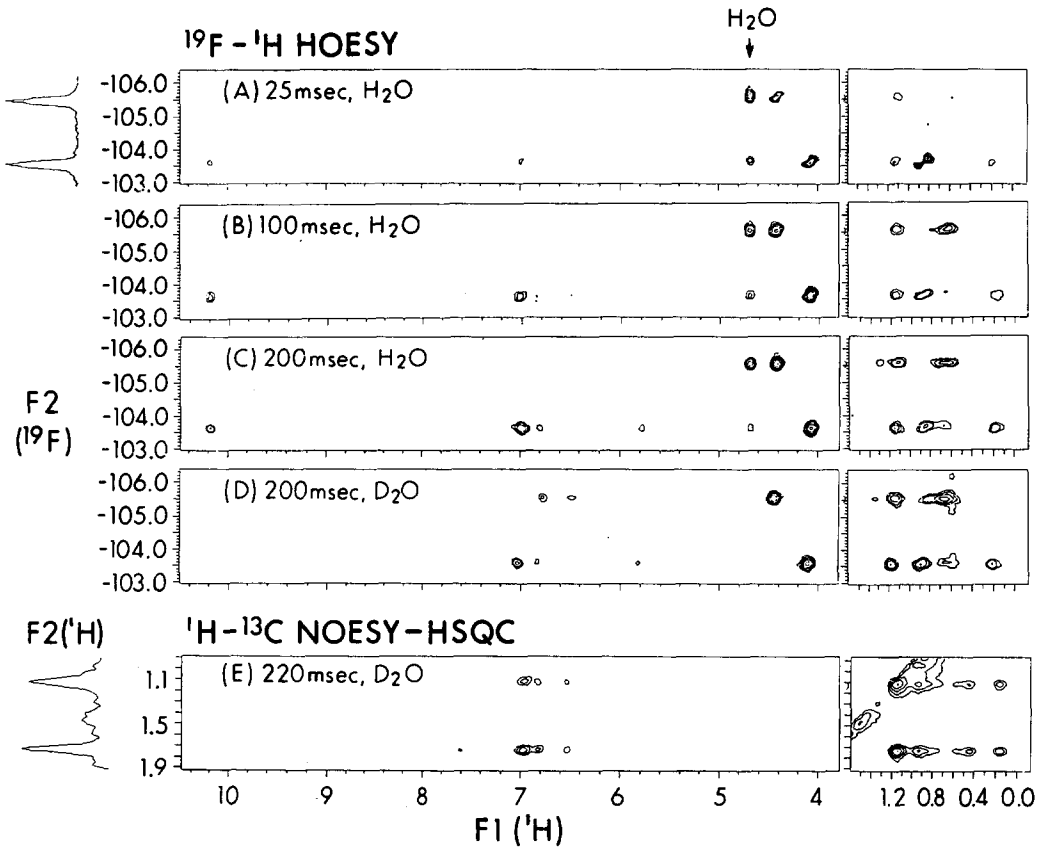

Fig. 1. (A–D) Two-dimensional ^{19}F-^{1}H HOESY spectra for a 1 : 1 complex of [2-mono-^{19}F]-palmitate with I-FABP; (E) two-dimensional ^{1}H-^{13}C NOESY-HSQC spectrum for a 1 : 1 complex of [2-^{13}C]-palmitate with I-FABP. All spectra were recorded at 25 °C and pH 7.2. A one-dimensional ^{19}F spectrum is displayed to the left of panel A. In panels A–C, HOESY spectra were accumulated at different mixing times in an H_2O-based buffer. Subsequently, the sample was lyophilized and rehydrated with 99.996% D_2O (panel D). The protein concentrations were 1.5 mM (A–D) and 2.8 mM (E). Spectra were accumulated using a three-channel Varian Unity-500 spectrometer equipped with two high-frequency amplifiers and a Nalorac 5 mm dual-coil HF500-5 probe. The HOESY pulse sequence employed was similar to that of Rinaldi (1983), except that no field gradient pulses were used. The HOESY spectra were acquired with spectral widths of 4000 Hz (F2) and 5997.9 Hz (F1) with a relaxation delay of 1.0 s (4–5 times T_1 for ^{19}F-palmitate), and were processed with 512 points zero-filled to 2048 (F2) and 256 points zero-filled to 1024 (F1) with exponential multiplication with a line broadening of 50 Hz (F2) and a Gaussian with a time constant of 0.017 (F1). The ^{19}F chemical shifts were referenced to external trifluoroacetic acid and ^{1}H shifts to external sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d_4 (TSP) in D_2O, all at 25 °C. The sample buffer was 20 mM potassium phosphate, 50 mM KCl and 0.05% NaN_3.

tate, respectively, both bound to I-FABP.* Because of the hydrogen-bonding and ion-pair network involving the fatty acid carboxyl group and Arg^{106} (Cistola et al., 1989; Sacchettini et al., 1989; Jakoby et al., 1993), the carboxyl end of the bound fatty acid has a fixed orientation in the binding cavity. The ^{19}F atoms in the *R*- and *S*-enantiomers experience different local environments and therefore have different chemical shifts. A ^{1}H-^{13}C NOESY-HSQC spectrum of I-FABP complexed with [2-^{13}C]-palmitate revealed a similar pattern of two well-resolved methylene protons at 1.1 and 1.8 ppm (Fig. 1E). The distinct environments of the two substituents of C-2 are thus not an artefact of the asymmetric ^{19}F substitution.

The x-axis of the HOESY spectra in Fig. 1, panels A–D, identifies those protons that are dipolar coupled (through space) to the fatty acid [2-^{19}F] nucleus. In practice, cross peaks are observed for protons that are within approximately 4–5 Å of the fluorine. Hence, these spectra provide a map of the nearest neighbors of the 2-position of the bound fatty acid.

To evaluate the HOESY maps, we have utilized the 1.98 Å crystal structure of I-FABP complexed with palmitate (Sacchettini et al., 1989; Brookhaven Protein Data Bank, file 2ifb). From this structure, which lacks fluorine, two models were generated by simply substituting a fluorine atom for each of the two methylene protons at the 2-

*The possibility that one of these resonances represents unbound palmitate was ruled out based on the following. The protein concentration used for NMR (1.5 mM) was more than 1000-fold greater than the dissociation constant for fatty acid binding (K_D (palmitate) = 2×10^{-7} M; Jakoby et al., 1993). Therefore, > 99.9% of the fatty acid was expected to be bound under these conditions. In addition, unbound palmitate forms a 1:1 acid–soap crystalline phase at this pH and temperature (Cistola et al., 1986,1988). This crystalline phase gives rise to extremely broad and undetectable NMR resonances.

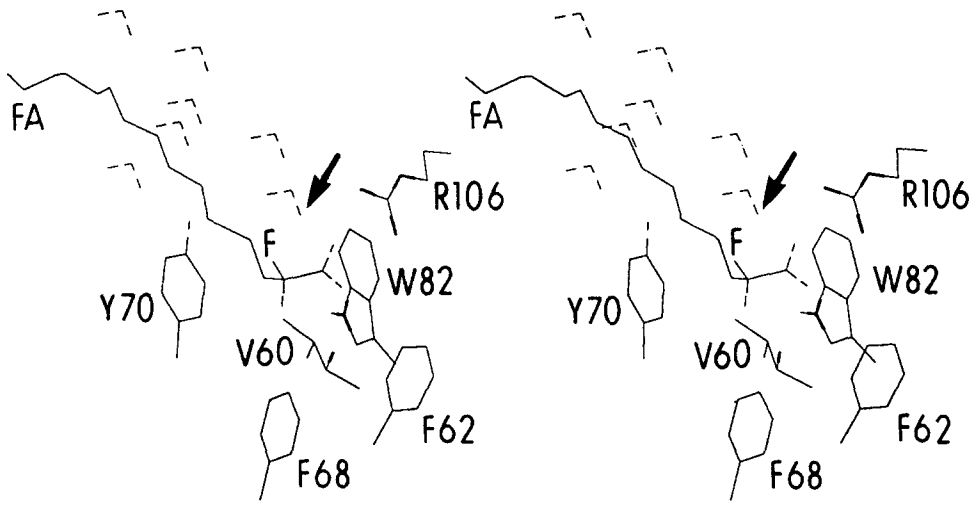

Fig. 2. Stereo diagram of the binding cavity of I-FABP containing *S*-[2-mono-^{19}F]-palmitate, based on the 1.98 Å crystal structure (Sacchettini et al., 1989). The water molecule closest to the fluorine is designated by an arrow. For clarity, only protons at the ring NH of Trp^{82}, the 2-position of the fatty acid and the waters are shown.

position of the ligand. A stereo diagram of the binding cavity containing the *S*-enantiomer of [2-^{19}F]-palmitate is shown in Fig. 2. The long axis of the fatty acid has two faces: a polar face, adjacent to seven ordered water molecules, and a nonpolar face, adjacent to nonpolar aromatic and aliphatic side chains. The ^{19}F atom in the *S*-enantiomer points toward the polar face and, notably, toward an interior water molecule. The distance between this fluorine and the closest water proton* (arrow in Fig. 2) is 2.1 Å. The distances between the fluorine atoms and their nearest neighbors, as determined from the structural models, are summarized in Table 1.

The HOESY cross peaks were assigned (Table 1), and several important observations were made. First, the cross peaks at a proton chemical shift of 4.75 ppm represent bound water. Their identity as water, rather than degenerate backbone α-protons from the protein, was confirmed by observing their selective disappearance in a HOESY spectrum collected in D_2O-based buffer (Fig. 1D). The other nearby exchangeable proton, the ring NH of Trp^{82} at 10.21 ppm proton chemical shift, was also absent in the D_2O spectrum. Second, the intensities of the water cross peaks at fluorine chemical shifts of −103.7 and −105.6 ppm were not equal; the latter was more intense at all mixing times examined, as can be seen in Figs. 1 and 3. This difference in cross-peak intensity implied that the distances between the two ^{19}F atoms and the nearest water molecule(s) were not equal. This conclusion is consistent with the X-ray structure. Finally, the patterns of nearest neighbor cross peaks observed in the ^{19}F-^{1}H HOESY and ^{1}H-^{13}C NOESY spectra were in agreement with each other and with the crystal structure.

The spectra in Fig. 1 employed a relaxation delay of 1 s and were essentially fully relaxed with respect to the fluorine (T_1 = 0.23 s), but not with respect to the water protons (T_1 = 3.0 s). Analysis of NOE buildup curves

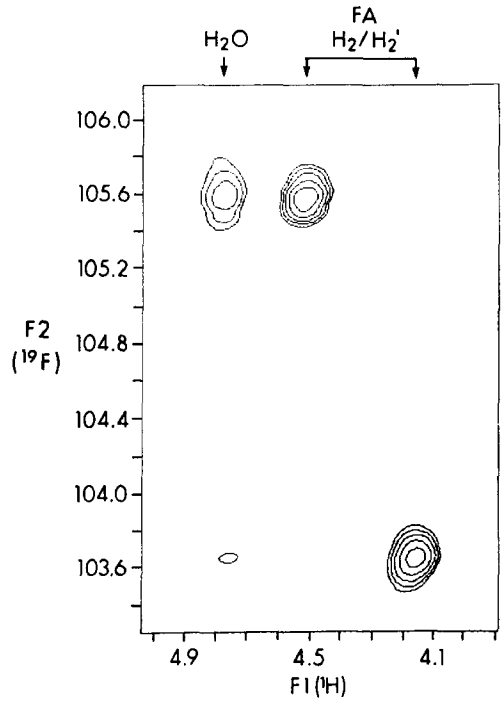

Fig. 3. Fully relaxed HOESY spectrum of a 1 : 1 complex of [2-mono-^{19}F]-palmitate with I-FABP. The sample conditions and spectral parameters were identical to those in Fig. 1, except that a 9 s relaxation delay and a narrow spectral width in F1 (2500 Hz) were employed. The spectral width was chosen so that no cross peaks folded into the region of interest, between 4 and 5 ppm, as shown in the figure.

*The positions of the bound water oxygens, not the protons, have been experimentally defined by X-ray crystallography. Therefore, the distance obtained from the structural model in Fig. 2 must be considered an estimate of the ^{19}F-^{1}H (water) distance.

TABLE 1
^{19}F-^{1}H NOE CROSS-PEAK ASSIGNMENTS AND DISTANCES[a]

^{1}H chemical shift (ppm)	Resonance assignment	Distance from 2-F (*R*) (Å)	Distance from 2-F (*S*) (Å)
10.21	Trp^{82}, ring NH	3.0	4.1
7.02/7.04[b]	Phe^{68}, ring 3/5H	2.6	4.2
6.97[b]	Phe^{68}, ring 4H	3.1	4.5
6.86[b]	Tyr^{70}, ring 2/6H	4.0	4.5
6.53[b]	Tyr^{70}, ring 3/5H	4.7	4.1
5.82–5.89[b]	Arg^{106}, guanidino NH	5.0	4.2
4.75	Internal water	3.5	2.1 (X-ray) 2.1 (NMR)
4.49[c]	Fatty acid H2 (*S*)	1.78	–
4.14[c]	Fatty acid H2 (*R*)	–	1.78
1.18–1.21	Val^{60}, methyl	2.1	3.0
0.95	Fatty acid H3/3'	2.3	2.6
0.85–0.88	Val^{49}, methyl	4.8	5.2
0.67–0.69	Fatty acid H4/4'	3.7	3.0
0.22–0.27	Fatty acid H3/3'	2.2/2.6	2.6/3.0

[a] Distance estimates were derived from the crystal structure model (Fig. 2), except where noted.
[b] The aromatic proton assignments are tentative and based on our initial interpretation of 2D NOESY and 3D ^{15}N-NOESY experiments, combined with sequence-specific assignments for the backbone and aliphatic segments of the protein side chains (Hodsdon et al., 1995).
[c] These assignments were established by comparing spectra of the lipid–protein complex containing [2-mono-^{19}F]- or [2,2-di-^{19}F]-palmitate; the latter lacked these cross peaks.

acquired under these conditions indicated that the intensities for the water cross peaks were artefactually high at mixing times < 75 ms. Therefore, additional spectra with mixing times of 0, 25 and 100 ms were accumulated with a 9 s relaxation delay for the purpose of quantitation and distance estimation. Since fully relaxed spectra require long accumulation times, a compromise value of 9 s was used and the F1 dimension was partially folded (see legend to Fig. 3). These spectra were considered to be essentially fully relaxed, since no NOE cross-peak intensity was observed at 0 ms mixing time.

The HOESY spectrum acquired with a mixing time of 100 ms is shown in Fig. 3. The cross peaks at 4.14 and 4.49 ppm represent ^{1}H nuclei covalently attached to the same carbon as the fluorine, in the *R*- and *S*-enantiomers, respectively. This ^{19}F-^{1}H distance is fixed at 1.78 Å by covalent bond geometry, and the intensities of these cross peaks could be used as an internal ruler to calibrate the distances for other ^{19}F-^{1}H cross peaks. Since the palmitate is rigidly bound in the I-FABP cavity, these two components share the same correlation time, and the covalent ^{1}H-^{19}F length should accurately reference their intermolecular distances. No information is available to describe the dynamics of the bound water, which could have a different correlation time. However, the fluorine–water distance, estimated by comparing cross-peak volumes*, was 2.1 Å, which is in good agreement with the distance estimated from the crystal structure model.

These data illustrate that the ^{19}F-detected HOESY experiment provides a simple, yet powerful method for probing bound water molecules in proteins. This approach, while complementary to existing NMR methods, has several key advantages. First, HOESY employs ^{19}F detection rather than ^{1}H detection. When combined with site- or residue-specific ^{19}F incorporation, the HOESY experiment yields a greatly simplified spectrum because of the substantial reduction in the number of observed peaks. This results in excellent resolution and eliminates the need for a third dimension, as in 3D ROESY-HSQC (Clore et al., 1990; Forman-Kay et al., 1991; Xu et al., 1993). Second, ^{19}F detection also eliminates the requirement for solvent suppression, a major advantage when probing bound water molecules which often have observed ^{1}H chemical shifts identical to that of bulk water. Third, the HOESY experiment is quite sensitive, and the NOE peaks between the fluorine and the water protons can be intense, even at low mixing times. The excellent sensitivity and resolution facilitates the generation of NOE buildup curves and the resulting estimation of internuclear distances. Fourth, the ^{19}F nucleus can be readily incorporated into proteins, at low cost compared with ^{13}C or ^{15}N; typical sites of incorporation include the aromatic residues phenylalanine, tyrosine or tryptophan. Also, many pharmaceutical agents that bind to proteins or nucleic acids naturally contain fluorine atoms. Notable examples include the sedative flurazepam (Dalmane), the anti-cancer agent 5-fluorouracil and its derivatives, and several non-steroidal anti-inflammatory agents: sulindac (Clinoril), flurbiprofen (Ansaid) and diflunisal (Dolobid).

*Initial ^{19}F-^{1}H (water) distance estimates were made according to expression [4] in Clore and Gronenborn (1985). Peak volumes were measured in the fully relaxed spectrum obtained with a 100 ms mixing time using the integration routine in NMR COMPASS v. 2.5 (Molecular Simulations, Inc.). A more in-depth analysis of distances using NOE buildup curves will be presented in a subsequent report.

The use of one or a few ^{19}F atoms to probe bound waters eliminates the requirement for sequence-specific ^{1}H resonance assignments, making this method potentially applicable to a wider variety of molecules. This would include larger proteins or macromolecular complexes, where the determination of sequential assignments is not feasible. Although the assignment database for I-FABP was used in the present case to identify other ligand–protein cross peaks and to validate the HOESY method, the database was not required to assign either the fluorine atoms or the water cross peaks.

A possible disadvantage of the ^{19}F-detected HOESY approach is the potential perturbation of the protein resulting from the substitution of a fluorine atom for a hydrogen in cases where the fluorine is not naturally present. Although fluorine has a van der Waals radius comparable to that of hydrogen, it is substantially more electronegative. It could lower the pKa and alter the ionization state of nearby groups. In addition, the fluorine atom, through an interaction analogous to a hydrogen bond, might attract a water molecule to a buried region of the molecule where one is not normally present. In most proteins, it is unlikely that the incorporation of one or a few ^{19}F atoms would introduce serious perturbations into the system. However, this possibility depends on the site of incorporation and must be assessed on a case-by-case basis.

In the case of I-FABP, the bound fatty acid already has a very low pKa in the absence of the fluorine atom and is normally bound as the carboxylate anion (Cistola et al., 1989; Jakoby et al., 1993). In addition, the nearest neighbor patterns observed for [2-^{13}C]- and [2-^{19}F]-palmitate were very similar (Fig. 1), indicating that the fluorinated ligand was bound in the same location and orientation as the non-fluorinated ligand. Moreover, the fluorine in the *R*-enantiomer did not attract a hydrogen-bonded water molecule to its site, since little NOE to water was observed, especially in the fully relaxed HOESY spectrum (Fig. 3). For the *S*-enantiomer, the nearby water molecule was observed in the crystal structure, which was determined in the absence of fluorine.

Finally, we note that the observation of strong NOE cross peaks between ^{19}F and water is not peculiar to I-FABP. We initially observed such peaks in HOESY spectra of U1A, an RNA-binding protein, with site-specific substitution of ^{19}F into the 3(5)-ring position of each of its four tyrosine residues. In U1A, the tyrosine rings are not buried, but are located at the solvent–protein interface. We have also observed ^{19}F-water cross peaks in [6-^{19}F-Trp]-dihydrofolate reductase, where the tryptophan residues are present in a variety of locations throughout the protein. Therefore, the HOESY method may prove to be useful for probing surface, as well as interior, bound water molecules. Surface-bound waters have been notoriously difficult to detect using ^{1}H-based NMR methods (Otting and Wüthrich, 1989; Otting et al., 1991a,b). Much more work is needed to test these possibilities and to further develop this method.

Acknowledgements

This work was supported by Grants from the National Science Foundation (MCB-9205665 to D.P.C.), National Institutes of Health (GM46318 to K.B.H.), the American Digestive Health Foundation and the Lucille P. Markey Charitable Trust (#90-47). The authors thank Dr. Changguo Tang for advice and technical assistance, James Kranz for his work on U1A, Michael Hodsdon for assistance with the assignments and Sydney Hoeltzli and Carl Frieden for providing a sample of dihydrofolate reductase. D.P.C. is a Johnson & Johnson/Merck Research Scholar of the American Digestive Health Foundation; K.B.H. is a Lucille P. Markey Scholar.

References

Berndt, K.D., Beunink, J., Schröder, W. and Wüthrich, K. (1993) *Biochemistry*, **32**, 4564–4570.

Cheng, X. and Schoenborn, B.P. (1991) *Acta Crystallogr.*, **A47**, 314–317.

Cistola, D.P., Atkinson, D., Hamilton, J.A. and Small, D.M. (1986) *Biochemistry*, **25**, 2804–2812.

Cistola, D.P., Hamilton, J.A., Jackson, D. and Small, D.M. (1988) *Biochemistry*, **27**, 1881–1888.

Cistola, D.P., Sacchettini, J.C., Banaszak, L.J., Walsh, M.T. and Gordon, J.I. (1989) *J. Biol. Chem.*, **264**, 2700–2710.

Clore, G.M. and Gronenborn, A.M. (1985) *J. Magn. Reson.*, **61**, 158–164.

Clore, G.M., Bax, A., Wingfield, P. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 5671–5676.

Clore, G.M., Bax, A., Omichinski, J.G. and Gronenborn, A.M. (1994) *Structure*, **2**, 89–94.

Edsall, J.T. and McKenzie, H.A. (1983) *Adv. Biophys.*, **16**, 53–183.

Forman-Kay, J.D., Gronenborn, A.M., Wingfield, P.T. and Clore, G.M. (1991) *J. Mol. Biol.*, **220**, 209–216.

Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 627–638.

Hodsdon, M.E., Toner, J.J. and Cistola, D.P. (1995) *J. Biomol. NMR*, in press.

Jakoby, M.G., Miller, K.R., Toner, J.J., Bauman, A., Cheng, L., Li, E. and Cistola, D.P. (1993) *Biochemistry*, **32**, 872–878.

Meyer, E. (1992) *Protein Sci.*, **1**, 1543–1562.

Otting, G. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 1871–1875.

Otting, G., Liepinsh, E. and Wüthrich, K. (1991a) *J. Am. Chem. Soc.*, **113**, 4363–4364.

Otting, G., Liepinsh, E. and Wüthrich, K. (1991b) *Science*, **254**, 974–980.

Rinaldi, P.L. (1983) *J. Am. Chem. Soc.*, **105**, 5167–5168.

Sacchettini, J.C., Gordon, J.I. and Banaszak, L.J. (1989) *J. Mol. Biol.*, **208**, 327–339.

Williams, M.A., Goodfellow, J.M. and Thornton, J.M. (1994) *Protein Sci.*, **3**, 1224–1235.

Xu, R.X., Meadows, R.P. and Fesik, S.W. (1993) *Biochemistry*, **32**, 2473–2480.