Journal of Biomolecular NMR, **14:** 79–83, 1999.
KLUWER/ESCOM

Application of amino acid type-specific ^{1}H- and ^{14}N-labeling in a ^{2}H-, ^{15}N-labeled background to a 47 kDa homodimer: Potential for NMR structure determination of large proteins

Mark J.S. Kelly[a,*], Cornelia Krieger[b], Linda J. Ball[a], Yihua Yu[c,**], Gerald Richter[b], Peter Schmieder[a], Adelbert Bacher[b] & Hartmut Oschkinat[a,c,*]
[a]*Forschungsinstitut für Molekulare Pharmakologie, Alfred-Kowalke-Strasse 4, D-10315 Berlin, Germany;* [b]*Lehrstuhl für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany;* [c]*EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany*

Received 4 January 1999; Accepted 3 March 1999

Key words: amino acid specific labeling, 3D heteronuclear NMR, deuteration, heteronuclear half-filter

Abstract

NMR investigations of larger macromolecules (>20 kDa) are severely hindered by rapid ^{1}H and ^{13}C transverse relaxation. Replacement of non-exchangeable protons with deuterium removes many efficient ^{1}H-^{1}H and ^{1}H-^{13}C relaxation pathways. The main disadvantage of deuteration is that many of the protons which would normally be the source of NOE-based distance restraints are removed. We report the development of a novel labeling strategy which is based on specific protonation and ^{14}N-labeling of the residues phenylalanine, tyrosine, threonine, isoleucine and valine in a fully deuterated, ^{15}N-labeled background. This allows the application of heteronuclear half-filters, ^{15}N-editing and ^{1}H-TOCSY experiments to select for particular magnetization transfer pathways. Results from investigations of a 47 kDa dimeric protein labeled in this way demonstrated that the method provides useful information for the structure determination of large proteins.

With the development of uniform ^{13}C- and ^{15}N-labeling and multi-dimensional heteronuclear NMR experiments, the structures of proteins with molecular weights of up to 20 kDa can now be routinely determined (Wagner, 1993; Clore et al., 1994). However, NMR investigations of larger macromolecules (>20 kDa) are severely hindered by rapid ^{1}H and ^{13}C transverse relaxation (LeMaster, 1990; Grzesiek et al., 1992; Wagner, 1993; Kay and Gardner, 1997). Substitution of non-exchangeable protons for deuterons allows the removal of many of these efficient ^{1}H-^{1}H and ^{1}H-^{13}C relaxation pathways, thereby improving the sensitivity and resolution which can be obtained (LeMaster, 1990; Kay and Gardner, 1997). Systematic substitution allows particular sets of protons to be retained in a fully deuterated background so that NOEs between only selected groups of protons or residues may be observed in the absence of overlap and where the relaxation properties are improved (LeMaster, 1990). Amino acid-type specific labeling in a fully deuterated background was effective in elucidating the structure of the 25 kDa tryptophan-repressor from *Escherichia coli* (Arrowsmith et al., 1990; Zhang et al., 1994). By combining deuteration with selective ^{1}H-, ^{13}C-, ^{15}N-labeling of only a limited number of amino acid types, a sufficient number of NOEs can be identified to determine the global folds of large proteins (Metzler et al., 1996; Smith et al., 1996). Similar results have also been achieved using an elegant technique where only the methyl groups of selected residue types are ^{1}H-labeled (Rosen et al., 1996; Gardner and Kay, 1997; Gardner et al., 1997).

*To whom correspondence should be addressed. E-mail: oschkinat@fmp-berlin.de; kelly@fmp-berlin.de
**Present address: Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 354 Fenglin Lu, Shanghai 200032, China.

During the resonance assignment and structure determination of 3,4-dihydroxy-2-butanone 4-phosphate synthase (Richter et al., 1992), a 47 kDa homodimer with 217 amino acid residues per subunit (employing ^{13}C-/^{15}N-labeled samples with either complete or random fractional deuteration: Grzesiek et al., 1993; Yamazaki et al., 1994a, b; Nietlispach et al., 1996; Shan et al., 1996), we encountered two specific problems which presented significant obstacles to the structure determination. Firstly, the spectra displayed severe overlap. Secondly, the linewidths of many proton resonances, in particular those of Hα protons, were large due to the strong dipolar relaxation by ^{13}C, as previously observed for large proteins (Grzesiek et al., 1992). We therefore developed a method to aid various aspects of the structure determination process (see below) and to provide specific information about several residue types which are often involved in the hydrophobic cores of proteins. The method is based on ^{1}H- and ^{14}N-labeling of the residues Phe, Tyr, Thr, Ile and Val in a fully deuterated and ^{15}N-labeled background, i.e. [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X. The main advantage of this type of isotope distribution is the potential to make use of heteronuclear ^{14}N/^{15}N-half-filters (Otting and Wüthrich, 1990), 3D ^{15}N-separated NOESY and ^{1}H-TOCSY experiments to yield spectra with dramatically reduced overlap, for large proteins (30–50 kDa), where ^{13}C-labeling would lead to problematic proton transverse-relaxation rates for many residues. The application of ^{14}N- or ^{15}N-(ω_2)-half-filters to select for NOEs to particular residue types in NOESY spectra is illustrated schematically for amino acids in two neighboring anti-parallel β-strands in Figure 1. As only the side-chains of [^{1}H,^{14}N]-FYTIV residues are ^{1}H-labeled, the NOEs which appear in the differently edited experiments are also subject to the distribution of these residues within the protein. For example, sequential NOEs originating from the side chains of FYTIV residues, which are observed in the ^{15}N-(ω_2)-edited NOESY experiment, can only arise from sequential fragments of the type F/Y/T/I/V-X (where X is a [^{2}H,^{15}N]-labeled residue; Figure 1). Similarly, sequential NOEs (i, i+1) within sequence fragments of [^{1}H,^{14}N]-FYTIV amino acids will only appear in ^{14}N-(ω_2)-edited spectra. Additionally, both types of edited NOESY experiments yield asymmetric spectra (HN–HN region) which can be used to confirm the sequential assignments obtained from the triple-resonance experiments.

By combining our labeling scheme with powerful filtering techniques (Otting and Wüthrich, 1990)

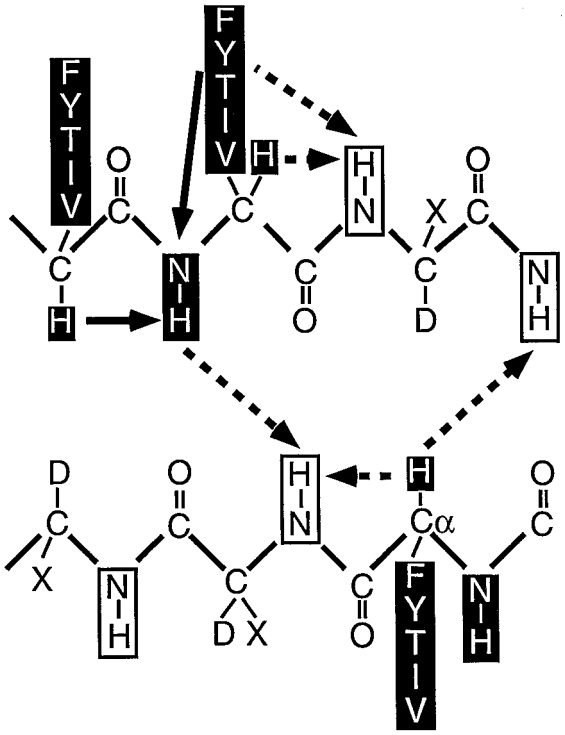

Figure 1. Expected NOEs for two neighboring anti-parallel β-strands (i and j) in ^{14}N/^{15}N-(ω_2)-edited NOESY experiments using the residue-specific labeled sample described in the text. [^{1}H,^{14}N]-FYTIV-labeled residues are indicated by the solid boxes and [^{2}H,^{15}N]-labeled residues (except prolines) are shown by open boxes. NOEs observed in ^{14}N- and ^{15}N-(ω_2)-edited NOESY experiments are shown by solid and dashed arrows, respectively.

we hoped to obtain the following information: (i) an unambiguous assignment of sequential residues; (ii) assignment of Phe and Tyr aromatic side chains, and the aliphatic side chains of Thr, Ile and Val; (iii) identification of NOEs defining elements of secondary structure, particularly of β-sheet regions where strong Hα(i)-HN(i+1) interactions attenuate the NOE build-up of the much weaker, but structurally important inter-strand NOEs (Phe, Tyr, Thr, Ile and Val are known to occur very frequently in β-strands; Chou and Fasman, 1974; Minor and Kim, 1994; Smith et al., 1994); (iv) measurement of long-range NOEs between [^{1}H,^{14}N]-FYTIV protons and either amide protons or other FYTIV side-chain protons, which would normally be significantly attenuated by short-range ^{1}H-^{1}H dipolar interactions; (v) observation of through-bond connectivities for [^{1}H,^{14}N]-FYTIV residues (the 2D ^{1}H-TOCSY spectra of the 47 kDa protein under study were of high quality; data not shown); (vi) measurement of certain NOEs from well-resolved 2D

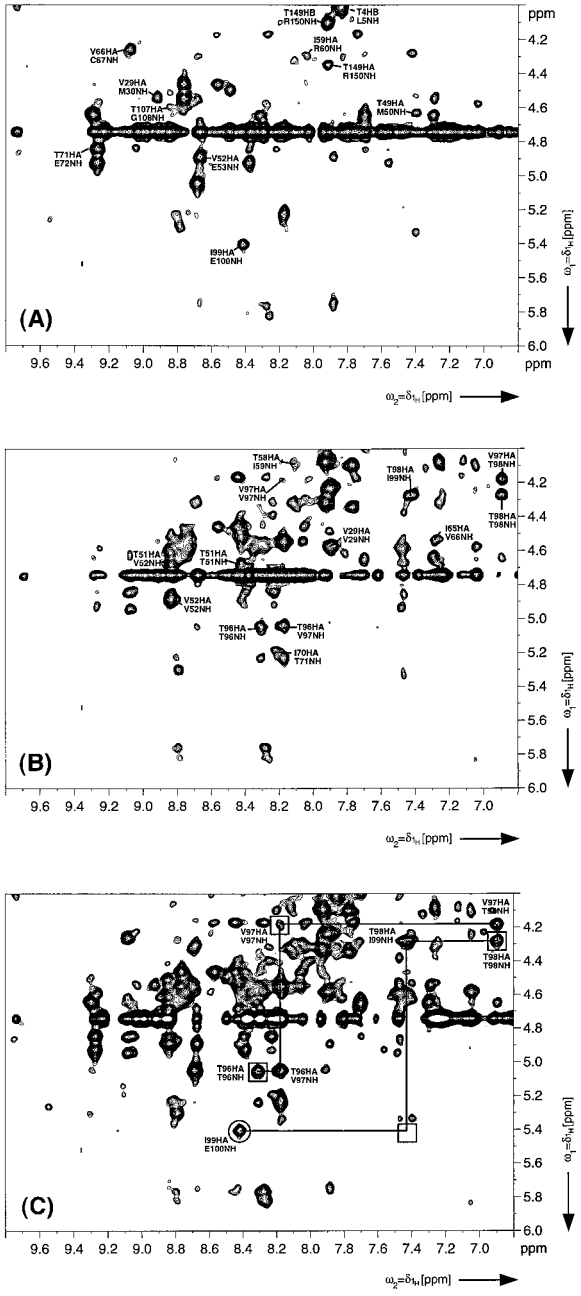

Figure 2. Finger print regions of 2D NOESY spectra of a [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X 3,4-dihydroxy-2-butanone 4-phosphate synthase sample (only a selection of assignments is shown for clarity). Mixing times were 40 ms in all cases. (a) 2D ^{15}N-(ω_2)-edited NOESY spectrum. (b) 2D ^{14}N-(ω_2)-edited NOESY spectrum. (c) 2D NOESY spectrum. The sequential assignment of a sequence fragment involving the residues T96-V97-T98-I99-E100 is indicated by the solid lines. Inter-residue correlations, also appearing in the 2D ^{1}H TOCSY, are boxed and peaks appearing in the 2D ^{15}N-(ω_2)-edited NOESY spectrum are circled. Spectra were recorded using a 2 mM sample at 750 MHz at a temperature of 313 K with a total measurement time of 24 h for each experiment.

homonuclear spectra where there is no transfer via a heteronucleus and where the number of spin-diffusion pathways is reduced, increasing the accuracy of the extracted distances for large proteins.

The amino acid sequence of the 217 residue protein under study (Richter et al., 1992) contains 7 Phe, 2 Tyr, 15 Val, 10 Ile and 21 Thr residues, with 10 sequence fragments where two or more [^{1}H,^{14}N]-FYTIV residues occur sequentially. A [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X-labeled sample of 3,4-dihydroxy-2-butanone 4-phosphate synthase was prepared by growing the *E. coli* strain M15/pREP4 (a derivative of pQE30; Qiagen GmbH, Hilden, Germany) expressing the protein in a minimal M9 medium (Maniatis et al., 1982; 15.0 g/l Na_2HPO_4, 3.0 g/l KH_2PO_4, 0.5 g/l NaCl, 0.13 g/l $MgSO_4$, 33 mg/l $CaCl_2$, 27 mg/l $FeCl_2$ $6.H_2O$, 33.8 mg/l $MnSO_4$ $4.H_2O$, 0.5 mg/l thiamine hydrochloride, 0.5 mg/l biotin, 150 mg/l ampicillin and 25 mg/l kanamycin sulfate) supplemented with 4.0 g/l [U-$^2H_8$]-glycerol, 1.0 g/l $^{15}NH_4Cl$ and 100 mg of each unlabelled amino acid (Phe, Tyr, Thr, Ile and Val) dissolved in 99.9% (v/v) D_2O. Freshly transformed cells from Luria-Bertani plates (Maniatis et al., 1982) were used to inoculate 30 ml of supplemented M9 medium containing 99.9% D_2O which was incubated for 24 h at 37 °C while shaking (200 r.p.m.). Following adaptation to growth in D_2O the cells were diluted into 1.5 l of fresh medium and protein expression was induced by addition of IPTG (isopropyl-β-D-thiogalactoside; 0.5 mM final concentration) when an 0.6 OD_{600} (apparent absorbance at 600 nm) was reached. The culture was incubated for a further 12 h and 3,4-dihydroxy-2-butanone 4-phosphate synthase was purified from the cells by chromatography under native conditions at pH 7.0 (Richter et al., 1999). From 1.5 l of culture, 70 mg of protein was isolated. NMR samples contained 2 mM protein in 50 mM potassium phosphate pH 6.0, 1 mM NaN_3 and 10:90 D_2O/H_2O.

The specificity of the labeling in the sample relies on the regulation of the amino acid biosynthesis pathways in *E. coli*. (Pittard, 1987; Umbarger, 1987). The selectivity of the information provided by the sample suffers if transamination of the proffered amino acids (Phe, Tyr, Thr, Ile and Val) occurs. As the sample was prepared in D_2O, transamination of a supplied [^{1}Hα,^{14}N]-amino acid would yield the [^{2}Hα,^{15}N]-labeled form of this amino acid in the protein. This process could lead to the appearance of cross peaks from the same residue in the amide regions of both the ^{14}N- and ^{15}N-(ω_2)-edited 2D NOESY spectra. We observed that the same HN-HN cross peaks

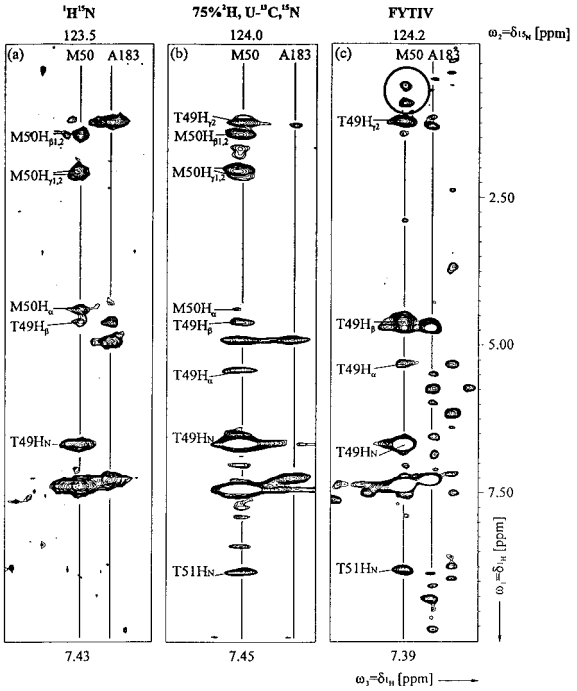

Figure 3. Three strips taken at the ^{1}HN and ^{15}N frequencies of the residue M50 from 3D ^{15}N-separated NOESY-HSQC spectra acquired with the following three samples: (a) a ^{1}H-,^{15}N-labeled sample (τ_m = 40 ms); (b) a 75% ^{2}H, 100% ^{15}N,^{13}C-labeled sample (τ_m = 85 ms) and (c) the [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X sample (τ_m = 80 ms). Inter- and intra-residue NOEs arising from M50 are indicated in the three spectra. Putative long-range NOEs are circled. Spectra were recorded at 750 MHz with 2 mM samples at a temperature of 313 K (with the exception of (b); 310 K) with a total measurement time of 4 days for each experiment.

arising from Val and Ala residues occurred in both the ^{15}N-(ω_2)-edited and ^{14}N-(ω_2)-edited 2D NOESY spectra. This can be explained by the activity of the *E. coli* transaminase C (*avt*A gene product) which catalyzes a reversible transamination between L-alanine and valine (Reitzer and Magasanik, 1987). The remaining [^{1}H,^{14}N]-residues (Phe, Tyr, Thr and Ile) were affected only to a minor extent by transamination (not more than 10% ^{15}N-enriched as estimated from 2D ^{1}H,^{15}N-HSQC spectra). No signals arising from aliphatic side chains of the remaining 15 [^{2}H,^{15}N]-labeled residues (non-Phe, Tyr, Thr, Ile and Val residues) were identified in NOESY spectra.

Previously obtained ^{1}HN and ^{15}N assignments were used to identify the spin systems of [^{1}H,^{14}N]-FYTIV residues in a 3D ^{15}N-separated NOESY of the [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X sample. The most intense cross peaks occurred where [^{2}H,^{15}N]-X residues were directly preceded by [^{1}H,^{14}N]-FYTIV residues in the sequence. These were used as starting points to confirm assignments of sequence fragments of [^{1}H,^{14}N]-FYTIV residues using the 2D edited NOESY spectra. The fingerprint region of the ^{15}N-(ω2)-edited NOESY spectrum (Figure 2a) shows a number of cross peaks, all of the type Hα(i−1)→HN(^{15}N; i) and Hα(i)→HN(^{15}N; j). Using this spectrum, in combination with the 3D ^{15}N-separated NOESY of the same sample, the Hα frequency of the last residue in a [^{1}H,^{14}N]-FYTIV sequence fragment was identified (Figure 2a; Hα(i−1)→HN(^{15}N; i) sequential cross peak). Assignments were extended in reverse (towards the N-terminus) through the [^{1}H,^{14}N]-FYTIV sequence fragments by the main-chain directed approach (Englander and Wand, 1987) using a 2D ^{14}N-(ω_2)-edited NOESY spectrum (Figure 2b) and the 20 ms 2D ^{1}H-TOCSY (data not shown). Figure 2c shows a sequence fragment involving the residues Thr96-Val97-Thr98-Ile99-Glu100. The reduced number of ^{1}H-^{1}H dipolar relaxation pathways is essential to allow the observation of certain interstrand NOEs in large proteins. This is apparent from the amide region of the 2D ^{15}N-(ω_2)-edited NOESY spectrum of the [^{1}H,^{14}N]-FYTIV, [^{2}H, ^{15}N]-X sample which shows a large improvement compared to NOESY spectra from a ^{1}H-,^{15}N-labeled sample (data not shown), as previously observed for highly deuterated samples (Torchia et al., 1988a, b).

The [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X sample also shows properties which indicate its usefulness for the detection of long-range NOEs. This becomes clear when the 3D ^{15}N-NOESY-HSQC spectra of ^{1}H-,^{15}N-labeled, 75% ^{2}H-, 100% ^{15}N-,^{13}C-labeled and [^{1}H,^{14}N]-FYTIV-[^{2}H,^{15}N]-X-labeled samples are compared. Figure 3 shows three strips taken at the HN and ^{15}N frequencies of M50. In the spectrum of the 75% randomly fractionally deuterated, ^{13}C-, ^{15}N-labeled sample many more signals which have narrower linewidths can be identified than in that of the ^{1}H-,^{15}N-labeled sample. In the case of the FYTIV sample, NOEs were identified which were not present in spectra of the other two samples. Sequential (i, i+1) NOEs lead to the appearance of the T49 spin system in the strip taken at the HN and ^{15}N frequencies of M50 in the spectrum of the FYTIV sample (Figure 3c; a number of signals to the right of those from M50 arise from A183 and ridges in the spectrum). Apart from the expected cross peaks from T49 (Hα-region and around 1.25 ppm), additional cross peaks can be seen between 0.5 and 1.0 ppm, which are absent in the other two spectra. These are likely to be long-range

NOEs (signals within the circle in Figure 3c). In our view, it will be possible to assign a large number of NOEs such as these, when high fields are used, due to the reduced overlap in the spectra.

The present investigation of the [^{1}H-^{14}N]-FYTIV-[^{2}H,^{15}N]-X sample of 3,4-dihydroxy-2-butanone 4-phosphate synthase has helped us considerably in the assignment of this 47 kDa protein. The recourse to ^{14}N-labeling in a ^{15}N-labeled background is similar to the reverse-labeling approach of Bax and co-workers (Vuister et al., 1994). Our method has the added advantage that it avoids the detrimental effects of ^{13}C relaxation on proton linewidths and removes most of the ^{1}H-^{1}H dipolar relaxation pathways resulting in the observation of many structurally important NOEs which were previously not detected. Due to the improved transfer efficiency of the ^{1}H TOCSY in the absence of ^{13}C relaxation, we were also able to overcome the problem of the assignment of many Hα and Hβ protons which were absent in the 3D HBHA(CBCACO)NNH spectrum (Grzesiek and Bax, 1993). Additionally, the dramatic simplification of the NOESY spectrum accelerated the assignment of short-range NOEs from sequential interactions, and also aided the identification of many of the side-chain resonances, including all of the Phe and Tyr aromatic ring protons. This labeling scheme will play an important role in the structure determinations of large proteins by supplementing data from ^{2}H-, ^{13}C-, ^{15}N-labeled samples as well as providing a way to study interactions with ligands or larger protein complexes.

Acknowledgements

L.J.B. and Y.Y. are grateful to EMBO and the Max-Planck Society for financial support.

References

Arrowsmith, C.H., Pachter, R., Altman, R.B., Iyer, S.B. and Jardetzky, O. (1990) *Biochemistry*, **29**, 6332–6341.

Chou, P.Y. and Fasman, G.D. (1974) *Biochemistry*, **13**, 211–222.

Clore, G.M. and Gronenborn, A.M. (1994) *Methods Enzymol.*, **239**, 349–363.

Englander, S.W. and Wand, A.J. (1987) *Biochemistry*, **26**, 5953–5958.

Gardner, K.H. and Kay, L.E. (1997) *J. Am. Chem. Soc.*, **119**, 7599–7600.

Gardner, K.H., Rosen, M.K. and Kay, L.E. (1997) *Biochemistry*, **36**, 1389–1401.

Grzesiek, S., Doebeli, H., Gentz, R., Garotta, G., Labhardt, A.M. and Bax, A. (1992) *Biochemistry*, **31**, 8180–8190.

Grzesiek, S., Anglister, J., Ren, H. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 4369–4370.

Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.

Kay, L.E. and Gardner, K.H. (1997) *Curr. Opin. Struct. Biol.*, **7**, 722–731.

LeMaster, D.M. (1990) *Quart. Rev. Biophys.*, **23**, 133–174.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Metzler, W.J., Wittekind, M., Goldfarb, V., Mueller, L. and Farmer II, B.T. (1996) *J. Am. Chem. Soc.*, **118**, 6800–6801.

Minor, D.L. and Kim, P.S. (1994) *Nature*, **367**, 660–663.

Nietlispach, D., Clowes, R.T., Broadhurst, R.W., Ito, Y., Keeler, J., Kelly, M., Ashurst, J., Oschkinat, H., Domaille, P.J. and Laue, E.D. (1996) *J. Am. Chem. Soc.*, **118**, 407–415.

Otting, G. and Wüthrich, K. (1990) *Quart. Rev. Biophys.*, **23**, 39–96.

Pittard, A.J. (1987) In *Escherichia coli and Salmonella Typhimurium* (Ed., Neidhardt, F.C.), Vol. 1, American Society for Microbiology, Washington, DC, pp. 368–394.

Reitzer, L.J. and Magasanik, B. (1987) In *Escherichia coli and Salmonella Typhimurium* (Ed., Neidhardt, F.C.), Vol. 1, American Society for Microbiology, Washington, DC, pp. 302–320.

Richter, G., Volk, R., Krieger, C., Lahm, H.-W., Röthlisberger, U. and Bacher, A. (1992) *J. Bacteriol.*, **174**, 4050–4056.

Richter, G., Kelly, M.J.S., Krieger, C., Yu, Y., Bermel, W., Karlsson, G., Bacher, A. and Oschkinat, H. (1999) *Eur. J. Biochem.*, **261**, 57–65.

Rosen, M.K., Gardner, K.H., Willis, R.C., Parris, W.E., Pawson, T. and Kay, L.E. (1996) *J. Mol. Biol.*, **263**, 627–636.

Shan, X., Gardner, K.H., Muhandiram, D.R., Rao, N.S. and Kay, L.E. (1996) *J. Am. Chem. Soc.*, **118**, 6570–6579.

Smith, B.O., Ito, Y., Raine, A., Teichmann, S., Ben-Tovim, L., Nietlispach, D., Broadhurst, R.W., Terada, T., Kelly, M., Oschkinat, H., Shibata, T., Yokoyama, S. and Laue, E.D. (1996) *J. Biomol. NMR*, **8**, 360–368.

Smith, C.K., Withka, J.M. and Regan, L. (1994) *Biochemistry*, **33**, 5510–5517.

Torchia, D.A., Sparks, S.W. and Bax, A. (1988a) *Biochemistry*, **27**, 5135–5141.

Torchia, D.A., Sparks, S.W. and Bax, A. (1988b) *J. Am. Chem. Soc.*, **110**, 2320–2321.

Umbarger, H.E. (1987) In *Escherichia coli and Salmonella Typhimurium* (Ed., Neidhardt, F.C.), Vol. 1, American Society for Microbiology, Washington, DC, pp. 352–367.

Vuister, G.W., Kim, S.-J. and Bax, A. (1994) *J. Am. Chem. Soc.*, **116**, 9206–9210.

Wagner, G. (1993) *J. Biomol. NMR*, **3**, 375–385.

Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994a) *J. Am. Chem. Soc.*, **116**, 11655–11666.

Yamazaki, T., Lee, W., Revington, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994b) *J. Am. Chem. Soc.*, **116**, 6464–6465.

Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C.H. and Jardetzky, O. (1994) *J. Mol. Biol.*, **238**, 592–614.