



A novel method for the biosynthesis of deuterated proteins with selective protonation at the aromatic rings of Phe, Tyr and Trp

Sundaresan Rajesh^a, Daniel Nietlispach^c, Hiroshi Nakayama^b, Koji Takio^b, Ernest D. Laue^c, Takehiko Shibata^a & Yutaka Ito^{a,b,*}

^aCellular and Molecular Biology Laboratory and ^bBiomolecular Characterization Division RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan; ^cDepartment of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

Received 21 March 2003; Accepted 30 April 2003

Key words: aromatic amino acids, perdeuteration, resonance assignment, selective protonation, shikimic acid

Abstract

A novel biosynthetic strategy is described for the preparation of deuterated proteins containing protons at the ring carbons of Phe, Tyr and Trp, using the aromatic amino acid precursor shikimic acid. Specific protonation at aromatic side chains, with complete deuteration at C^{α/β} positions was achieved in proteins overexpressed in bacteria grown in shikimate-supplemented D₂O medium. Co-expression of a shikimate transporter in prototrophic bacteria resulted in protonation levels of 62–79%, whereas complete labeling was accomplished using shikimate auxotrophic bacteria. Our labeling protocol permits the measurement of important aromatic side chain derived distance restraints in perdeuterated proteins that could be utilized to enhance the accuracy of NMR structures calculated using low densities of NOEs from methyl selectively protonated samples.

NMR studies of proteins over 250 amino acids often employ perdeuteration to alleviate spectral overlap, resolution and sensitivity problems commonly encountered in such large systems (Grzesiek et al., 1993; Yamazaki et al., 1994a, b). This and the introduction of various advances in high resolution NMR such as TROSY (Salzmann et al., 1998), CRINEPT (Riek et al., 1999), the use of residual dipolar couplings (Prestegard et al., 2000) as well as hardware improvements with high field magnets and cryoprobes, etc., have allowed NMR analysis of proteins to extend past the 30 kDa size limit. The perdeuteration induced reduction in the number of proton-proton NOEs, was overcome through the reintroduction of protons either by using labeled amino acids (Smith et al., 1996; Metzler et al., 1996) or methyl selectively protonated metabolic precursors of aliphatic amino acids (Rosen et al., 1996; Gardner et al., 1997a; Goto et al., 1998). The merits of the methyl-selective protonation

approach in a perdeuterated background was successfully demonstrated in the global fold determination of proteins from a limited set of NOEs comprising H^N-H^N, H^N-CH₃ and CH₃-CH₃ restraints (Gardner et al., 1997b; Mueller et al., 2000).

However, the quality of structures determined using the methyl-selective protonation approach suffers from a strong dependence on the secondary structure and topology of the proteins (Gardner et al., 1997b). The method proves less effective for highly helical proteins, even with the use of extensive dipolar coupling restraints (Clare et al., 1999). Smith et al. (1996) showed that NOEs involving aromatic protons are an important source of additional distance restraints in the structure calculation of perdeuterated proteins, on account of their predominance in the hydrophobic protein core. Inclusion of aromatic proton derived distance information has been shown to increase the accuracy of a family of structures calculated using the low densities of NOEs (Smith et al., 1996; Clare et al., 1999, Agazhadeh et al., 1998, Medek et al.,

*To whom correspondence should be addressed. E-mail: ito@louise.riken.go.jp

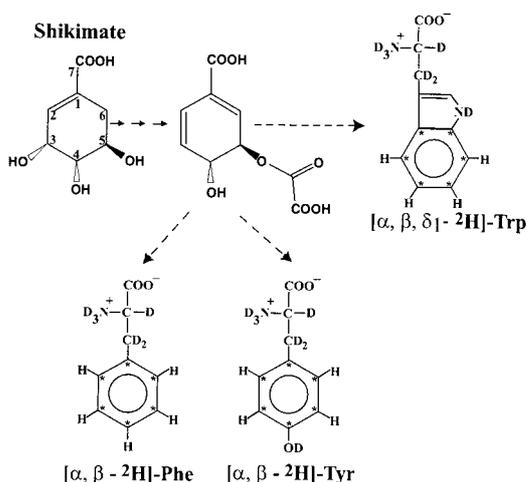


Figure 1. Biosynthetic route for incorporation of $^1\text{H}/^{12}\text{C}$ -shikimic acid into side chains of aromatic amino acids in the proposed selective protonation approach, using *E. coli* grown in D_2O containing media (adapted from Pittard, 1996). Sites labelled by shikimate are represented by *.

2000). In this communication, we introduce a novel cost-effective biosynthetic method for the selective reverse labeling of the aromatic amino acids Phe, Tyr and Trp in perdeuterated proteins, using natural abundance shikimic acid ($^1\text{H}/^{12}\text{C}$). This strategy selectively protonates aromatic rings, while the labeling of the backbone α - and side chain β -positions can be controlled independently via the D_2O and glucose added to the growth medium, thus ensuring the use of standard approaches for the main chain assignment of aromatic residues.

Shikimic acid, a major intermediate in the aromatic amino acid and aromatic vitamin biosynthetic pathway in bacteria and higher plants (Haslam, 1993), is a precursor to the aromatic rings of Phe, Tyr and Trp (Figure 1). The α - and β -sites of Phe and Tyr are derived from phosphoenolpyruvate, whereas phosphoribosyl pyrophosphate and serine contribute to the remainder of the Trp sites. Thus, highly specific labeling of aromatic ring positions is possible using the *Escherichia coli* biosynthetic pathway via shikimic acid. The shikimate based aromatic ring-selective protonation (henceforth referred to as 'aro- ^1H ') was validated on three different proteins with varying numbers of aromatic amino acid residues: Two Phe and one Tyr in the case of ubiquitin (Ub, 76 residues); one Phe, three Tyr and two Trp in the RecA C-terminal domain (RecA_{268–330}, 63 residues); twelve Phe, eight Tyr and three Trp in yeast ubiquitin hydrolase 1 (YUH1, 236 residues). A description of the growth conditions

Table 1. Levels of aromatic side chain protonation^a

Residue type	Efficiency ^b (%) \pm std. dev.	
	BL21(DE3)	BL21(DE3)/shiA
Phe		
$\text{H}^{\epsilon 1/\epsilon 2}$ (2)	44 \pm 1 (8)	64 \pm 2 (8)
H^{ζ} (1)	40 \pm 1 (5)	68 \pm 0.6 (5)
$\text{H}^{\delta 1/\delta 2}$ (2)	48 \pm 1.5 (7)	68 \pm 2 (7)
Tyr		
$\text{H}^{\epsilon 1/\epsilon 2}$ (1)	56 \pm 1 (3)	79 \pm 1.3 (3)
$\text{H}^{\delta 1/\delta 2}$ (1) ^c	45 \pm 0.5 (3)	62 \pm 0.5 (3)

^aThe protonation levels were determined from the 1D amide-rejected proton spectrum of 99.9% $^2\text{H}/^{15}\text{N}$ -Ub samples overexpressed in either BL21(DE3) or in BL21(DE3)/ShiA, in the presence of externally added shikimate, with the control sample consisting of 99.9% $^2\text{H}/^{15}\text{N}$ -Ub, overexpressed in BL21(DE3) to which natural abundance Phe and Tyr were added.

^bThe average protonation levels in the aromatic side chains were obtained by integration of the various peaks from shikimate added samples and dividing by the integral of the corresponding peaks from the control sample (normalized for differences in sample concentration). The values in parentheses, against residue type, indicate the number of peaks analysed, while those against the efficiency indicate the number of measurements.

^cOverlap of the Tyr $\text{H}^{\delta 1/\delta 2}$ peak with one of the two Phe $\text{H}^{\epsilon 1/\epsilon 2}$ peaks in the spectrum of Ub, limited the accuracy of this measurement.

to obtain aro- ^1H samples using shikimic acid is provided as supplementary material. Figure 2 shows a 1D ^1H NMR spectrum recorded on an aro- ^1H , perdeuterated ^{15}N -labeled YUH1. The intense proton signals in the amide and aromatic regions, with negligible protonation at the methyl, α - and β -regions of the spectrum (Figures 2A and 2C) clearly demonstrate the specific nature of shikimic acid incorporation into the aromatic ring side chain positions. It also emphasizes the lack of apparent scrambling of protons from shikimic acid into other sites or residues through alternative biosynthetic pathways. The spectrum in Figure 2B, where all the signals from amide attached protons are suppressed by amide filtering (Otting and Wüthrich, 1990), further highlights the selectivity of the aro- ^1H approach. Detection of strong intra-residue NOEs between aromatic ring protons and amide protons of aromatic residues and the absence of amide to aliphatic NOEs for all the three amino acid types clearly implicated externally added shikimic acid as the source of side chain protons in the aromatic amino acids targeted (Figure 2C).

We next investigated the levels of shikimate-derived protonation in aromatic amino acids, using Ub as the test sample. Table 1 lists the levels of proto-

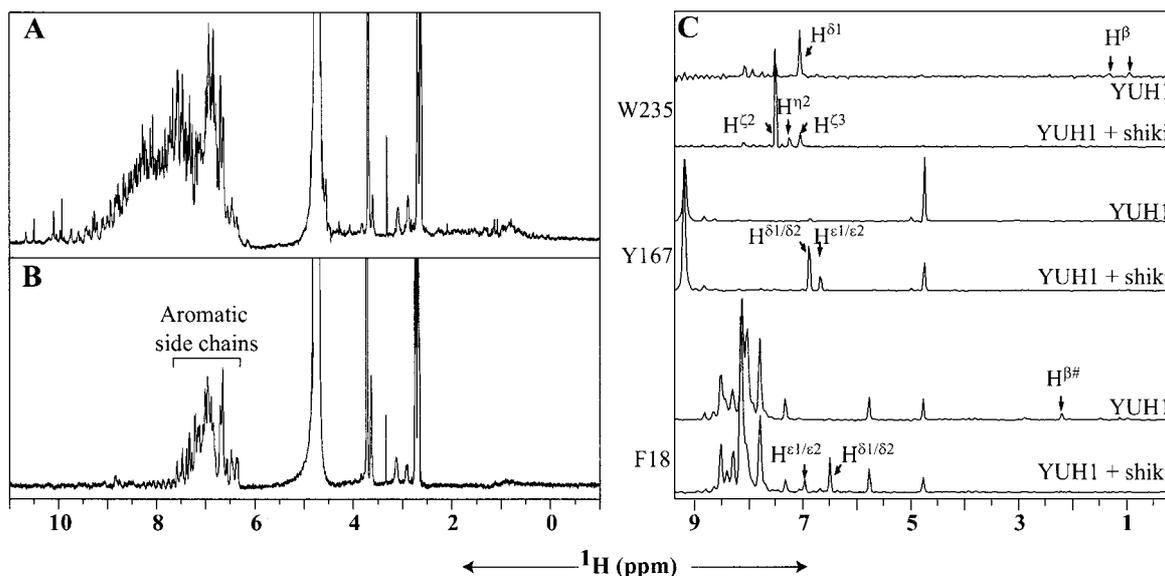


Figure 2. 1D ^1H spectrum (A), and amide-rejected 1D ^1H spectrum (B), acquired at 600 MHz on a 99.9% $^2\text{H}/\text{u}$ - ^{15}N -labelled YUH1 in 90% $\text{H}_2\text{O}/10\%\text{D}_2\text{O}$ buffer, selectively protonated with shikimic acid, as described in the text. Aromatic side chain specific proton signals, derived from shikimic acid are indicated. (C) Comparison of F1 (^1H) cross sections, at particular F2 (^{15}N) chemical shifts for three representative aromatic residues from 3D ^{15}N -separated NOESY-HSQC spectra of 95% $^2\text{H}/^{15}\text{N}$ -YUH1 (YUH1) and aro- ^1H protonated 99% $^2\text{H}/^{15}\text{N}$ -YUH1 (YUH1 + shiki) protein. Both spectra were recorded over 5 days with a mixing time of ~ 200 ms on a 1 mM sample of ^{15}N -labeled protein. The spectra were processed under identical conditions. Intraresidue NOE cross peaks to amide protons are represented with solid arrows.

nation at the observable side chain positions of Phe and Tyr in aro- ^1H perdeuterated/ ^{15}N labeled ubiquitin samples. Amide filtered 1D ^1H NMR spectra (Otting and Wüthrich, 1990) measured on aro- ^1H $^2\text{H}/^{15}\text{N}$ -Ub, expressed in the *E. coli* BL21(DE3) strain revealed incorporation levels of 40–48% for Phe and 45–56% for Tyr aromatic protons. Changes in the time of shikimate addition prior to protein induction or increasing the concentration of added shikimate did not improve the levels of protonation (data not shown). Apparently the low proton densities at these sites are due to the inefficient uptake of shikimate by BL21(DE3). One possible reason for this would be the catabolite repression of genes encoding transporters of non-glucose carbon sources by the glucose in the growth medium (Saier et al., 1996). Studies by Frost and co workers (Knop et al., 2001) suggest a similar scenario for shikimate transporters in *E. coli*. Overexpression of ShiA, a membrane bound transporter of shikimate, was shown to increase the uptake of shikimic acid from the growth medium in shikimate auxotrophs of *E. coli* K-12 (Whipp et al., 1998). Hence, we examined the effect of overexpression of recombinant ShiA, on the protonation levels in the above system. As expected, co-expression of recombinant ShiA in BL21(DE3) [referred to as BL21(DE3)/shiA] markedly increased the

levels of side chain protonation to 66–68% for Phe and 62–79% for Tyr (Table 1). The levels quantified by NMR were independently confirmed by MS analyses of lysylendopeptidase digests from the aro- ^1H Ub samples (data not shown).

Finally, complete protonation of aromatic positions by total incorporation of shikimate was achieved using an auxotrophic bacterial strain defective in endogenous shikimate production. The *E. coli* K-12 strain, AB2826, harbors a mutation in the *aroB* gene and fails to synthesize an upstream precursor of shikimate, thus requiring an external source of shikimate, or aromatic amino acids for growth (Brown and Doy, 1976). The *shiA* plasmid was also introduced into the AB2826 strain to ensure increased transport of shikimate, thereby increasing the growth rate of the bacteria (see supplementary material for details). The total incorporation of shikimate is demonstrated by the ^1H - ^{13}C HSQC spectra of protonated u- $^{13}\text{C}/^{15}\text{N}$ -YUH1 overexpressed in AB2826(DE3)/shiA, in minimal media containing $^1\text{H}/^{13}\text{C}$ -glucose and $^{15}\text{NH}_4\text{Cl}$, supplemented with natural abundance shikimic acid (Figure 3B). The absence of aromatic side chain crosspeaks from Phe, Tyr and Trp (excluding $\delta 1$) in the above sample, compared to that expressed in BL21(DE3) host minus shikimate (Figure 3A), clearly

pling constants (Grzesiek and Bax, 1995; Yamazaki et al., 1993; Löhr and Rüterjans, 1996; Prompers et al., 1998), tend to become insensitive in case of larger proteins as the fast transverse relaxation of $^1\text{H}^\beta/^{13}\text{C}^\beta$ coherences decreases the transfer efficiency and broadens the signals. Moreover, severe overlap of the already broadened signals in the $^{13}\text{C}^\beta$ dimension further complicate the assignment process, with the applicability of these methods being further limited by perdeuteration of all non-exchangeable protons. Hence, the assignment of aromatic side chain protons in aro- ^1H perdeuterated proteins were made by correlating the signals of the through-bond coupled protons in amide filtered 2D homonuclear ^1H -TOCSY spectra, to the intra-residue aromatic- H^N NOEs in 3D ^{15}N -separated NOESY-HSQC spectra (Figure 4A). The amide ^1H and ^{15}N assignments for the three proteins tested were obtained from earlier studies (Rajesh et al., 1999; Aihara et al., 1997; Sakamoto et al., 1999). For YUH1 (Figure 4A), complete aromatic proton assignment was obtained for the seven Tyr and the three Trp residues. The $\text{H}^{\delta 1}$ of Trp could not be observed in aro- ^1H samples, as is expected from the biosynthetic pathway of Trp (Figure 1). All of the other aromatic ring protons of Trp ($\text{H}^{\epsilon 3}$, $\text{H}^{\zeta 2}$, $\text{H}^{\zeta 3}$ and $\text{H}^{\eta 2}$) were successfully assigned. A further advantage of this protocol is the incorporation of ^{15}N label from the $^{15}\text{NH}_4\text{Cl}$ in the growth medium at the $\epsilon 1$ position of tryptophan residues. The $\text{H}^{\epsilon 1}$ proton is often well resolved and provides additional intraresidue NOEs to side chain protons, assisting in the complete assignment of the Trp ring protons. As an example, Figure 4A shows the complete aromatic ring proton assignment of W235 of YUH1, which was obtained from the intra-residue aromatic to amide and $\text{H}^{\epsilon 1}$ NOEs. In addition, the $\text{H}^{\epsilon 1}$ proton also provides crucial long range NOE information that helps in the rapid determination of the global fold of proteins (long range NOEs from $\text{H}^{\epsilon 1}$ proton of W81 and aromatic to H^N NOEs for few other amino acids is indicated in Figure 4B). Assignments for most of the $\text{H}^{\epsilon 1/\epsilon 2}$ and $\text{H}^{\delta 1/\delta 2}$ resonances in the twelve Phe residues were obtained for YUH1. Although TOCSY based H^ζ assignment was obtained for one Phe residue in YUH1, indicating that shikimic acid labeling also targets ζ -positions of Phe, other H^ζ protons could not be observed, probably as a result of chemical shift overlap with the ϵ - or δ -protons. In comparison, a fractional 50%- $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled YUH1 yielded only half of the aromatic ^1H resonance assignments (data not shown), demonstrating the advantages of our protocol. By virtue of the aro-

^1H approach, complete aromatic side chain proton assignments could be obtained also for RecA_{262–330} and Ub (see Supplementary material available from the author) in a similar manner.

The longer mixing times used in the NOESY experiments on aro- ^1H perdeuterated samples, can lead to some ambiguity in the assignment of aromatic ring ^1H spin systems, due to spin diffusion amongst the aromatic protons. Nevertheless, the improved sensitivity of the amide to aromatic NOE transfer, facilitated by the deuteration of the C^α and C^β positions, provides efficient intra-residue H^N to aromatic side chain NOE assignments (Figure 4B). Any resulting ambiguities in the proton position assignment can then be treated using the ambiguous distance restraints approach (Nilges, et al., 1998) for the structure calculation. The absence of ^{13}C labels in the aromatic side chains on account of reverse labeling by shikimate, though beneficial in case of Phe due to the limited dispersion of aromatic ^1H and ^{13}C resonances, would probably be limiting in the assignment of Tyr and Trp rings that often have much better dispersion of aromatic ^{13}C resonances. In the proteins tested in this study, however, the absence of aromatic ^{13}C labels did not prove to be problematic for the assignments of the Tyr and Trp side chains. Floss and co-workers (Cho et al., 1992) have reported the chemical synthesis of ^{13}C labeled shikimate at the 1 or 2 positions (Figure 1), which could be employed in the proposed labeling scheme for site specific labeling of aromatic side chains with ^{13}C in order to overcome the ambiguities in aromatic side chain assignments of proteins with numerous aromatic residues, and to provide unambiguous aromatic derived NOEs. The versatility of this labeling approach allows for experimenting with different ^{13}C -labeling patterns at aromatic side chains in high molecular weight proteins using appropriately designed chemically synthesized ^{13}C -shikimic acid.

In summary, we have described a novel, cost-effective and easy to implement biosynthetic procedure for the selective protonation of aromatic side chain ring positions in otherwise fully perdeuterated proteins. The divergent metabolic pathways of aromatic amino acids (Phe, Tyr and Trp) and aliphatic amino acids (Leu, Ile and Val) allow one to effectively combine the above shikimate protonation with the methyl-selective protonation approach (Gardner et al., 1997b), providing the much-needed additional long-range aromatic distance restraints in perdeuterated proteins. The above selective protonation scheme will prove to be highly beneficial for the rapid and

accurate global fold determination of high molecular weight proteins from a limited set of NOEs.

Acknowledgements

The authors express their sincere gratitude to Prof. Jim Pittard for the ShiA plasmid and AB2826 strain; to Dr Toshiyuki Kohno for the YUH1 and Ub plasmids. We thank Prof Masahiro Shirakawa, Dr Brian Smith and Dr Tsutomu Mikawa for a critical reading of the manuscript. This work was supported by Grants in-Aid for Scientific Research on Priority Areas (13014221) from the Ministry of Education, Culture, Sports, and Technology (MEXT), as well as MR Science and Molecular Ensemble programs (RIKEN, Japan) to YI.

References

- Aghazadeh, B., Zhu, K., Kubiseski, T.J., Liu, G.A., Pawson, T., Zheng, Y. and Rosen, M.K. (1998) *Nat. Struct. Biol.*, **5**, 1098–1107.
- Aihara, H., Ito, Y., Kurumizaka, H., Terada, T., Yokoyama, S. and Shibata, T. (1997) *J. Mol. Biol.*, **274**, 213–221.
- Brown, K.D. and Doy, C.D. (1976) *Biochim. Biophys. Acta*, **428**, 550–662.
- Cho, H., Heide, L., and Floss, H.G. (1992) *J. Label. Comp. Radiopharm.*, **31**, 589–592.
- Clare, G.M., Starich, M.R., Bewely, C.A., Cai, M. and Kuszewski, J. (1999) *J. Am. Chem. Soc.*, **121**, 6513–6514.
- Farmer II, B.T. and Venters, R.A. (1995) *J. Am. Chem. Soc.*, **117**, 4187–4188.
- Gardner, K.H. and Kay, L.E. (1997a) *J. Am. Chem. Soc.*, **119**, 7599–7600.
- Gardner, K.H., Rosen, M.K. and Kay, L.E. (1997b) *Biochemistry*, **36**, 1389–1401.
- Goto, N.K., Gardner, K.H., Mueller, G.A., Willis, R.C. and Kay, L.E. (1999) *J. Biomol. NMR*, **13**, 369–374.
- Grzesiek, S. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 6527–6531.
- Grzesiek, S., Anglister, J., Ren, H and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 4369–4370.
- Haslam, E. (1993) *Shikimic Acid: Metabolism and Metabolites*, Wiley & Sons, New York.
- Knop, D.R., Draths, K.M., Chandran, S.S., Barker, J.L., von Daeniken, R., Weber, W. and Frost, J.W. (2001) *J. Am. Chem. Soc.*, **123**, 10173–10182.
- Lohr, F. and Rüterjans, H. (1996) *J. Magn. Reson.*, **B112**, 259–268.
- Medek, A., Olejniczak, E.T., Meadows, R.P. and Fesik, S.W. (2000) *J. Biomol. NMR*, **18**, 229–238.
- Metzler, W.J., Wittekind, M., Goldfarb, V., Mueller, L. and Farmer II, B.T. (1996) *J. Am. Chem. Soc.*, **118**, 6800–6801.
- Mueller, G.A., Choy, W.Y., Yang, D., Forman-Kay, J.D., Venters, R.A. and Kay, L.E. (2000) *J. Mol. Biol.*, **300**, 197–212.
- Nilges, M., Macias, M.J., O'Donoghue, S.I. and Oschkinat, H. (1997) *J. Mol. Biol.*, **269**, 408–422.
- Otting, G. and Wüthrich, K. (1990) *Quart. Rev. Biophys.*, **23**, 39–96.
- Pittard, A.J. (1996) In *Escherichia coli and Salmonella Typhimurium: Cellular and Molecular Biology*, Chapter 28. Niedhardt, F. C. (Ed.), ASM Press, Washington, DC.
- Prestegard, J.H., Al-Hashimi, H.N. and Tolman, J.R. (2000) *Quart. Rev. Biophys.*, **33**, 371–424.
- Rajesh, S., Sakamoto, T., Iwamoto-Sugai, M., Shibata, T., Kohno, T. and Ito, Y. (1999) *Biochemistry*, **38**, 9242–9253.
- Riek, R., Wider, G., Pervushin, K. and Wüthrich, K. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 4918–4923.
- Rosen, M.K., Gardner, K.H., Willis, R.C., Parris, W.E., Pawson, T. and Kay, L.E. (1996) *J. Mol. Biol.*, **26**, 627–636.
- Saier, M.H. Jr., Ramsier, T.M. and Reizer, J. (1996) In *Escherichia coli and Salmonella Typhimurium: Cellular and Molecular Biology*, Chapter 85. Niedhardt, F.C. (Ed.), ASM Press, Washington, DC.
- Sakamoto, T., Tanaka, T., Ito, Y., Rajesh, S., Iwamoto-Sugai, M., Koderia, Y., Tsuchida, N., Shibata, T. and Kohno, T. (1999) *Biochemistry*, **38**, 11634–11642.
- Salzmann, M., Pervushin, K., Wider, G., Senn, H. and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 13585–13590.
- 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.
- Vuister, G.W., Kim, S.J., Wu, C. and Bax, A. (1994) *J. Am. Chem. Soc.*, **116**, 9206–9210.
- Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.
- Whipp, M.J., Camakaris, H. and Pittard, A.J. (1998) *Gene*, **209**, 185–192.
- Wüthrich, K. (1986) In *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.
- Yamazaki, T., Forman-Kay, J.D. and Kay, L. (1993) *J. Am. Chem. Soc.*, **115**, 11054–11055.
- Yamazaki, T., Lee, W., Revingtom, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H. and Kay, L.E. (1994a) *J. Am. Chem. Soc.*, **116**, 6464–6465.
- Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994b) *J. Am. Chem. Soc.*, **116**, 11655–11666.