Journal of Biomolecular NMR, 4 (1994) 863–870
ESCOM

J-Bio NMR 221

Differential deuterium isotope shifts and one-bond ^{1}H-^{13}C scalar couplings in the conformational analysis of protein glycine residues

David M. LeMaster*, John C. LaIuppa and Diana M. Kushlan

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208, U.S.A.

Received 2 June 1994
Accepted 25 July 1994

Keywords: Deuterium isotope shift; Glycine conformation; Methylene stereochemistry; Hyperconjugation

SUMMARY

The one-bond deuterium isotope shift effect for glycine C^{α} resonances exhibits a conformational dependence comparable to that of the corresponding $^{1}J_{HC}$ scalar coupling in both magnitude (~ 11 Hz at 14.1 T) and dihedral angle dependence. The similarity in the conformational dependence of the $^{1}J_{HC}$ and deuterium isotope shift values suggests a common physical basis. Given the known distribution of (ϕ,ψ) main-chain dihedral angles for glycine residues, the deuterium isotope shifts and the $^{1}J_{HC}$ scalar couplings can determine conformations in the left- and right-handed helical-to-bridge regions of the (ϕ,ψ) plane to an accuracy of approximately 13°. In the absence of stereochemical assignments, the differential deuterium isotope shifts and the $^{1}J_{HC}$ scalar couplings can be combined with limited independent structural information (e.g., the sign of ϕ) to determine the chirality of the deuterium substitution.

In the course of observing the ^{2}H-decoupled ^{1}H-^{13}C correlation peaks of *E. coli* thioredoxin (LeMaster and Kushlan, 1993) and staphylococcal nuclease H124L, enriched with a mixture of chirally deuterated [2-^{13}C]glycine samples, it became apparent that the ^{13}C chemical shifts for the pro-*R* and pro-*S* resonances (arising from the $^{2}H_{S}$ and $^{2}H_{R}$ samples, respectively) of the same residue could vary as much as 70 ppb (11 Hz at 14.1 T). Such a differential isotope shift is illustrated in Fig. 1 for Gly^{20} of staphylococcal nuclease in the Ca^{2+} · pTp inhibitor complex. Apparent as well is the variation in the $^{1}J_{H\alpha C\alpha}$ values in this ^{13}C-coupled spectrum. The conformational dependence of $^{1}J_{H\alpha C\alpha}$ for peptide groups was observed earlier (Egli and Von Philipsborn, 1981). These authors interpreted their data in terms of an increased coupling resulting from the alignment of the H-C bond with the p_z orbital of the attached nitrogen. Conversely, alignment of the H-C bond with the π orbital of the carbonyl carbon results in a smaller coupling constant.

*To whom correspondence should be addressed.

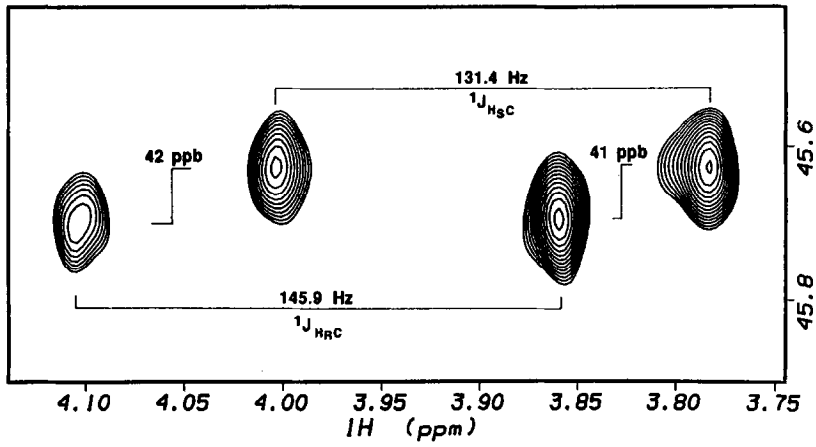

Fig. 1. ^{2}H-decoupled, ^{13}C-coupled HSQC spectrum of [^2H_R,2-^{13}C]glycine + [^2H_S,2-^{13}C]glycine enriched staphylococcal nuclease H124L in the Ca^{2+} · pTp ternary complex at 14.1 T. The cross peaks corresponding to Gly^{20} are illustrated. A 2 kHz ^{2}H decoupling field was used.

More recently (Mierke et al., 1992; Vuister et al., 1992,1993), the anticipated trigonometric dependence of $^1J_{H\alpha C\alpha}$ on the main-chain dihedral angles has been extensively verified experimentally using isotopically enriched protein samples. These studies have demonstrated that this conformational dependence of $^1J_{H\alpha C\alpha}$ can potentially serve as a useful constraint for protein structural analysis. In analogy, the present study analyzes the readily obtained differential isotope shifts for their corresponding utility as structural constraints.

Due to stereochemical ambiguity of the glycine methylene resonances, the earlier analyses of $^1J_{H\alpha C\alpha}$ values used only the methine C^α positions in their correlations. To examine the similarity of the dihedral angle dependencies for the glycyl $^1J_{H\alpha C\alpha}$ values, [^2H_R,2-^{13}C]glycine and [^2H_S,2-^{13}C]glycine were incorporated into *E. coli* thioredoxin and staphylococcal nuclease H124L. In this case, since two separate $^1J_{H\alpha C\alpha}$ values are available to characterize each glycine residue, the analyses were carried out in terms of both the individual $^1J_{H\alpha C\alpha}$ values and their difference $\Delta^1J_{H\alpha C\alpha}$. The observed $\Delta^1J_{H\alpha C\alpha}$ values are given in Table 1, along with the corresponding (ϕ,ψ) values from high-resolution X-ray structure determinations (Loll and Lattman, 1989; Katti et al., 1990). The individual glycine $^1J_{H\alpha C\alpha}$ were fit via singular value decomposition (Press et al., 1989) to the trigonometric function:

$$^1J_{H\alpha C\alpha} = 142.6 + 3.1\cos^2\phi' - 7.6\cos^2\psi' - 1.1\sin\phi' - 3.0\sin\psi' \quad (1)$$

where $\phi' = \phi + 30°$ and $\psi' = \psi - 30°$ for the pro-*R* hydrogen and $\phi' = \phi + 150°$ and $\psi' = \psi - 150°$ for the pro-*S* hydrogen. In terms of both the magnitude and relative contributions of the ϕ and ψ dependencies, these glycine parameters agree quite closely with those reported by Bax and coworkers (Vuister et al., 1992) for the main-chain methine resonances. With the present data, the use of the $\sin\phi$ term for the fifth empirical parameter gave equivalent results as compared to the use of a 12° phase shift of ψ proposed in this earlier study. The observed scalar couplings fit this empirical function with an rmsd of 1.7 Hz. Although this deviation is reasonably small compared to the full range of the Δ^1J_{HC} values (± 15 Hz), as noted earlier (Vuister et al., 1993), it is

appreciably larger than the error in the coupling constant measurements, consistent with the presence of other minor contributions to the observed $^1J_{HC}$ values.

The fact that the geometric dependence of the $^1J_{HC}$ values for both the pro-*R* and pro-*S* protons can be well fit to the same parameterization indicates that all of the structurally useful information on (ϕ,ψ) is contained in the differential $^1J_{HC}$ values. The analogous four-parameter equation was used to fit the trigonometric dependence of the differential deuterium isotope shift data ($\Delta(^1\Delta^{13}C(^2H))$ in ppb).

$$\begin{aligned}\Delta(^1\Delta^{13}C(^2H)) = {} & 12\,(\cos^2\phi'_r - \cos^2\phi'_s) - 43\,(\cos^2\psi'_r - \cos^2\psi'_s) \\ & + 17\,(\sin\phi'_r - \sin\phi'_s) - 21\,(\sin\psi'_r - \sin\psi'_s)\end{aligned} \tag{2}$$

The contour plot of this trigonometric fitting function is given in Fig. 2. The experimental differential isotope shifts fit this empirical function with an rmsd of 13 ppb (2.0 Hz at 14.1 T). Assuming independent, normally distributed errors, the standard deviations for the four coefficients in Eq. 2 are 5.5, 4.0, 1.4 and 2.6, respectively. The true error values for the sinϕ and sinψ terms are likely underestimated, due to the presence of significant covariance with the $\cos^2\psi$ and $\cos^2\phi$ terms, respectively. In Fig. 3 the regression plot of the experimental vs. the predicted

TABLE 1
DIFFERENTIAL ISOTOPE SHIFTS AND $^1J_{HC}$ COUPLINGS FOR GLYCYL RESIDUES OF *E. coli* THIOREDOXIN AND STAPHYLOCOCCAL NUCLEASE

Residue	ϕ (°)	ψ (°)	$\Delta(^1\Delta^{13}C(^2H))$	$\Delta^1J_{H^\alpha C^\alpha}$
***E. coli* thioredoxin**				
Gly^{33}	−61	−54	71	11.2
Gly^{51}	89	−21	41	4.4
Gly^{65}	−82	−26	16	7.6
Gly^{71}	68	30	−27	−10.1
Gly^{74}	−150	167	−15	2.8
Gly^{84}	81	10	18	−5.3
Gly^{92}	98	160	41	0.4
Gly^{97}	−64	−39	65	14.6
Staphylococcal nuclease H124L				
Gly^{20}	−60	−34	41	14.5
Gly^{29}	73	15	9	−0.6[a]
Gly^{50}	97	−154	10	−1.7[a]
Gly^{79}	−102	−145	−58	−4.6
Gly^{86}	83	16	11	−6.3
Gly^{88}	−69	145	−19	11.3
Gly^{96}	76	9	22	−1.9
Gly^{107}	75	25	−30	−7.4

Residues Gly^{21} of *E. coli* thioredoxin and Gly^{55} and Gly^{141} of staphylococcal nuclease are not included, due to ^1H-^1H geminal degeneracy (Wang et al., 1990). The differential isotope shift is defined as $\delta(^2H_S) - \delta(^2H_R)$ $(=(\delta(^1H) - \delta(^2H_R)) - (\delta(^1H) - \delta(^2H_S)) = \Delta^2H_R - \Delta^2H_S)$ (in ppb), consistent with the most common convention (Hansen, 1988). The differential coupling constants are defined as $^1J_{HRC} - {}^1J_{HSC}$.

[a] Denotes differential coupling constants obtained from the corresponding resonances of the apo-nuclease spectrum.

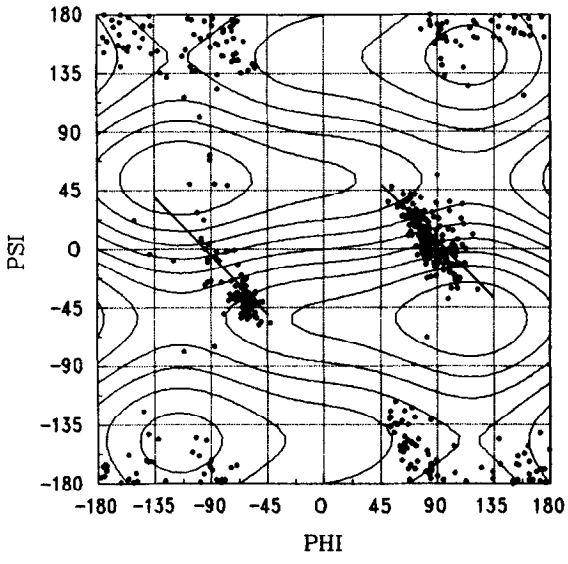

Fig. 2. Contour plot of Eq. 2, illustrating the predicted dihedral angle dependence of the differential one-bond deuterium isotope shift for the glycine C^{α} resonances. The contour interval is 20 ppb, with the zero-level contour passing through the origin. Also plotted are the (ϕ,ψ) values for 498 glycine residues observed in 37 high-resolution X-ray structures. The diagonal lines illustrate the best fit (Eq. 3) to the 281 glycine residues found in the helix-to-bridge regions of the conformational plane (i.e. $-90° < \psi < 90°$).

differential isotope shifts is given, based on Eq. 2. A linear regression coefficient $r^2 = 0.85$ is obtained.

The differential isotope shift and coupling constant values bear considerable similarities in terms of both the opposite signs of the $\cos^2\phi$ and $\cos^2\psi$ dependencies and the approximate threefold dominance of the $\cos^2\psi$ term. On the other hand, there is a clear contrast in the relative magnitude of the geometric dependence of the $^1J_{HC}$ values and the isotope shift effects. As a function of (ϕ,ψ), the $^1J_{HC}$ constants vary over a range of only 10% of the average $^1J_{HC}$ value. In contrast, the relative variation in $^1\Delta^{13}C(^2H)$ is nearly 30% of the average $^1\Delta^{13}C(^2H)$ (0.25 ppm).

As compared to the more commonly used ^3J values, the obvious practical weakness of $^1J_{H^{\alpha}C^{\alpha}}$ analysis in structural studies is the dependence of a single parameter on two independent dihedral angles. Since the differential isotope shifts exhibit a quite similar conformational dependence, they will not significantly enhance the ability to define the two-dimensional information needed to define an arbitrary position in the (ϕ,ψ) plane. This would appear to be a particularly significant limitation in the case of glycine residues, for which the standard dipeptide conformational analysis (Ramachandran and Sasisekharan, 1968) suggests that nearly the entire (ϕ,ψ) plane is potentially accessible.

In practice, the observed glycine dihedral angle distribution shows considerably more clustering than is anticipated from the prediction of permissible conformations based on the simple hard-sphere model, as illustrated in Fig. 2 for 498 glycine residues, drawn from 37 high-resolution X-ray structures (Bernstein et al., 1977) having refinement parameters of at least approximately 1.7 Å resolution, with an R-factor of 17%. In particular, almost 60% of all glycine residues have (ϕ,ψ) dihedral angles lying in the α-helical-to-bridge regions (here operationally defined as

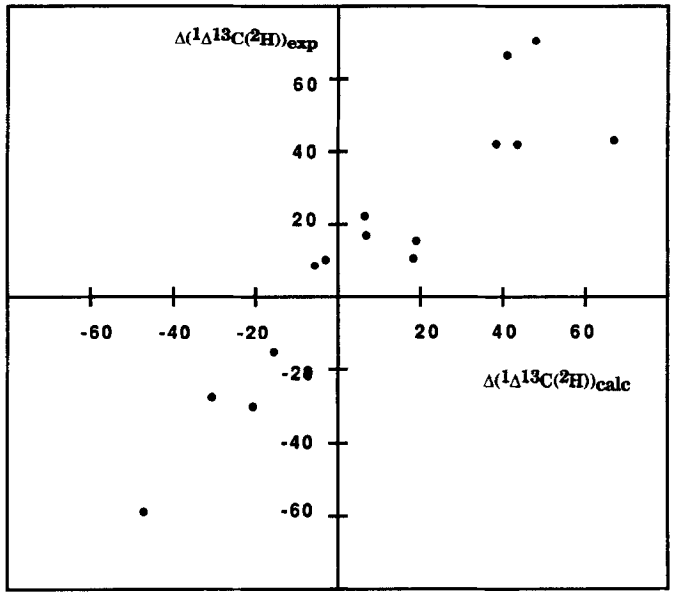

Fig. 3. Linear regression analysis of the experimental differential isotope shifts (i.e., $\Delta(^1\Delta^{13}C(^2H))$) of Table 1 vs. those calculated using the trigonometric fit function of Eq. 2, relating these data (in ppb) to the corresponding (ϕ,ψ) dihedral angles of the high-resolution X-ray structures. A linear regression coefficient $r^2 = 0.85$ is obtained.

$-90° < \psi < 90°$) of either the left- or right-handed conformations (i.e. $\phi > 0$ or $\phi < 0$). A least-squares analysis of the elongated distributions in these two regions yields a linear fit of the form:

$$\phi = -0.94\psi \pm 87° \tag{3}$$

for the $\phi > 0$ and $\phi < 0$ regions, respectively. The rmsd for the fit to these 281 glycine residues in the (ϕ,ψ) plane is 11.3°.

The relevance to this research of an efficient linear approximation of the (ϕ,ψ) distribution for this set of glycine residues stems from the fact that this linear fit function lies approximately parallel to the gradients of the trigonometric fit functions for both the $^1J_{H^{\alpha}C^{\alpha}}$ (Vuister et al., 1992) and the differential isotope shifts (Fig. 2). Hence, if independent structural information can discriminate between a glycine conformation in the extended regions vs. the helix-to-bridge regions, the combination of the $^1J_{H^{\alpha}C^{\alpha}}$ and $\Delta(^1\Delta^{13}C(^2H))$ data should provide a useful estimate of the best choice along the curve of Eq. 3. Assignment of a residue to the left- vs. right-handed helix-to-bridge region is carried out as discussed below in the section on determination of stereochemistry.

A least-squares analysis of the observed differential $^1J_{HC}$ and $\Delta(^1\Delta^{13}C(^2H))$ values for the 11 glycine residues of these conformational regions for which complete data is available yields the following fit for $\phi < 0$ ($\Delta(^1\Delta^{13}C(^2H))$ in ppb and Δ^1J_{HC} in Hz):

$$\phi = 0.18\ \Delta(^1\Delta^{13}C(^2H)) + 0.38\ \Delta^1J_{HC} - 81 \tag{4a}$$

$$\psi = -0.37\ \Delta(^1\Delta^{13}C(^2H)) - 0.78\ \Delta^1J_{HC} - 12 \tag{4b}$$

while for $\phi > 0$, the signs of the constant terms are reversed. Using these equations one can determine for each individual glycine residue how well the observed Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values reproduce the best choice along the $\phi = -0.94\psi \pm 87°$ parametric fit for the observed (ϕ,ψ) X-ray values at that residue. For the 11 glycine residues, the rmsd is 6.5°.

Using these rmsd values, an estimate can be made for the quality of the determination of (ϕ,ψ) based on the Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values. The rmsd for the perpendicular distance between the observed (ϕ,ψ) values and the $\phi = -0.94\psi \pm 87°$ parametric fit is 11.3°, while the estimated error of 6.5° resulting from using the Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values to place an estimate along the (ϕ,ψ) parametric fit is orthogonal. Assuming the independence of these errors, an overall rmsd error of 13° is obtained for the determination of (ϕ,ψ) in these conformational regions based on experimental couplings and isotope shifts.

The accuracy of this error estimate is dependent on two distinct factors. The parameters of the best fit line are population statistics, determined from a large number of the best quality X-ray structures available and hence are surely comparatively reliable. On the other hand, the estimate of how well a given pair of Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values determines a position along that best fit line is dependent on several factors. First, this estimation depends on the accuracy of the X-ray determination of the particular glycine residues considered, although it may be noted that for these two proteins the X-ray temperature factors for the glycine residues of the helix-to-bridge regions do not deviate substantially from the bulk of the main-chain atoms. In addition, the 22 Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values for these glycines, used to derive the two independent parameters of Eq. 4, are in turn being used to estimate the reliability of those parameters and thus the error assessment is not strictly independent.

However, although the estimate of a 6.5° rmsd may somewhat overstate the accuracy to which the differential isotope shifts and coupling constants can determine a position along the best fit line to the glycine conformational population, the analysis strongly suggests that this source of error is significantly smaller than that introduced by the initial assumption of the projection of the (ϕ,ψ) values onto the best fit line (i.e. 11.3°). Hence, the estimate of an overall rmsd of 13° is unlikely to be substantially modified by the limitations discussed above.

Since the magnitudes of the differential isotope shifts in ppb are four- to fivefold larger than the range of Δ^1J_{HC} in Hz, for the fit described above in Eq. 4 the isotope shifts provide an approximately twofold larger contribution to the estimation of ϕ and ψ than do the scalar coupling data. However, it should be noted that these estimates are robust with respect to the relative weighting of the Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values.

The analyses given above all presuppose the known stereochemical assignment of the deuterium labeling. It is of considerable practical interest to reverse this question to ask how useful is the combination of Δ^1J_{HC} and $\Delta(^1\Delta^{13}C(^2H))$ values in the determination of stereochemistry. For all of the 16 glycine residues listed in Table 1, a determination of chirality can be readily obtained using only the additional constraint that the sign of ϕ is known. A simple protocol can be given as follows:

(1) If $|\Delta^1J_{HC}|$ is greater than 6 Hz, the pro-*R* hydrogen has the larger coupling for $\phi < 0$ and the smaller coupling for $\phi > 0$.

(2) Otherwise, the pro-*R* hydrogen is the most upfield shifted resonance for $\phi < 0$ and most downfield shifted for $\phi > 0$.

Although the glycine residues of the two proteins considered represent only a modest subset of the

observed glycine conformations illustrated in Fig. 2, these results suggest that reliable stereochemical assignments can be obtained from the isotope shift and scalar coupling data with only fairly limited additional structural information, without recourse to more involved chiral labeling techniques.

In order to consider what generality this approach offers for conformational analysis at methylene sites, a brief discussion of the physical basis of isotope shifts is warranted. The role of hyperconjugation in isotope shift effects remains controversial (Hansen, 1988; Berger, 1990). These reviews discuss the argument that explaining isotope shift effects via hyperconjugation effects implies a breakdown of the Born–Oppenheimer approximation. This concern does not appear valid for the one-bond isotope shift data presented herein. Both theoretical and experimental studies (e.g. DeFrees et al., 1979; Wiberg et al., 1984; Farnham et al., 1985; Rathna and Chandrasekhar, 1991) have demonstrated bond length variations as a function of the orientation of adjacent lone pair and π orbitals. These geometrically dependent bond length variations are correlated with corresponding variations in the vibrational frequencies (Thomas et al., 1994). As the isotope shift arises from the difference in the zero-point vibrational frequencies for the proton and deuteron, conformationally dependent changes in the vibrational manifold will give rise to differential isotope shifts. As the two geminal glycine H-C bonds in general occupy two different orientations with respect to the adjacent π orbitals, the differential ^{13}C shifts result. Such ab initio results have often been heuristically interpreted in terms of a hyperconjugation interaction between the σ–σ^* orbitals of the H-C bond and an adjacent lone pair or π orbital. Other heuristic models, such as variation in hybridization state, appear substantially less compelling. Earlier correlations (Wesener et al., 1985) of one-bond isotope shifts with fractional s-character indicate a change of 43 ppb on going from sp to sp^3, a change which is only 60% of the range in differential isotope shifts demonstrated here.

In this context it is worthwhile to note that such one-bond differential isotope shifts are not limited to hyperconjugation with π systems. Saturated hydrocarbons have also been argued to exhibit comparable effects, resulting from the periplanar alignment of vicinal H-C bonds (Gunther and Aydin, 1981; Williams, 1986). We note that differential isotope shifts can be observed for various methylene types in protein side chains (LeMaster, D.M. and Kushlan, D.M., manuscript in preparation).

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of General Medical Sciences (GM-38779) and the National Science Foundation (DMB-8957336). The 600 MHz NMR facility used in these studies was funded by the W.M. Keck Foundation, the National Institutes of Health, the National Science Foundation and Northwestern University.

REFERENCES

Berger, S. (1990) *NMR Basic Princ. Prog.*, **22**, 1–29.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **112**, 535–542.

DeFrees, D.J., Taagepera, M., Levi, B.A., Pollack, S.K., Summerhays, K.D., Taft, R.W., Wolfsberg, M. and Hehre, W.J. (1979) *J. Am. Chem. Soc.*, **101**, 5532–5536.

Egli, H. and Von Philipsborn, W. (1981) *Helv. Chim. Acta*, **46**, 976–988.
Farnham, W.B., Smart, B.E., Middleton, W.J., Calabrese, J.C. and Dixon, D.A. (1985) *J. Am. Chem. Soc.*, **107**, 4565–4567.
Gunther, H. and Aydin, R. (1981) *Angew. Chem.*, **93**, 1000–1001.
Hansen, P.E. (1988) *Prog. NMR Spectrosc.*, **20**, 207–255.
Katti, S., LeMaster, D.M. and Eklund, H. (1990) *J. Mol. Biol.*, **212**, 167–184.
LeMaster, D.M. and Kushlan, D.M. (1993) *J. Biomol. NMR*, **3**, 701–708.
Loll, P.J. and Lattman, E.E. (1989) *Protein Struct. Funct. Genet.*, **5**, 183–201.
Mierke, D.F., Grdadolnik, S.G. and Kessler, H. (1992) *J. Am. Chem. Soc.*, **114**, 8283–8284.
Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1989) *Numerical Recipes* (FORTRAN Version), Cambridge University Press, Cambridge, pp. 515–520.
Ramachandran, G.N. and Sasisekharan, V. (1968) *Adv. Protein Chem.*, **23**, 283–437.
Rathna, A. and Chandrasekar, J. (1991) *J. Chem. Soc., Perkins Trans. II*, 1661–1666.
Thomas, H.D., Chen, K. and Allinger, N.L. (1994) *J. Am. Chem. Soc.*, **116**, 5887–5897.
Vuister, G.W., Delaglio, F. and Bax, A. (1992) *J. Am. Chem. Soc.*, **115**, 9674–9675.
Vuister, G.W., Delaglio, F. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 67–80.
Wang, J., Hinck, A.P., Loh, S.N. and Markley, J.L. (1990) *Biochemistry*, **29**, 102–113.
Wesener, J.R., Moskau, D. and Gunther, H. (1985) *J. Am. Chem. Soc.*, **107**, 7307–7311.
Wiberg, K.B., Walters, V. and Colson, S.D. (1984) *J. Phys. Chem.*, **88**, 4723–4728.
Williams, I.H. (1986) *J. Chem. Soc., Chem. Commun.*, 627–628.